



Τεχνολογικό
Πανεπιστήμιο
Κύπρου

Τμήμα Επικοινωνίας και
Σπουδών Διαδικτύου

Διδακτορική διατριβή

ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ

ΣΕ ΗΛΕΚΤΡΟΝΙΚΟΥΣ ΣΥΜΒΟΥΛΟΥΣ ΨΗΦΟΥ

Μαριλένα Αγαθοκλέους

Λεμεσός, 30 Μαΐου 2018

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ
ΣΧΟΛΗ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΜΕΣΩΝ ΕΝΗΜΕΡΩΣΗΣ
ΤΜΗΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΣΠΟΥΔΩΝ ΔΙΑΔΙΚΤΥΟΥ

Διδακτορική διατριβή

ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ
ΣΕ ΗΛΕΚΤΡΟΝΙΚΟΥΣ ΣΥΜΒΟΥΛΟΥΣ ΨΗΦΟΥ
της

Μαριλένας Αγαθοκλέους

Λεμεσός, 30 Μαΐου 2018

Έντυπο έγκρισης

Διδακτορική διατριβή

ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ

ΣΕ ΗΛΕΚΤΡΟΝΙΚΟΥΣ ΣΥΜΒΟΥΛΟΥΣ ΨΗΦΟΥ

Παρουσιάστηκε από τη
Μαριλένα Αγαθοκλέους

Επιβλέπων καθηγητής: Καθηγητής Νικόλας Τσαπατσούλης

Υπογραφή _____

Μέλος επιτροπής: Καθηγητής Ανδρέας Λανίτης

Υπογραφή _____

Μέλος επιτροπής: Καθηγητής Ανδρέας Ευσταθιάδης

Υπογραφή _____

Τεχνολογικό Πανεπιστήμιο Κύπρου
Λεμεσός, 30 Μαΐου 2018

Πνευματικά Δικαιώματα

Copyright © Μαριλενα Αγαθοκλέους, 2018

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της διδακτορικής διατριβής από το Τμήμα Επικοινωνίας και Σπουδών Διαδικτύου του Τεχνολογικού Πανεπιστημίου Κύπρου δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Θα ήθελα να ευχαριστήσω ιδιαίτερα τη γιαγιά μου Ελένη που πρώτη πίστεψε σε εμένα! Επιπρόσθετα θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου για τις πολύτιμες συμβουλές του, καθώς επίσης την ακαδημαϊκή κοινότητα αλλά και την οικογένεια μου για τη στήριξη που μου παρείχαν. Αφιερωμένο στα παιδιά μου Αλέξανδρο και Μαρία και στο σύζυγο μου Λούκα για την υπομονή που έδειξε όλα αυτά τα χρόνια.

Περίληψη

Τα Συστήματα Συστάσεων (ΣΣ) είναι μέθοδοι και εργαλεία λογισμικού που κάνουν προτάσεις για προϊόντα και υπηρεσίες προς πιθανούς ενδιαφερόμενους χρήστες. Οι Ηλεκτρονικοί Σύμβουλοι Ψήφου (Voting Advice Applications - VAAs) είναι online εργαλεία, τα οποία συγκρίνουν τις πολιτικές απόψεις των χρηστών με τις πολιτικές θέσεις των πολιτικών κομμάτων ή υποψηφίων και, έμμεσα, προτείνουν στο χρήστη τον υποψήφιο ή το κόμμα που έχει όμοιες απόψεις μαζί του. Η παρούσα διατριβή καταπάνεται με τη χρήση τεχνικών και μεθοδολογιών από τα Συστήματα Συστάσεων στην πανευρωπαϊκή εφαρμογή ηλεκτρονικού συμβούλου ψήφου “EUvox”¹, για τη διερεύνηση συγκεκριμένων ερευνητικών ερωτημάτων. Αρχικά γίνεται εισαγωγή στο θέμα, στο σκοπό και στους στόχους της διατριβής. Ακολουθεί αναφορά στην υπάρχουσα βιβλιογραφία σχετικά με τα ΣΣ και τα VAAs. Η υπόλοιπη εργασία παρουσιάζει τη μεθοδολογία που χρησιμοποιήθηκε και τα πειράματα με τη σειρά που διεξήχθησαν για να απαντήσουν τα ερευνητικά ερωτήματα. Γίνεται σχολιασμός των αποτελεσμάτων και σύγκριση με την υπάρχουσα έρευνα.

Τα αποτελέσματα έδειξαν μεταξύ άλλων πως οι μη γραμμικές μέθοδοι αποδίδουν καλύτερα από άλλες τεχνικές μηχανικής μάθησης, καθώς επίσης η χρήση συμπληρωματικών ερωτήσεων αυξάνει την απόδοση της πρόβλεψης ψήφου. Τεχνικές παραγοντοποίησης πινάκων μπορούν να εφαρμοστούν με μεγάλη επιτυχία για την εκτίμηση των συμπληρωματικών ερωτήσεων που δεν απαντήθηκαν από τους χρήστες. Υπάρχει ανομοιογένεια εντός των κομμάτων, ενώ παρατηρούνται πολλοί χρήστες να δίνουν όμοιες απαντήσεις με άλλους χρήστες που υποστηρίζουν διαφορετικά κόμματα. Όσον αφορά την ταυτότητα του Ευρωπαϊού ψηφοφόρου, φαίνεται να είναι άντρας, νέος, με υψηλή μόρφωση και μεγάλο ενδιαφέρον για την πολιτική. Οι άντρες χρήστες που είναι μορφωμένοι τείνουν να ενδιαφέρονται περισσότερο για την πολιτική από ότι οι γυναίκες ή αυτοί με χαμηλό μορφωτικό επίπεδο. Καθώς επίσης η ηλικιακή ομάδα φαίνεται να μην αποτελεί σημαντικό παράγοντα για το ενδιαφέρον του χρήστη προς την πολιτική.

Λέξεις κλειδιά: Συστήματα Συστάσεων, Ηλεκτρονικοί Σύμβουλοι Ψήφου, Συνεργατικό Φιλτράρισμα, Μηχανική Μάθηση, Μηχανές Διανυσμάτων Υποστήριξης

¹<http://www.euvox2014.eu/>

Abstract

Recommender Systems (RSs) are software tools and techniques that make suggestions for products and services to potentially interested users. Voting Advice Applications (VAAs) are online tools that compare the political views of users with the political positions of political parties or candidates and indirectly suggest to the user the candidate or party with similar views. This thesis uses techniques and methods from RSs to the EU-wide VAA “EUvox”² to explore specific research questions. At the beginning, the subject, purpose and objectives of the thesis are introduced. Then, reference is made to the existing literature on RSs and VAAs. The rest of the work presents the methodology used and the experiments conducted in order to answer the research questions. The results of the experiments are compared to the existing research.

The results showed, among other things, that non-linear methods perform better than other machine learning techniques, and the use of supplementary questions increases the performance of the voting prediction. Matrix factorization techniques can be implemented with great success in estimating supplementary questions that were not answered by users. There is heterogeneity within the parties, and many users give similar responses with other users who support different parties. As far as the identity of the European voter is concerned, he seems to be a man, young, highly educated and very interested in politics. Male educated users tend to be more interested in politics than women or those with a low level of education. As well as the age group does not seem to be an important factor for the user’s interest in politics.

Keywords: Recommender Systems, Voting Advice Applications, Collaborative Filtering, Machine Learning, Support Vector Machines

²<http://www.euvox2014.eu/>

Περιεχόμενα

Περίληψη	vii
Abstract	ix
Περιεχόμενα	x
Κατάλογος σχημάτων	xii
Κατάλογος πινάκων	xvii
1 Εισαγωγή και Υπόβαθρο	1
1.1 Εισαγωγή	1
1.1.1 Σκοπός	2
1.1.2 Ερευνητικά Ερωτήματα	3
1.1.3 Συνεισφορά	3
1.1.4 Δομή της Διατριβής	4
1.2 Θεωρητικό Υπόβαθρο	5
1.3 Συστήματα Συστάσεων	6
1.3.1 Η λειτουργία των Συστημάτων Συστάσεων	6
1.3.2 Εξόρυξη Δεδομένων και Μηχανική Μάθηση	7
1.3.3 Προσεγγίσεις Συστημάτων Συστάσεων	8
1.3.4 Συνεργατικό Φιλτράρισμα (Collaborative Filtering)	9
1.3.5 Εφαρμογές των Συστημάτων Συστάσεων	11
1.4 Ηλεκτρονικοί Σύμβουλοι Ψήφου	11
1.4.1 Η “ηλεκτρονική δημοκρατία”	11
1.4.2 Το αντίκτυπο των Ηλεκτρονικών Συμβούλων Ψήφου	12
1.5 Σχεδιασμός Ηλεκτρονικών Συμβούλων Ψήφου	14
1.5.1 Επιλογή ερωτήσεων	14
1.5.2 Κωδικοποίηση των θέσεων των κομμάτων	17
1.5.3 Δημιουργία της πρότασης ψήφου	19
1.6 Η ποιότητα και ο “καθαρισμός” των δεδομένων των VAAs	25
1.7 Μεθοδολογία Έρευνας	26
1.7.1 Μαθηματικό Υπόβαθρο	29
1.7.2 Εκπαίδευση, δοκιμή και αξιολόγηση	36
1.7.3 Δεδομένα	37
2 Χρήση Μηχανικής Μάθησης στα VAAs	39
2.1 Δημιουργία της σύστασης ψήφου με βάση τις ερωτήσεις του VAA	39
2.1.1 Οι παράγοντες που επηρεάζουν τη δημιουργία της σύστασης ψήφου	39
2.1.2 Χρήση τεχνικών μηχανικής μάθησης στο VAA	42
2.1.3 Η απόδοση της πρόβλεψης ψήφου με χρήση μόνο των απαντήσεων στις δηλώσεις πολιτικής	43
2.2 Ενίσχυση της σύστασης ψήφου με χρήση των συμπληρωματικών ερωτήσεων	53
2.2.1 Συμπληρωματικές ερωτήσεις	54
2.2.2 Η πρόβλεψη της πρόθεσης ψήφου συνδυάζοντας τις δηλώσεις πολιτικής μαζί με κάποιες από τις συμπληρωματικές ερωτήσεις	59
2.3 Εκτίμηση των απαντήσεων στις συμπληρωματικές ερωτήσεις του VAA	66
2.3.1 Τεχνικές για τον υπολογισμό των ελλειπουσών τιμών	66

2.3.2	Η εκτίμηση των ελλειπουσών τιμών των χρηστών	68
3	Πολιτικές - Κοινωνικές επιπτώσεις των VAAs	73
3.1	Ιδεολογική ταύτιση κόμματος - χρήστη	73
3.1.1	Αναζήτηση της ιδεολογικής ταύτισης	74
3.2	Το ενδιαφέρον του Ευρωπαίου χρήστη των VAAs για την πολιτική	91
3.2.1	Το ενδιαφέρον του χρήστη για την πολιτική, σύμφωνα με τα δημογραφικά του χαρακτηριστικά	92
3.2.2	Τι δηλώνουν τα κριτήρια ψήφου του χρήστη για το ενδιαφέρον του για την πολιτική	94
3.3	Το πρόβλημα σύστασης στα VAAs ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλών κριτηρίων	97
3.3.1	Η απόδοση του VAA με την ομαδοποίηση των δηλώσεων πολιτικής του VAA	98
4	Συζήτηση	103
4.1	Σχολιασμός Αποτελεσμάτων	103
4.1.1	Η βελτιστοποίηση της πρόβλεψης ψήφου	103
4.1.2	Η ταυτότητα και συμπεριφορά του Ευρωπαίου χρήστη στο EUnox	108
4.2	Συμπεράσματα	109
4.3	Μελλοντική Έρευνα	110
4.4	Επίλογος	111
	Βιβλιογραφία	115
I	Τα ονόματα των κομμάτων σε κάθε χώρα	123
II	Αναλυτικοί πίνακες αποτελεσμάτων	127
III	Η κατανομή του δείγματος	149
IV	Αποτελέσματα των συνδυασμών μεταξύ των δηλώσεων πολιτικής και των συμπληρωματικών ερωτήσεων	159
V	Αποτελέσματα από τη διερεύνηση της ιδεολογικής ταύτισης κόμματος-χρήστη	185
VI	Αποτελέσματα από τη διερεύνηση της ιδεολογικής ταύτισης κόμματος-χρήστη - Μέρος Β	211
VII	Αποτελέσματα από τη διερεύνηση της ιδεολογικής ταύτισης κόμματος-χρήστη - Μέρος Γ	237
	Ευρετήριο	265

Κατάλογος σχημάτων

1.1	Παράδειγμα δήλωσης πολιτικής από το <i>EUVOx</i> 2014	16
1.2	Παράδειγμα συμπληρωματικών ερωτήσεων δημογραφικού, κυρίως, περιεχομένου από το <i>EUVOx</i> 2014	17
1.3	Παράδειγμα της γραφικής απεικόνισης των αποτελεσμάτων από το <i>Kieskompas</i> για τις ομοσπονδιακές εκλογές της 24ης Σεπτεμβρίου 2017 στη Γερμανία. Η σταγόνα με την ενσωματωμένη κουκκίδα αναπαριστά την τοποθέτηση του χρήστη από το VAA με βάση τις δηλώσεις του στις δηλώσεις πολιτικής	20
1.4	Παράδειγμα της γραφικής απεικόνισης των αποτελεσμάτων του <i>StemWijzer</i> για τις κοινοβουλευτικές εκλογές της 15ης Μαρτίου 2017 στην Ολλανδία	21
1.5	Παράδειγμα της γραφικής απεικόνισης των αποτελεσμάτων από το <i>Smartvote</i> με τη χρήση του “διαγράμματος αράχνης” για τις κοινοτικές εκλογές της 24ης Σεπτεμβρίου 2017 στο Κόνιζ της Ελβετίας	24
1.6	Η συμπληρωματική ερώτηση για την πρόθεση ψήφου όπως εμφανίζεται στο <i>EUVOx</i> 2014	25
1.7	Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης για την πρόβλεψη της πρόθεσης ψήφου	35
2.1	Μια δήλωση πολιτικής από το <i>EUVOx</i> 2014 μαζί με τις επιλογές που δίνονται στο χρήστη για να απαντήσει	40
2.2	Ταξινόμηση των κομμάτων βάσει της ομοιότητας μεταξύ κόμματος και χρήστη, όπως υπολογίστηκε με την παραδοσιακή προσέγγιση των VAAs, <i>EUVOx</i> 2014	41
2.3	Ταξινόμηση των κομμάτων βάσει της ομοιότητας μεταξύ του μοντέλου κόμματος και των απαντήσεων του χρήστη, όπως υπολογίστηκε με την κοινωνική προσέγγιση των VAAs, <i>EUVOx</i> 2014	41
2.4	Επιπρόσθετες ερωτήσεις που υπάρχουν στο <i>EUVOx</i> 2014	54
2.5	Ο τρόπος που εμφανίζεται η ερώτηση για την αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη στο <i>EUVOx</i> 2014 σε σχέση με την κλίμακα “λιγότερη ΕΕ - περισσότερη ΕΕ”	57
2.6	Αναπαράσταση της μεθόδου SVD	67
3.1	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Γαλλία (β) Γερμανία (γ) Ουγγαρία	75
3.2	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (7 ομάδες)	76
3.3	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (11 ομάδες)	77
3.4	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (14 ομάδες)	77
3.5	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (21 ομάδες)	78
3.6	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 7 ομάδες	79
3.7	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 11 ομάδες	79
3.8	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 14 ομάδες	79
3.9	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 21 ομάδες	80
3.10	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 7 ομάδες	81
3.11	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 11 ομάδες	82
3.12	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 14 ομάδες	82
3.13	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 21 ομάδες	83

3.14	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 7 ομάδες	84
3.15	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 11 ομάδες	84
3.16	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 14 ομάδες	85
3.17	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 21 ομάδες	86
3.18	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 5 ομάδες	87
3.19	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 8 ομάδες	87
3.20	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 10 ομάδες	88
3.21	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 15 ομάδες	88
3.22	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 5 ομάδες	89
3.23	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 8 ομάδες	90
3.24	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 10 ομάδες	90
3.25	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 15 ομάδες	90
III.1	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Αυστρία (β) Βουλγαρία (γ) Κύπρος . . .	150
III.2	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Τσεχία (β) Δανία (γ) Εσθονία	151
III.3	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Γαλλία (β) Γερμανία (γ) Ουγγαρία . . .	152
III.4	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Αγγλία (β) Β.Ιρλανδία (γ) Σκωτία . . .	153
III.5	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ουαλία (β) Ισπανία (γ) Φινλανδία . . .	154
III.6	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ελλάδα (β) Κροατία (γ) Ιρλανδία . . .	155
III.7	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ιταλία (β) Λιθουανία (γ) Λετονία . . .	156
III.8	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ολλανδία (β) Πολωνία (γ) Ρουμανία . .	157
III.9	Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Πορτογαλία (β) Σουηδία (γ) Σλοβακία .	158
V.1	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Αυστρία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες	186
V.2	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 8 ομάδες	187
V.3	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 12 ομάδες	187
V.4	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 16 ομάδες	187
V.5	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 24 ομάδες	188
V.6	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Βουλγαρία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες	189
V.7	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 5 ομάδες	190
V.8	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 8 ομάδες	190
V.9	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 10 ομάδες	190
V.10	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 15 ομάδες	191
V.11	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Κύπρος (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	192
V.12	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 6 ομάδες	193
V.13	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 9 ομάδες	193
V.14	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 12 ομάδες	193
V.15	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 18 ομάδες	194
V.16	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Τσεχία (α) 10 ομάδες (β) 15 ομάδες (γ) 20 ομάδες (δ) 30 ομάδες	195
V.17	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 10 ομάδες	196
V.18	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 15 ομάδες	196
V.19	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 20 ομάδες	196
V.20	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 30 ομάδες	197

V.21	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Δανία (α) 7 ομάδες (β) 11 ομάδες (γ) 14 ομάδες (δ) 21 ομάδες	198
V.22	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 7 ομάδες	199
V.23	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 11 ομάδες	199
V.24	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 14 ομάδες	199
V.25	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 21 ομάδες	200
V.26	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Εσθονία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	201
V.27	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 6 ομάδες	202
V.28	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 9 ομάδες	202
V.29	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 12 ομάδες	202
V.30	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 18 ομάδες	203
V.31	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες	204
V.32	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 5 ομάδες	205
V.33	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 8 ομάδες	205
V.34	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 10 ομάδες	205
V.35	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 15 ομάδες	206
V.36	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	207
V.37	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 6 ομάδες	208
V.38	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 9 ομάδες	208
V.39	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 12 ομάδες	208
V.40	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 18 ομάδες	209
VI.1	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	212
VI.2	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 6 ομάδες	213
VI.3	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 9 ομάδες	213
VI.4	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 12 ομάδες	213
VI.5	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 18 ομάδες	214
VI.6	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	215
VI.7	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 6 ομάδες	216

VI.8	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 9 ομάδες	216
VI.9	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 12 ομάδες	216
VI.10	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 18 ομάδες	217
VI.11	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ισπανία (α) 12 ομάδες (β) 18 ομάδες (γ) 24 ομάδες (δ) 36 ομάδες	218
VI.12	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 12 ομάδες	219
VI.13	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 18 ομάδες	220
VI.14	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 24 ομάδες	221
VI.15	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 36 ομάδες	222
VI.16	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Φινλανδία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες	223
VI.17	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 8 ομάδες	224
VI.18	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 12 ομάδες	224
VI.19	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 16 ομάδες	224
VI.20	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 24 ομάδες	225
VI.21	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ελλάδα (α) 9 ομάδες (β) 14 ομάδες (γ) 18 ομάδες (δ) 27 ομάδες	226
VI.22	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 9 ομάδες	227
VI.23	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 14 ομάδες	227
VI.24	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 18 ομάδες	228
VI.25	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 27 ομάδες	229
VI.26	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Κροατία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες	230
VI.27	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 5 ομάδες	231
VI.28	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 8 ομάδες	231
VI.29	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 10 ομάδες	231
VI.30	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 15 ομάδες	232
VI.31	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ιρλανδία (α) 7 ομάδες (β) 11 ομάδες (γ) 14 ομάδες (δ) 21 ομάδες	233
VI.32	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 7 ομάδες	234
VI.33	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 11 ομάδες	234
VI.34	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 14 ομάδες	234
VI.35	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 21 ομάδες	235
VII.1	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ιταλία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες	238
VII.2	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 8 ομάδες	239
VII.3	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 12 ομάδες	239
VII.4	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 16 ομάδες	239
VII.5	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 24 ομάδες	240
VII.6	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Λιθουανία (α) 4 ομάδες (β) 6 ομάδες (γ) 8 ομάδες (δ) 12 ομάδες	241
VII.7	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 4 ομάδες	242
VII.8	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 6 ομάδες	242

VII.9	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 8 ομάδες	242
VII.10	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 12 ομάδες	243
VII.11	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Λετονία (α) 3 ομάδες (β) 5 ομάδες (γ) 6 ομάδες (δ) 9 ομάδες	244
VII.12	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 3 ομάδες	245
VII.13	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 5 ομάδες	245
VII.14	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 6 ομάδες	245
VII.15	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 9 ομάδες	245
VII.16	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ολλανδία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	246
VII.17	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 6 ομάδες	247
VII.18	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 9 ομάδες	247
VII.19	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 12 ομάδες	247
VII.20	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 18 ομάδες	248
VII.21	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Πολωνία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες	249
VII.22	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 8 ομάδες	250
VII.23	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 12 ομάδες	250
VII.24	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 16 ομάδες	250
VII.25	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 24 ομάδες	251
VII.26	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Πορτογαλία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες	252
VII.27	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 6 ομάδες	253
VII.28	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 9 ομάδες	253
VII.29	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 12 ομάδες	253
VII.30	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 18 ομάδες	254
VII.31	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ρουμανία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες	255
VII.32	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ρουμανία – 5 ομάδες	256
VII.33	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ρουμανία – 8 ομάδες	256
VII.34	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ρουμανία – 10 ομάδες	256
VII.35	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ρουμανία – 15 ομάδες	257
VII.36	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Σουηδία (α) 10 ομάδες (β) 15 ομάδες (γ) 20 ομάδες (δ) 30 ομάδες	258
VII.37	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 10 ομάδες	259
VII.38	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 15 ομάδες	259
VII.39	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 20 ομάδες	260
VII.40	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 30 ομάδες	261
VII.41	Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Σλοβακία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες	262
VII.42	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 8 ομάδες	263
VII.43	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 12 ομάδες	263
VII.44	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 16 ομάδες	263
VII.45	Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 24 ομάδες	264

Κατάλογος πινάκων

1.1	Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει της City-block (Manhattan) απόστασης	20
1.2	Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει της Ευκλείδειας απόστασης	21
1.3	Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει της θεωρίας της εκλογικής συμπεριφοράς των Rabinowitz & MacDonald [2]	22
1.4	Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει του συνδυασμού της City-block (Manhattan) απόστασης της και της θεωρίας της εκλογικής συμπεριφοράς των Rabinowitz & MacDonald [2]	22
1.5	Δεδομένα	38
2.1	Τα συνολικά αποτελέσματα των ταξινομητών για Αυστρία, Βουλγαρία, Κύπρο και Τσεχία	44
2.2	Τα συνολικά αποτελέσματα των ταξινομητών για Γερμανία, Δανία, Εσθονία και Ηνωμένο Βασίλειο-Αγγλία	45
2.3	Τα συνολικά αποτελέσματα των ταξινομητών για Ηνωμένο Βασίλειο-Βόρεια Ιρλανδία, Ηνωμένο Βασίλειο-Σκωτία, Ηνωμένο Βασίλειο-Ουαλία και Ισπανία	48
2.4	Τα συνολικά αποτελέσματα των ταξινομητών για Φινλανδία, Γαλλία, Ελλάδα και Κροατία	49
2.5	Τα συνολικά αποτελέσματα των ταξινομητών για Ουγγαρία, Ιρλανδία, Ιταλία και Λιθουανία	50
2.6	Τα συνολικά αποτελέσματα των ταξινομητών για Λετονία, Ολλανδία, Πολωνία και Πορτογαλία	51
2.7	Τα συνολικά αποτελέσματα των ταξινομητών για Ρουμανία, Σουηδία και Σλοβακία	52
2.8	Κατανομή με βάση το φύλο σε κάθε σύνολο δεδομένων	55
2.9	Κατανομή των χρηστών των VAAs ανά ηλικιακή ομάδα για κάθε σύνολο δεδομένων	56
2.10	Κατανομή των χρηστών των VAAs ανά μορφωτικό επίπεδο για κάθε σύνολο δεδομένων	58
2.11	Τα κριτήρια με τα οποία ψηφίζουν οι χρήστες των VAAs σε κάθε χώρα	59
2.12	Το ενδιαφέρον των χρηστών των VAAs για την πολιτική σε κάθε χώρα	61
2.13	Το ποσοστό των χρηστών που απάντησαν σε όλες τις συμπληρωματικές ερωτήσεις για κάθε χώρα . .	62
2.14	Η μέση αύξηση στην απόδοση της πρόβλεψης ψήφου που παρατηρήθηκε για όλες τις χώρες ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις . . .	63
2.15	Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Φινλανδία	64
2.16	Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Γαλλία	65
2.17	Το ποσοστό των ελλειπουσών τιμών στο δείγμα	69
2.18	Η απόδοση της SVD για την εκτίμηση των ελλειπουσών τιμών των χρηστών	70
2.19	Σύγκριση μεταξύ των αποτελεσμάτων στην Ενότητα 2.1 και 2.2	71
3.1	Η εξίσωση της γραμμικής παλινδρόμησης για κάθε χώρα (η σχέση μεταξύ δημογραφικών χαρακτηριστικών και ενδιαφέροντος για την πολιτική)	93

3.2	Η εξίσωση της γραμμικής παλινδρόμησης για κάθε χώρα (η σχέση μεταξύ κριτηρίων ψήφου και ενδιαφέροντος για την πολιτική)	95
3.3	Η θεματική ενότητα για κάθε σύνολο ερωτήσεων	99
3.4	Η απόδοση της πρόβλεψης ψήφου ομαδοποιώντας τις δηλώσεις πολιτικής σε κριτήρια εισόδου για κάθε χώρα ξεχωριστά – Μέθοδος Α	100
3.5	Η απόδοση της πρόβλεψης ψήφου ομαδοποιώντας τις δηλώσεις πολιτικής σε κριτήρια εισόδου για κάθε χώρα ξεχωριστά – Μέθοδος Β	101
I.1	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Α	123
I.2	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Β	124
I.3	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Γ	124
I.4	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Δ	125
I.5	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Ε	125
I.6	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Στ	126
I.7	Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Η	126
II.1	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Αυστρία	127
II.2	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Βουλγαρία	128
II.3	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Κύπρος	128
II.4	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Τσεχία	129
II.5	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Γερμανία	130
II.6	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Δανία	131
II.7	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Εσθονία	132
II.8	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Αγγλία	132
II.9	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Βόρεια Ιρλανδία	133
II.10	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Σκωτία	134
II.11	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ουαλία	135
II.12	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ισπανία	136
II.13	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Φινλανδία	137
II.14	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Γαλλία	138
II.15	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ελλάδα	139
II.16	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Κροατία	140
II.17	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ουγγαρία	140
II.18	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ιρλανδία	141
II.19	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ιταλία	142
II.20	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Λιθουανία	143
II.21	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Λετονία	143
II.22	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ολλανδία	144
II.23	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Πολωνία	145
II.24	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Πορτογαλία	146
II.25	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ρουμανία	146
II.26	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Σουηδία	147
II.27	Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Σλοβακία	148
IV.1	Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Αυστρία	160

IV.25	Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Σλοβακία	184
-------	---	-----

Κεφάλαιο 1

Εισαγωγή και Υπόβαθρο

1.1 Εισαγωγή

Ο Παγκόσμιος Ιστός (World-Wide Web) είναι μια τεράστια συλλογή από ψηφιακά έγγραφα που με τη χρήση διαθέσιμης τεχνολογίας παρέχει την ευκαιρία στο χρήστη να ψάξει και να ενημερωθεί για το θέμα που επιθυμεί μέσα σε μόλις λίγα δευτερόλεπτα [3]. Σήμερα διανύει τη δεύτερη γενιά του (Web 2.0) και εκτός από την αναζήτηση πληροφοριών επιτρέπει και την αλληλεπίδραση των χρηστών, οι οποίοι συνεργάζονται, επικοινωνούν και ανταλλάσσουν πληροφορίες. Αυτό είχε ως αποτέλεσμα να αυξηθεί δραματικά ο όγκος των ψηφιακών δεδομένων στον παγκόσμιο ιστό και να συνεχίζει να μεγαλώνει, δυσκολεύοντας τους χρήστες όταν αναζητούν συγκεκριμένα ψηφιακά έγγραφα ή/και υπηρεσίες. Απάντηση στο πρόβλημα αυτό προσπαθούν να δώσουν τα Συστήματα Συστάσεων (ΣΣ), τα οποία με τη χρήση ειδικών αλγορίθμων βοηθούν τους χρήστες να προσδιορίσουν τι είναι αυτό που χρειάζονται μέσα από ένα μεγάλο όγκο αγαθών ή υπηρεσιών.

Ένας από τους πιο δημοφιλείς τρόπους σύστασης που χρησιμοποιείται από τα ΣΣ είναι το Συνεργατικό Φιλτράρισμα [4][5] (pp.1-35), σύμφωνα με το οποίο οι χρήστες που υπάρχουν στο σύστημα συγκρίνονται μεταξύ τους και προτείνονται σε ένα συγκεκριμένο άτομο αντικείμενα ή υπηρεσίες που δεν έχει δει ή δοκιμάσει ακόμη, αλλά άφησαν ικανοποιημένους χρήστες, οι οποίοι έχουν παρόμοια γούστα με αυτόν. Η σύσταση σε αυτή την περίπτωση συνήθως στηρίζεται σε ένα και μόνο κριτήριο, τη συνολική βαθμολογία που δίνει ο χρήστης και αντικατοπτρίζει τη συνολική του εντύπωση για το εν λόγω προϊόν ή υπηρεσία. Υπάρχουν όμως και οι περιπτώσεις που ένα πρόβλημα σύστασης μπορεί να αντιμετωπίζεται ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων, όπου εκτός από τη συνολική βαθμολογία υπολογίζονται και οι βαθμολογίες που έδωσαν οι χρήστες σε επιμέρους κριτήρια που αφορούν χαρακτηριστικά του προϊόντος ή της υπηρεσίας.

Τα ΣΣ έχουν εφαρμοστεί ευρέως σε τομείς όπως το ηλεκτρονικό εμπόριο, η ψυχαγωγία και οι διαδικτυακές υπηρεσίες. Για την εκπόνηση της παρούσας διατριβής μελετούμε την εφαρμογή των ΣΣ σε Ηλεκτρονικούς Συμβούλους Ψήφου, οι οποίοι προτείνουν στους χρήστες έναν υποψήφιο ή ένα υποψήφιο κόμμα από επερχόμενες εκλογές. Οι Ηλεκτρονικοί Σύμβουλοι Ψήφου (Voting Advice Applications-VAAs) είναι εφαρμογές του διαδικτύου που καλούν τους χρήστες να τοποθετήσουν την άποψη τους σε μια σειρά από ερωτήσεις που αφορούν σημαντικά πολιτικά, οικονομικά και κοινωνικά ζητήματα και βρίσκονται online στην ιστοσελίδα του VAA. Στη συνέχεια συγκρίνουν τις πολιτικές απόψεις των χρηστών με αυτές των κομμάτων ή υποψηφίων, συστήνοντας τελικά στον κάθε χρήστη, τον υποψήφιο ή το κόμμα που του ταιριάζει καλύτερα. Με αυτόν τον τρόπο βοηθούν έναν ψηφοφόρο να καταλήξει ή να επιβεβαιωθεί για την ψήφο του. Ενώ τα παραδοσιακά VAAs βασίζονται στις δηλωμένες ή ακαδημαϊκά κωδικοποι-

ημένες θέσεις των κομμάτων/υποψηφίων, μια πρόσφατη καινοτομία έχει εισαγάγει την «κοινωνική» σύσταση ψήφου που στηρίζεται στις βασικές αρχές του συνεργατικού φιλτραρίσματος. Τα κοινωνικά VAAs (ή διαφορετικά τα SVAAAs-Social Voting Advice Applications) συγκρίνουν τις πολιτικές θέσεις των χρηστών και προτείνουν σε έναν χρήστη, το πολιτικό κόμμα/υποψήφιο που δήλωσε πως προτίθεται να ψηφίσει η πλειοψηφία των χρηστών με κοντινές πολιτικές ιδέες με αυτόν.

1.1.1 Σκοπός

Η συγκεκριμένη διατριβή επικεντρώνεται στη χρήση τεχνικών και μεθοδολογιών από τα ΣΣ για τη βελτιστοποίηση της πρόβλεψης ψήφου που γίνεται από τα VAAs και τη διερεύνηση της συμπεριφοράς του Ευρωπαίου χρήστη.

Η παρούσα εργασία διεξάγεται με σκοπό να μελετήσει συγκεκριμένα ερευνητικά ερωτήματα που όλα μαζί θα δημιουργήσουν το προφίλ του Ευρωπαίου Ψηφοφόρου. Αφού γίνει επεξεργασία των δεδομένων από το EUvox (την πανευρωπαϊκή εφαρμογή συμβούλου ψήφου για τις εκλογές του 2014 για το Ευρωπαϊκό Κοινοβούλιο), ακολουθούν διάφορες πειραματικές διαδικασίες, οι οποίες αναλύονται σε ενότητες και αποσκοπούν στη μελέτη διάφορων ζητημάτων που προκύπτουν μέσα από δύο μεγάλα θέματα.

Το πρώτο θέμα σχετίζεται με τον τρόπο που αντιδρούν οι Ευρωπαίοι χρήστες στο περιβάλλον του VAA. Συγκεκριμένα διερευνάται κατά πόσο ο λόγος που επιλέγει ένας χρήστης να ψηφίσει ένα κόμμα δηλώνει και το ενδιαφέρον του για την πολιτική, εάν ψηφίζει, για παράδειγμα, το κόμμα που προτιμούν η οικογένεια και οι φίλοι του ή υποστηρίζει αυτό που ταιριάζει καλύτερα με την ιδεολογία του. Επίσης υπολογίζεται η σχέση μεταξύ των δημογραφικών χαρακτηριστικών ενός χρήστη και του ενδιαφέροντος του για την πολιτική καθώς επίσης το πόσο πολύ ταυτίζεται με τις πολιτικές θέσεις των κομμάτων της χώρας του.

Για πρώτη φορά μελετώνται τα ιδεολογικά χαρακτηριστικά των κομμάτων με βάση τους ψηφοφόρους τους και εάν υπάρχει ποικιλομορφία στα κόμματα. Όπως επίσης για πρώτη φορά χρησιμοποιούνται τεχνικές παραγοντοποίησης για να υπολογίσουν τις ελλείπουσες τιμές του χρήστη σε συμπληρωματικές ερωτήσεις βάσει των δηλώσεων πολιτικής που έχει καταχωρήσει στο σύστημα. Επιπρόσθετα εξετάζεται κατά πόσο το ενδιαφέρον του χρήστη για την πολιτική καθώς και άλλα στοιχεία, όπως είναι για παράδειγμα ο τρόπος που τοποθετεί τον εαυτό του στον πολιτικό χάρτη και τα κριτήρια ψήφου, επηρεάζουν τον τρόπο με τον οποίο καταλήγει στο κόμμα της προτίμησης του. Ακόμη διερευνάται αν ο Ευρωπαίος χρήστης διαφέρει από χώρα σε χώρα. Επιπλέον μελετάται κατά πόσο τα κοινωνικά VAAs είναι αποτελεσματικότερα από τα παραδοσιακά, κατά πόσο αυτό αλλάζει ανάλογα με τη χώρα που εξετάζεται και σε ποιους λόγους πιθανόν να ευθύνεται.

Η δεύτερη θεματική ενότητα επικεντρώνεται στους τρόπους με τους οποίους μπορεί να βελτιστοποιηθεί η απόδοση ενός VAA. Διεξάγονται πειράματα που μελετούν το ενδεχόμενο οι απαντήσεις των χρηστών σε συμπληρωματικές ερωτήσεις, όταν λαμβάνονται υπόψη μαζί με τις δηλώσεις πολιτικής, να οδηγούν σε μια πιο αποτελεσματική σύσταση ψήφου. Επίσης, μέσα από αυτήν την πειραματική διαδικασία, μπορούν να παρατηρηθούν οι συμπληρωματικές ερωτήσεις που βοηθούν περισσότερο στην πρόβλεψη της πρόθεσης ψήφου και κατά πόσο αυτό διαφέρει από χώρα σε χώρα. Επιπρόσθετα γίνεται αναφορά σε τεχνικές μηχανικής μάθησης που μπορούν να εφαρμοστούν για τη βελτιστοποίηση της απόδοσης της πρόβλεψης ψήφου. Ακόμη διερευνάται κατά πόσο οι δηλώσεις πολιτικής του VAA μπορούν να ομαδοποιηθούν κάτω από τρεις κατηγορίες με συγκεκριμένη θεματολογία, έτσι ώστε το σύστημα στην είσοδο να λαμβάνει μόνο τρεις τιμές.

1.1.2 Ερευνητικά Ερωτήματα

Κατά την εκπόνηση της διατριβής διερευνήθηκαν τα ακόλουθα ερευνητικά ερωτήματα:

1. Πόσο καλά μπορεί να προβλεφθεί η πρόθεση ψήφου ενός χρήστη του VAA, γνωρίζοντας μόνο τις απαντήσεις που έδωσε στις δηλώσεις πολιτικής; Σύγκριση σε πανευρωπαϊκό επίπεδο.
2. Μπορεί να ενισχυθεί η απόδοση της πρόβλεψης ψήφου, όταν μαζί με τις δηλώσεις πολιτικής λαμβάνονται υπόψη και κάποιες συμπληρωματικές ερωτήσεις;
3. Μπορεί να γίνει εκτίμηση των απαντήσεων στις συμπληρωματικές ερωτήσεις του VAA, χρησιμοποιώντας τις δηλώσεις πολιτικής;
4. Υπάρχει ιδεολογική ταύτιση μεταξύ κόμματος και χρήστη;
5. Μπορούν να εξαχθούν συμπεράσματα για το ενδιαφέρον του Ευρωπαίου χρήστη του VAA για την πολιτική, ανάλογα με τα κριτήρια ψήφου και τα δημογραφικά του στοιχεία;
6. Μπορούν οι απαντήσεις στο VAA ερωτηματολόγιο να χωριστούν σε κατηγορίες - κριτήρια, έτσι ώστε για την πρόβλεψη ψήφου να χρησιμοποιούνται μόνο οι συνολικές βαθμολογίες που προκύπτουν από το άθροισμα των ερωτήσεων σε κάθε κατηγορία - κριτήριο;

1.1.3 Συνεισφορά

Το ερευνητικό πεδίο γύρω από τα VAAs κυριαρχείται σε μεγάλο βαθμό από προσεγγίσεις της πολιτικής επιστήμης. Ωστόσο, ερευνητές από διάφορους κλάδους ασχολούνται με προβλήματα που πηγάζουν από τα VAAs. Για παράδειγμα, τα SVAAAs (κοινωνική σύσταση των VAAs) είναι ένας συνδυασμός μεταξύ των κοινωνικών επιστημών και των προσεγγίσεων της επιστήμης των υπολογιστών.

Η παρούσα διατριβή προσεγγίζει το θέμα της σύστασης ψήφου από την πλευρά της τεχνητής νοημοσύνης και ασχολείται με ζητήματα σχετικά με το προφίλ του Ευρωπαίου Ψηφοφόρου. Τα ζητήματα αυτά συνοψίζονται ως ακολούθως:

1. Πόσο καλά μπορεί να προβλεφθεί η πρόθεση ψήφου ενός χρήστη του VAA, γνωρίζοντας μόνο τις απαντήσεις που έδωσε στις ερωτήσεις του ερωτηματολογίου; Είναι αρκετές οι απαντήσεις των χρηστών στο ερωτηματολόγιο ή χρειάζεται και κάποια επιπλέον γνώση έτσι ώστε να βελτιωθεί η απόδοση του συστήματος; Οι απαντήσεις στο ερώτημα αυτό ποικίλουν ανάλογα με τη χώρα. Σε κάθε περίπτωση, όμως, η επιπλέον γνώση βοηθά το σύστημα να “γνωρίσει” καλύτερα το χρήστη, έτσι ώστε να προβεί σε αποτελεσματικότερες προβλέψεις ψήφου.
2. Εάν οι απαντήσεις των χρηστών στις δηλώσεις πολιτικής συνδυαστούν με κάποιες συμπληρωματικές ερωτήσεις, τότε το σύστημα θα μπορεί να αποκτήσει αυτήν την επιπλέον γνώση που χρειάζεται; Οι απαντήσεις στις συμπληρωματικές ερωτήσεις ενισχύουν την πρόβλεψη για την πρόθεση ψήφου του χρήστη. Το πρόβλημα σε αυτήν την περίπτωση είναι πως δεν απαντούν όλοι οι χρήστες στις συμπληρωματικές ερωτήσεις. Για αυτό προτείνεται μέσα από τη συγκεκριμένη διατριβή η χρήση τεχνικών παραγοντοποίησης πινάκων για την εκτίμηση των απαντήσεων στα αναπάντητα ερωτήματα.
3. Ο Ευρωπαίος Ψηφοφόρος διαφέρει από χώρα σε χώρα και έτσι οι συμπληρωματικές ερωτήσεις που βοηθούν στην πρόβλεψη ψήφου δεν είναι σε όλες τις περιπτώσεις οι ίδιες.

4. Ποιο ποσοστό των χρηστών που υποστηρίζουν το ίδιο κόμμα, δίνουν και τις ίδιες απαντήσεις στις δηλώσεις πολιτικής του VAA; Υπάρχει ποικιλομορφία στα κόμματα, όπως επίσης παρατηρούνται πολλοί χρήστες που δίνουν παρόμοιες απαντήσεις με χρήστες που υποστηρίζουν άλλα κόμματα. Σε αυτές τις περιπτώσεις, μπορούν να ομαδοποιηθούν και να εξαχθούν συμπεράσματα για τα ιδεολογικά χαρακτηριστικά του κάθε κόμματος, κάτι που γίνεται για πρώτη φορά. Επίσης μπορεί να εξεταστεί η συσπείρωση των χρηστών σε κάθε κόμμα.

Ταυτόχρονα η διατριβή επιχειρεί να εξηγήσει πώς μπορούν να εφαρμοστούν τεχνικές από ένα ευρέως εφαρμοσμένο πεδίο, όπως είναι τα Συστήματα Συστάσεων (ΣΣ), σε ένα καινούργιο πεδίο που είναι οι Ηλεκτρονικοί Σύμβουλοι Ψήφου (VAAs) για την επίλυση προβλημάτων που απασχολούν το σχεδιασμό και την υλοποίηση τους. Ενδεικτικά, τέτοιου είδους προβλήματα είναι τα πιο κάτω:

1. Η βελτιστοποίηση της αποτελεσματικότητας της σύστασης που γίνεται με τη χρήση τεχνικών που αποδίδουν καλύτερα και γρηγορότερα και με αποφάσεις σχετικά με το ποιες και πόσες από τις απαντήσεις των χρηστών μπορούν να χρησιμοποιούνται για την καλύτερη πρόβλεψη ψήφου.
2. Η αποτελεσματικότητα της κοινωνικής προσέγγισης των Ηλεκτρονικών Συμβούλων Ψήφου (Social Voting Advice Applications - SVAAs) και η υπεροχή τους έναντι της παραδοσιακής σύστασης, για κάθε χώρα ξεχωριστά. Ποιες τεχνικές μηχανικής μάθησης δίνουν καλύτερα αποτελέσματα στα SVAAs σε κάθε περίπτωση και που μπορεί να οφείλεται αυτό;
3. Οι συμπληρωματικές ερωτήσεις ενισχύουν την απόδοση της πρόβλεψης ψήφου, αλλά δεν απαντούν όλοι οι χρήστες σε αυτές. Υπό αυτές τις συνθήκες, πώς μπορούν να χρησιμοποιηθούν τεχνικές παραγοντοποίησης πινάκων για να προβλέψουν κάποιες από τις απαντήσεις στις συμπληρωματικές ερωτήσεις έχοντας στη διάθεση τους τις δηλώσεις πολιτικής του βασικού VAA ερωτηματολογίου;

Επιπρόσθετα εξετάζεται εάν οι δηλώσεις πολιτικής μπορούν να ομαδοποιηθούν σύμφωνα με κάποια κριτήρια για να μειωθούν οι διαστάσεις του προβλήματος σύστασης. Με αυτόν τον τρόπο για πρώτη φορά το πρόβλημα σύστασης στα VAAs αντιμετωπίζεται ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων που είναι ένα καλά εδραιωμένο πεδίο στον τομέα της Επιστήμης Αποφάσεων.

Το περιεχόμενο της διατριβής εμπίπτει στο ενδιαφέρον τόσο αυτών που ασχολούνται με τα πεδία της Επιστήμης των Υπολογιστών, των Συστημάτων Συστάσεων και της Επιστήμης των Δεδομένων, όσο και αυτών που καταπίνουν με τα πεδία της Πολιτικής Επιστήμης, της Επικοινωνίας και της Εξόρυξης Δεδομένων που εργάζονται σε θέματα που βρίσκονται στη διασταύρωση της πολιτικής και του διαδικτύου.

1.1.4 Δομή της Διατριβής

Η διατριβή χωρίζεται σε τέσσερα κύρια Κεφάλαια. Το περιεχόμενο τους διαμορφώνεται ως ακολούθως:

1. Το πρώτο Κεφάλαιο αποτελείται από την εισαγωγή και το θεωρητικό υπόβαθρο πίσω από τα ΣΣ και τα VAAs. Γίνεται αναφορά στον τρόπο λειτουργίας τους, στις εφαρμογές και στο αντίκτυπο τους καθώς και σε διάφορες πτυχές της τρέχουσας έρευνας. Επιπρόσθετα περιλαμβάνει την Ενότητα 1.7 που εκθέτει τις μεθόδους που εφαρμόστηκαν για την υλοποίηση των πειραμάτων και το δείγμα που χρησιμοποιήθηκε.
2. Το δεύτερο Κεφάλαιο αναφέρεται στην εφαρμογή της Μηχανικής Μάθησης στα VAAs και χωρίζεται στις πιο κάτω Ενότητες:

- Η Ενότητα 2.1 διερευνά την απόδοση της πρόβλεψης ψήφου που γίνεται τόσο με τον παραδοσιακό (VAAs) όσο και με τον κοινωνικό τρόπο σύστασης (SVAAs) για κάθε χρήστη. Χρησιμοποιούνται μόνο οι απαντήσεις που έδωσε ο κάθε χρήστης στις δηλώσεις πολιτικής του βασικού ερωτηματολογίου και γίνεται σύγκριση των αποτελεσμάτων σε πανευρωπαϊκό επίπεδο καθώς και σύγκριση της κοινωνικής με την παραδοσιακή προσέγγιση των VAAs, αναλύοντας ποιες από τις τεχνικές που εφαρμόστηκαν, δίνουν καλύτερα αποτελέσματα και για ποιο λόγο.
 - Η Ενότητα 2.2 μελετά εάν τα αποτελέσματα από την Ενότητα 2.1 βελτιώνονται όταν μαζί με τις απαντήσεις στις δηλώσεις πολιτικής, το σύστημα, λαμβάνει υπόψη και τις απαντήσεις των χρηστών σε συμπληρωματικές ερωτήσεις.
 - Η Ενότητα 2.3 διερευνά το πώς και αν μπορούν να προβλεφθούν κάποιες από τις απαντήσεις στις συμπληρωματικές ερωτήσεις, από όσους χρήστες δεν τις συμπλήρωσαν, με βάση τις απαντήσεις τους στις δηλώσεις πολιτικής.
3. Στο τρίτο Κεφάλαιο αναλύονται οι πολιτικές κοινωνικές επιπτώσεις των VAAs στις ακόλουθες Ενότητες:
- Η Ενότητα 3.1 διαχωρίζει τους χρήστες που προτίθενται να ψηφίσουν κάθε κόμμα σε ομάδες ομοίων χαρακτηριστικών και εξετάζει σε ποιο βαθμό υπάρχει ιδεολογική ταύτιση μεταξύ του κόμματος και των ομάδων αυτών.
 - Η Ενότητα 3.2 εξετάζει κατά πόσο μπορεί να υπολογιστεί το ενδιαφέρον του Ευρωπαίου χρήστη για την πολιτική, έχοντας ως δεδομένα τα κριτήρια με τα οποία επιλέγει να ψηφίσει ένα κόμμα και τα δημογραφικά του στοιχεία.
 - Η Ενότητα 3.3 αναφέρεται στην αντιμετώπιση του προβλήματος σύστασης στα VAAs ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων.
4. Στο τελευταίο Κεφάλαιο γίνεται σχολιασμός των αποτελεσμάτων και παρουσίαση των βασικών συμπερασμάτων της διατριβής, καθώς αναφέρονται και προτάσεις για μελλοντική έρευνα.

1.2 Θεωρητικό Υπόβαθρο

Ο Παγκόσμιος Ιστός (World-Wide Web) είναι ένα παγκόσμιο σύμπαν πληροφοριών, που επινοήθηκε το 1989 από τον Tim Berners - Lee, σε μια προσπάθεια του να αρχειοθετήσει τις επιστημονικές μελέτες των συνεργατών του στο κέντρο CERN¹. Γνώρισε γρήγορη και μεγάλη επιτυχία και λόγω της ευχρηστίας του ενσωματώθηκε στις υπηρεσίες του Διαδικτύου έχοντας τεράστια απήχηση. Τα τελευταία χρόνια διανύει τη δεύτερη γενιά του παρέχοντας στο χρήστη ένα δυναμικό περιβάλλον. Το Web 2.0, χάρη στο δυναμικό του χαρακτήρα επιτρέπει πλέον όχι μόνο την αναζήτηση πληροφοριών αλλά και την αλληλεπίδραση των χρηστών, οι οποίοι μέσα από την ανάπτυξη νέων εργαλείων και υπηρεσιών μπορούν εύκολα να συνεργαστούν, να επικοινωνήσουν και να μοιραστούν πληροφορίες μεταξύ τους. Το όνομα Web 2.0 δόθηκε από την O' Reilly media το 2004 και κάνει πράξη αυτό που χρόνια πριν οραματίστηκαν οι McLuhan [6], Νεγρεπόντε [7] και Δερτουζος [8], μια τεχνολογία, δηλαδή, που μετατρέπει τον κόσμο σε ένα παγκόσμιο χωριό επιτρέποντας την πρόσβαση από οπουδήποτε και οποτεδήποτε, αρκεί να υπάρχει σύνδεση στο Διαδίκτυο [9].

Η εμφάνιση του παγκόσμιου ιστού και η επακόλουθη αύξηση των διαθέσιμων δεδομένων που βρίσκονται σε απευθείας σύνδεση (online), οδήγησαν στη δημιουργία ενός τεράστιου όγκου ψηφιακών δεδομένων στον παγκόσμιο ιστό, ο οποίος συνεχίζει μέχρι και σήμερα να μεγαλώνει. Μάλιστα η έκθεση του

¹<http://ben.home.cern.ch/ben/TCPHIST.html>

2014 με τίτλο “The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things”² της εταιρείας EMC για το Ψηφιακό Σύμπαν, που δημιουργήθηκε σε συνεργασία με την IDC (International Data Corporation) για την ποσοτικοποίηση των δεδομένων που παράγονται ετησίως στον παγκόσμιο ιστό, διαπιστώνει ότι το ψηφιακό σύμπαν διπλασιάζεται κάθε δύο χρόνια με αποτέλεσμα το 2020 να προσεγγίσει τα 44 τρισεκατομμύρια gigabytes.

1.3 Συστήματα Συστάσεων

Τα Συστήματα Συστάσεων (Recommender Systems) είναι τεχνικές και εργαλεία λογισμικού που βοηθούν τους χρήστες να προσδιορίσουν τι χρειάζονται πραγματικά μέσα από ένα μεγάλο όγκο προϊόντων ή υπηρεσιών, παρέχοντάς τους εξατομικευμένες συστάσεις. Σκοπός των Συστημάτων Συστάσεων (ΣΣ) είναι να υιοθετούν τεχνικές για να μετατρέπουν τα δεδομένα που συλλέγουν, σχετικά με τους χρήστες και τις προτιμήσεις τους, σε μοντέλα χρηστών που μπορούν να χρησιμοποιηθούν για μελλοντικές προβλέψεις [10, 11]. Οι προβλέψεις αυτές αφορούν προτάσεις οι οποίες γίνονται στους χρήστες και βοηθούν στη λήψη διάφορων αποφάσεων, όπως την αγορά ενός αντικειμένου ή υπηρεσίας, το είδος της μουσικής που θα επέλεγαν να ακούσουν, το περιεχόμενο ενός βιβλίου που θα τους άρεσε να διαβάσουν, κ.ο.κ. Τα συστήματα αυτά είναι ιδιαίτερα χρήσιμα, αφού ο τεράστιος όγκος δεδομένων που διαχειρίζονται οι βάσεις δεδομένων πολλών σύγχρονων εφαρμογών, συχνά δυσκολεύουν τους χρήστες να ξεχωρίσουν τις πληροφορίες που πραγματικά σχετίζονται με τα ενδιαφέροντά τους [10].

Ο όρος ΣΣ εξελίχθηκε μέσα στα χρόνια καθώς διάφορες προσεγγίσεις των συστημάτων αυτών έκαναν την εμφάνισή τους [12]. Αν και ο ορισμός τους αρχικά αναπτύχθηκε μέσα από την τεχνική του συνεργατικού φιλτραρίσματος (Collaborative Filtering) [13, 14], αργότερα επεκτάθηκε και πλέον αφορά τις προτάσεις που γίνονται στους χρήστες και παράγονται με οποιοδήποτε τρόπο, χρησιμοποιώντας συχνά εκτός από το συνεργατικό φιλτράρισμα και προσεγγίσεις όπως αυτές που βασίζονται στο περιεχόμενο (Content-Based), σε συγκεκριμένη γνώση (Knowledge-based), στην κοινότητα (Community-based), στα δημογραφικά χαρακτηριστικά του χρήστη (Demographics), αλλά και σε συνδυασμό των διάφορων τεχνικών (Hybrid Recommender Systems) [15, 4, 16]. Ωστόσο, ανεξάρτητα από τον τρόπο που τα ΣΣ προσεγγίζουν ένα πρόβλημα πρότασης, η λογική των συστημάτων αυτών παραμένει η ίδια: Οι χρήστες αλληλεπιδρούν με τα ΣΣ για να αντλήσουν πληροφορίες, που μπορεί να τους χρησιμεύσουν στην αξιολόγηση ενός αντικειμένου ή υπηρεσίας.

1.3.1 Η λειτουργία των Συστημάτων Συστάσεων

Στην απλούστερη τους μορφή τα Συστήματα Συστάσεων (ΣΣ) λειτουργούν ως online υπηρεσίες, οι οποίες προσφέρουν εξατομικευμένες συστάσεις στους χρήστες χρησιμοποιώντας καταλόγους κατάταξης που περιέχουν τα πιο κατάλληλα προϊόντα ή υπηρεσίες για αυτούς, ανάλογα με το γούστο και τα ιδιαίτερα χαρακτηριστικά τους. Για να ολοκληρωθεί μια τέτοια υπολογιστική εργασία, τα ΣΣ συλλέγουν πληροφορίες για τις προτιμήσεις των χρηστών είτε μέσα από τις αξιολογήσεις που έκαναν για διάφορα προϊόντα είτε ερμηνεύοντας κάποιες από τις ενέργειες τους στον παγκόσμιο ιστό (Ricci et al. [5], pp. 1-36). Για παράδειγμα, η πλοήγηση του χρήστη σε μια συγκεκριμένη ιστοσελίδα, η οποία περιλαμβάνει λεπτομερείς πληροφορίες για ένα αντικείμενο και η παραμονή του στην ιστοσελίδα αυτή για ένα μεγάλο χρονικό διάστημα, μπορεί να θεωρηθεί από το ΣΣ ως μια ένδειξη για την προτίμηση που έχει ο χρήστης στο συγκεκριμένο προϊόν [4].

²<https://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>

Οι αλγόριθμοι που χρησιμοποιούνται στα ΣΣ παίρνουν ως είσοδο δεδομένα των χρηστών αλλά και των προϊόντων που εξετάζουν, όπως επίσης και στοιχεία από προηγούμενες συμπεριφορές που είχε ο ίδιος ο χρήστης ή και παρόμοιοι χρήστες με αυτόν.

Γενικά τα δεδομένα που μπορεί να επεξεργάζεται ένα ΣΣ είναι τα ακόλουθα:

- χαρακτηριστικά και ιδιότητες των χρηστών όπως είναι η ηλικία, το φύλο, το επάγγελμα τους [17],
- γνωρίσματα των προϊόντων που δείχνουν να προτιμούν, όπως για παράδειγμα το είδος της μουσικής, το συγγραφέα ενός βιβλίου, το περιεχόμενο μιας ταινίας, τα οποία μπορεί να διακρίνονται μέσα από τις βαθμολογήσεις τους σε προϊόντα που έχουν αγοράσει στο παρελθόν ή και από το ιστορικό των αναζητήσεων τους [18],
- προτιμήσεις των κοινωνικών φίλων και κοντινών τους ατόμων [19] καθώς και προτιμήσεις ανθρώπων με τους οποίους φαίνεται να έχουν παρόμοια γούστα [20, 21], να μοιράζονται τα ίδια χαρακτηριστικά και να αλληλεπιδρούν μεταξύ τους [22].

Όλα τα πιο πάνω δεδομένα αποθηκεύονται στη βάση δεδομένων του συστήματος, ταξινομούνται με βάση το εκτιμώμενο ενδιαφέρον του χρήστη και χρησιμοποιούνται στην επόμενη αλληλεπίδραση του με το σύστημα, έτσι ώστε να του προταθούν νέα αντικείμενα ή υπηρεσίες που πολύ πιθανόν να τον ενδιαφέρουν. Για τη δημιουργία συστάσεων εφαρμόζονται τεχνικές και μεθοδολογίες από πεδία όπως είναι η Αλληλεπίδραση Ανθρώπου Υπολογιστή (Human Computer Interaction - HCI) και η Ανάκτηση Πληροφορίας (Information Retrieval - IR). Ωστόσο, τα περισσότερα από τα συστήματα αυτά βασίζονται κυρίως σε τεχνικές Εξόρυξης Δεδομένων (Data Mining - DM) [23].

1.3.2 Εξόρυξη Δεδομένων και Μηχανική Μάθηση

Η Εξόρυξη Δεδομένων (ΕΔ) σχετίζεται με την εύρεση μοτίβων στα δεδομένα, κάτι που έκαναν οι άνθρωποι από την αρχή της ύπαρξης τους. Οι κυνηγοί έψαχναν και ψάχνουν μοτίβα στη συμπεριφορά των αποδημητικών ζώων, οι πολιτικοί στη συμπεριφορά των ψηφοφόρων και οι επιστήμονες στον τρόπο που λειτουργεί ο κόσμος. Τα μοτίβα βοηθούν στην ανάπτυξη θεωριών, οι οποίες μπορούν να χρησιμοποιηθούν σε προβλέψεις για το τι μπορεί να συμβεί σε καινούργιες καταστάσεις [24]. Το πρόβλημα που καλείται να λύσει ένα ΣΣ, το οποίο κάνει συστάσεις σε παρόμοιους χρήστες, ακολουθεί το μοτίβο: “Τα άτομα που φαίνονται παρόμοια, δρουν και παρόμοια” [25].

Η ΕΔ είναι ένα πεδίο που έχει ως βάση τη μάθηση, έτσι χρησιμοποιούνται τεχνικές μηχανικής μάθησης για την εύρεση και την περιγραφή δομικών σχημάτων στα δεδομένα, ως ένας τρόπος για να εξηγηθούν τα δεδομένα και να μπορούν να γίνουν προβλέψεις από αυτά [25]. Ο όρος “δεδομένα” στα ΣΣ αναφέρεται στα αντικείμενα και τα γνωρίσματα τους, όπου ως αντικείμενα μπορούν να δηλώνονται οι καταχωρήσεις του χρήστη, διάφορα στοιχεία, σημεία, δείγματα, παρατηρήσεις ή παραδείγματα και ως γνωρίσματα θεωρούνται οι μεταβλητές, τα πεδία, τα χαρακτηριστικά ή οι δυνατότητες που έχουν τα αντικείμενα. Τα πραγματικά δεδομένα που προέρχονται από τους χρήστες χρειάζονται μια προ-επεξεργασία, όπως είναι το “καθάρισμα” ή/και η τροποποίηση τους, έτσι ώστε να είναι σε θέση να χρησιμοποιηθούν από αλγόριθμους μηχανικής μάθησης [23].

Ο Marsland (2015) διακρίνει τους αλγόριθμους μάθησης σε τέσσερις κατηγορίες:

- την Επιβλεπόμενη Μάθηση (Supervised Learning),
- τη Μάθηση Χωρίς Επιβλεψη-Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning),

- την Ενισχυτική Μάθηση (Reinforcement Learning), και την
- την Εξελικτική Μάθηση (Evolutionary Learning).

Στην επιβλεπόμενη μάθηση παρέχεται ένα σύνολο εκπαίδευσης που περιλαμβάνει ένα αριθμό παραδειγμάτων με τις σωστές απαντήσεις, τις οποίες προσπαθεί να προβλέψει ο αλγόριθμος όταν εκπαιδευτεί. Κατά τη μη επιβλεπόμενη μάθηση οι ορθές απαντήσεις δεν παρέχονται, αντίθετα ο αλγόριθμος προσπαθεί να εντοπίσει ομοιότητες μεταξύ των εισόδων, έτσι ώστε οι εισοδοί που έχουν κάτι κοινό να κατηγοριοποιούνται μαζί. Η ενισχυτική μάθηση είναι ένας συνδυασμός μεταξύ της επιβλεπόμενης και μη επιβλεπόμενης μάθησης. Όσο αφορά την εξελικτική μάθηση πρόκειται για μάθηση που προσπαθεί να μοντελοποιήσει τη βιολογική εξέλιξη για να αξιολογήσει πόσο καλή είναι μια τρέχουσα λύση.

Ο πιο γνωστός τύπος μάθησης είναι η επιβλεπόμενη μάθηση. Τα δεδομένα εκπαίδευσης που λαμβάνει αποτελούνται από ένα σύνολο ζευγών εισόδου - εξόδου. Οι εισοδοί είναι τα γνωστά δεδομένα που λαμβάνει ο αλγόριθμος και οι εξοδοί είναι τα αποτελέσματα που καλείται να παράγει, σύμφωνα πάντα με την είσοδο που παίρνει. Στη διάρκεια της μάθησης και οι εξοδοί είναι γνωστοί, έτσι ώστε ο αλγόριθμος να “βλέπει” τα αποτελέσματα που έχει η κάθε είσοδος και να “μαθαίνει” από αυτά (μάθηση μέσω παραδειγμάτων - learning by example).

Η παλινδρόμηση (regression) και η ταξινόμηση (classification) είναι τεχνικές επιβλεπόμενης μάθησης. Στην πιο απλή τους μορφή αποτελούνται από μια είσοδο x και μια έξοδο y και έχουν ως στόχο να κατανοήσουν πώς η είσοδος x καταλήγει στην συγκεκριμένη έξοδο y . Ο τρόπος με τον οποίο προσεγγίζονται τέτοιου είδους προβλήματα από τη μηχανική μάθηση είναι μέσω της δημιουργίας ενός μοντέλου που ορίζεται από ένα σύνολο παραμέτρων: $y = g(x|\theta)$, όπου $g(\cdot)$ είναι το μοντέλο και θ είναι παράμετροι του. Στην περίπτωση της παλινδρόμησης το y είναι ένας αριθμός, ενώ στην ταξινόμηση είναι ένας κώδικας που δηλώνει μια κατηγορία (π.χ., 0/1). Η μηχανική μάθηση βελτιστοποιεί τις παραμέτρους θ , έτσι ώστε το προσεγγιστικό σφάλμα να ελαχιστοποιείται, δηλαδή, οι εκτιμήσεις να είναι όσο το δυνατόν πλησιέστερες προς τις σωστές τιμές που δίνονται στο σύνολο εκπαίδευσης [26]. Μόλις βρεθεί ένας κανόνας που ταιριάζει με τα δεδομένα του παρελθόντος, αν το μέλλον είναι παρόμοιο με το παρελθόν, τότε μπορεί να γίνουν σωστές προβλέψεις για νέες περιπτώσεις [26, 23, 27].

1.3.3 Προσεγγίσεις Συστημάτων Συστάσεων

Υπάρχουν διάφοροι τρόποι, με τους οποίους ένα ΣΣ προσεγγίζει το πρόβλημα της δημιουργίας εξατομικευμένων συστάσεων. Ωστόσο μέχρι σήμερα οι πιο δημοφιλείς είναι το συνεργατικό φιλτράρισμα (Collaborative filtering) και το φιλτράρισμα με βάση το περιεχόμενο (Content-Based) [28]. Με το φιλτράρισμα βάσει του περιεχομένου προτείνονται αγαθά ή υπηρεσίες σύμφωνα με το περιεχόμενό τους, δημιουργώντας συνήθως ένα προφίλ που περιέχει τις λέξεις-κλειδιά από τα αρεστά αντικείμενα ενός χρήστη, το οποίο χρησιμοποιείται για την εύρεση νέων προϊόντων που ταιριάζουν στο προφίλ αυτό και πολύ πιθανόν να ανήκουν στις προτιμήσεις του χρήστη. Σύμφωνα με αυτό τον τύπο, προτείνονται στο χρήστη αντικείμενα, τα οποία είναι παρόμοια με αυτά που έδειξε να του αρέσουν στο παρελθόν. Η στρατηγική αυτή αντιμετωπίζει προβλήματα όταν το περιεχόμενο είναι δύσκολο να αναλυθεί [29].

Στο Συνεργατικό Φιλτράρισμα συγκρίνονται οι χρήστες μεταξύ τους έτσι ώστε να βρεθούν οι όμοιοι χρήστες, αυτοί δηλαδή που αξιολογούν τα ίδια αντικείμενα με παρόμοιο τρόπο και για αυτό υπολογίζεται πώς έχουν παρόμοιο γούστο. Στη συνέχεια, το ΣΣ ψάχνει τα αντικείμενα που δεν αξιολογήθηκαν από όλους τους όμοιους χρήστες, πράγμα που σημαίνει πως δε χρησιμοποιήσαν όλοι ακόμη τα συγκεκριμένα αντικείμενα. Ακολούθως προτείνει όσα από αυτά τα αντικείμενα πήραν υψηλή βαθμολογία στους χρήστες που δεν τα χρησιμοποίησαν ακόμη. Με αυτή τη μέθοδο, γίνεται πρόταση σε ένα συγκεκριμένο άτομο

για αντικείμενα ή υπηρεσίες που άρεσαν στους χρήστες που έχουν παρόμοια γούστα με αυτόν [29].

Μια άλλη προσέγγιση των ΣΣ είναι αυτή που βασίζεται σε συγκεκριμένη γνώση (Knowledge-based) και προτείνει τα προϊόντα στο χρήστη που μπορούν να καλύψουν τις δεδομένες ανάγκες του. Επιπρόσθετα ένα ΣΣ μπορεί να στηρίζει τις συστάσεις του στην κοινότητα του χρήστη (Community-based) κάνοντας προτάσεις για προϊόντα ή υπηρεσίες που φαίνεται να αρέσουν στους φίλους και τα κοντινά του πρόσωπα. Ακόμη, μπορούν να προταθούν αντικείμενα ή υπηρεσίες στο χρήστη βάσει των δημογραφικών χαρακτηριστικών του (Demographic-based), όπως για παράδειγμα ένας χρήστης να οδηγηθεί σε μια συγκεκριμένη ιστοσελίδα σύμφωνα με τη χώρα και τη γλώσσα του ή να του γίνει μια πρόταση ανάλογα με το φύλο ή την ηλικία του. Όταν ένα ΣΣ κάνει προτάσεις για αντικείμενα στο χρήστη βάσει των δημογραφικών χαρακτηριστικών του, τότε μπορεί να βασίζεται στη χώρα που βρίσκεται, στη μητρική του γλώσσα, στο φύλο ή/και στην ηλικία που έχει. Για παράδειγμα σε μια ιστοσελίδα που πωλεί βρεφικά ρούχα, μπορεί να γίνει σύσταση μιας μπλε φανέλας που έχει τυπωμένους ελληνικούς χαρακτήρες σε μια μητέρα που έχει αγοράκι και έχει την ελληνική ως μητρική γλώσσα. Βέβαια υπάρχουν και οι περιπτώσεις όπου οι πιο πάνω προσεγγίσεις μπορούν να συνδυαστούν, έτσι ώστε να αξιοποιηθούν τα πλεονεκτήματα της μιας μεθόδου για να αντιμετωπιστούν τα μειονεκτήματα κάποιας άλλης (Hybrid Recommender Systems) [4, 30].

Για την εκπόνηση της συγκεκριμένης διατριβής, τα διάφορα ερευνητικά προβλήματα προσεγγίστηκαν με τεχνικές του Συνεργατικού Φιλτραρίσματος. Ως εκ τούτου η μέθοδος του Συνεργατικού Φιλτραρίσματος επεξηγείται με περισσότερες λεπτομέρειες στην επόμενη υποενότητα.

1.3.4 Συνεργατικό Φιλτράρισμα (Collaborative Filtering)

Το συνεργατικό φιλτράρισμα είναι ένας δημοφιλής αλγόριθμος σύστασης που βασίζει τις προβλέψεις και τις συστάσεις του στη συμπεριφορά ή στις αξιολογήσεις των χρηστών σχετικά με προϊόντα ή υπηρεσίες που χρησιμοποίησαν στο παρελθόν. Η θεμελιώδης παραδοχή που κρύβεται πίσω από τη μέθοδο αυτή είναι πως εάν βρεθούν άτομα που φαίνεται να μοιράζονται το ίδιο γούστο με τον ενεργό χρήστη, τότε οι απόψεις τους μπορούν να επιλεγούν και να συναθροιστούν με τέτοιο τρόπο ώστε να δημιουργηθεί μια εύλογη πρόβλεψη σχετικά με την προτίμηση που πιθανόν να δείξει ο ενεργός χρήστης σε ένα αντικείμενο, για το οποίο δε γνωρίζει ή δεν έτυχε να το δοκιμάσει αλλά παρόμοιοι χρήστες σχημάτισαν θετική άποψη για αυτό. Διαισθητικά η προσέγγιση αυτή βασίζεται στην υπόθεση πως αν οι χρήστες συμφωνούν για την ποιότητα ή την καταλληλότητα κάποιων συγκεκριμένων αντικειμένων που αγόρασαν ή γνώρισαν με κάποιο τρόπο, τότε κατά πάσα πιθανότητα θα συμφωνήσουν και για άλλα αντικείμενα στο μέλλον [21]. Για παράδειγμα, εάν ένας χρήστης A αξιολόγησε κάποιο βιβλίο με υψηλή βαθμολογία και ένας άλλος χρήστης B έπραξε το ίδιο, τότε θα προταθούν στο χρήστη A βιβλία που δεν έχει αγοράσει και βρίσκονται στο ιστορικό των αγορών του χρήστη B με υψηλή βαθμολογία, γεγονός που δείχνει ότι έμεινε ικανοποιημένος.

Η δημιουργία σύστασης με βάση το συνεργατικό φιλτράρισμα έχει χαρακτηριστεί ως η προσέγγιση “από άνθρωπο σε άνθρωπο” (“από άνθρωπο σε άνθρωπο” προώθηση ονομάζεται η δωρεάν προώθηση προϊόντων και υπηρεσιών, κατά την οποία ευχαριστημένοι καταναλωτές μεταφέρουν από στόμα σε στόμα -word of mouth³- την ικανοποίησή τους για το συγκεκριμένο προϊόν ή την υπηρεσία που χρησιμοποίησαν, συμβάλλοντας στην προώθηση τους). Είναι από τους πλέον επιτυχημένους τρόπους διαφήμισης προϊόντων, με υψηλό δείκτη αξιοπιστίας, καθώς οι καταναλωτές που συμμετέχουν στη διαδικασία αυτή λειτουργούν αυθόρμητα, χωρίς την ύπαρξη προσωπικού συμφέροντος. Έτσι το σύστημα συστάσεων σε αυτή την περίπτωση παίζει το ρόλο του μεσάζοντα, αφού στην ουσία μεταφέρει την ικανοποίηση του

³https://en.wikipedia.org/wiki/Word_of_mouth

ενός χρήστη στον άλλον [31].

Στο Συνεργατικό Φιλτράρισμα οι αλγόριθμοι σύστασης συνήθως χωρίζονται σε αυτούς που βασίζονται στη μνήμη (memory based) και σε αυτούς που βασίζονται σε μοντέλο (model based). Από τη μια πλευρά, οι αλγόριθμοι που βασίζονται στη μνήμη χρησιμοποιούν όλα τα δεδομένα που έχει στη μνήμη του το σύστημα για να υπολογίσουν την ομοιότητα ανάμεσα σε προϊόντα ή χρήστες, έτσι ώστε να είναι σε θέση να κάνουν προβλέψεις σχετικά με αντικείμενα που πολύ πιθανόν να ήθελε να δει ή να αγοράσει ο χρήστης. Από την άλλη, το συνεργατικό φιλτράρισμα που έχει ως βάση του το μοντέλο, χρησιμοποιεί τα δεδομένα σχετικά με ένα χρήστη, που είναι αποθηκευμένα στο σύστημα, για να φτιάξει ένα μοντέλο, το οποίο θα συμβουλευτεί για τις μελλοντικές του προτάσεις [20]. Αυτό που κάνει είναι να μοντελοποιεί τις αλληλεπιδράσεις που έχουν οι χρήστες με τα προϊόντα, έχοντας εις γνώση τους παράγοντες που εκπροσωπούν τα χαρακτηριστικά του χρήστη και των προϊόντων στο σύστημα. Το μοντέλο εκπαίδευεται χρησιμοποιώντας τα διαθέσιμα δεδομένα και στη συνέχεια εφαρμόζεται για να προβλέψει τις βαθμολογίες των χρηστών σε νέα προϊόντα.

Οι αλγόριθμοι που βασίζονται στη μνήμη θεωρούνται πιο παραδοσιακοί τρόποι προσέγγισης του προβλήματος της σύστασης μέσω Συνεργατικού Φιλτραρίσματος καθώς όλες οι βαθμολογίες των χρηστών αποθηκεύονται στη μνήμη του συστήματος και χρησιμοποιούνται από εκεί κατευθείαν για τη δημιουργία προτάσεων. Στη σημερινή εποχή έχουν αναπτυχθεί πιο σύνθετες προσεγγίσεις ΣΦ που βασίζονται σε μοντέλα εφαρμόζοντας τεχνικές από διάφορους τομείς, όπως είναι η μηχανική μάθηση, η ανάκτηση πληροφορίας και η εξόρυξη δεδομένων, καθώς επίσης και μεθόδους που βασίζονται σε τεχνικές Γραμμικής Άλγεβρας, όπως η ανάλυση του πίνακα σε ιδιάζουσες τιμές (Singular Value Decomposition – SVD), για να προβλέψουν τις αξιολογήσεις που θα έκαναν οι χρήστες σε αντικείμενα, αλλά και για να μάθουν πώς να ταξινομούν σωστά με σειρά προτεραιότητας προϊόντα χρήσιμα για ένα χρήστη. Ο σχεδιασμός και η ανάπτυξη μοντέλων μπορεί να επιτρέψει σε ένα ΣΣ να μάθει να αναγνωρίζει πολύπλοκα σχήματα βάσει των δεδομένων εκπαίδευσης, ώστε να είναι σε θέση να κάνει έξυπνες προβλέψεις για τον τρόπο που θα αντιδράσει το συνεργατικό φιλτράρισμα στα δεδομένα δοκιμών (test data) ή στα πραγματικά δεδομένα (real-world data), με βάση τα μοντέλα που δημιουργήθηκαν.

Διάφορες τεχνικές εξόρυξης δεδομένων, όπως οι τεχνικές ταξινόμησης (π.χ. τα Bayesian δίκτυα - Bayesian networks) και οι μηχανές διανυσμάτων υποστήριξης (support vector machines), οι τεχνικές ομαδοποίησης (π.χ. ο αλγόριθμος k-means) και οι κανόνες συσχέτισης, έχουν ερευνηθεί ως απάντηση στα προβλήματα επεκτασιμότητας που αντιμετωπίζουν οι τεχνικές ΣΦ με βάση τη μνήμη (memory based) λόγω των δεκάδων εκατομμυρίων χρηστών και αντικειμένων που υπάρχουν στις βάσεις δεδομένων τους [4, 30].

Τα παραδοσιακά ΣΣ που χρησιμοποιούν συνεργατικό φιλτράρισμα στηρίζουν τις συστάσεις τους σε ένα και μόνο κριτήριο. Όταν εξετάζουν ένα προϊόν ή μια υπηρεσία για τη χρησιμότητά τους σε ένα συγκεκριμένο χρήστη, λαμβάνουν ως μοναδικό κριτήριο τη συνολική βαθμολογία που δόθηκε από όμοιους χρήστες για να δηλώσουν κατά πόσο έμειναν ικανοποιημένοι από το εν λόγω προϊόν ή την εν λόγω υπηρεσία. Από την άλλη πλευρά, ένα πρόβλημα σύστασης μπορεί να αντιμετωπίζεται ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων. Δηλαδή, αντί να βασίζεται μόνο στη συνολική βαθμολογία, το πρόβλημα σύστασης μπορεί να λαμβάνει υπόψη και τις βαθμολογίες που έδωσε ο χρήστης επιπρόσθετα σε άλλα κριτήρια που αφορούν συγκεκριμένα χαρακτηριστικά του προϊόντος ή της υπηρεσίας. Για παράδειγμα σε ένα ΣΣ που χρησιμοποιείται για ταινίες, εκτός από τη συνολική βαθμολογία μπορούν να δίνονται βαθμολογίες και σε κάποια επιμέρους κριτήρια όπως είναι το σενάριο, οι ηθοποιοί, η σκηνοθεσία και τα οπτικά εφέ. Αυτό βοηθά στο να γίνει κατανοητό πώς ένας χρήστης αξιολογεί μια σειρά από κριτήρια για να καταλήξει στη συνολική του βαθμολογία και ως εκ τούτου ποια είναι η διαδικασία λήψης αποφάσεων που ακολουθεί. Η διαδικασία λήψης αποφάσεων θεωρείται ως η “γνωστική διαδικασία, η

οποία οδηγεί στην επιλογή μιας πορείας δράσης μεταξύ διάφορων εναλλακτικών σεναρίων”. Το αποτέλεσμα της διαδικασίας αυτής περιλαμβάνει μια τελική απόφαση, η οποία μπορεί να είναι μια ενέργεια ή μια άποψη [32].

1.3.5 Εφαρμογές των Συστημάτων Συστάσεων

Τα τελευταία χρόνια τα ΣΣ έχουν τύχει ευρείας εφαρμογής, έχοντας απήχηση τόσο στον εμπορικό [33] όσο και στον ερευνητικό τομέα, όπου η προσοχή εστιάζεται κυρίως στην ανάπτυξη νέων τρόπων προσέγγισης του προβλήματος σύστασης [12, 34]. Τα συστήματα αυτά συναντώνται κυρίως στο ηλεκτρονικό εμπόριο, στην ηλεκτρονική διακυβέρνηση, στο ηλεκτρονικό επιχειρείν, σε ηλεκτρονικές βιβλιοθήκες, καθώς και στην ηλεκτρονική μάθηση [35].

Μια από τις πιο δημοφιλείς περιπτώσεις, όπου χρησιμοποιούνται ΣΣ, είναι το ηλεκτρονικό κατάστημα *Amazon*. Η ιστοσελίδα αυτή χρησιμοποιεί τις αξιολογήσεις που έκανε ο κάθε χρήστης μετά από κάθε αγορά του για να καταλήξει σε παρόμοια προϊόντα, τα οποία στη συνέχεια συγκεντρώνει σε μια λίστα και τα συστήνει σε χρήστες που δεν τα έχουν αγοράσει [36]. Επίσης τα συστήματα αυτά εμφανίζονται σε μηχανές αναζήτησης, όπως είναι το *Google ads*, τα μέσα κοινωνικής δικτύωσης όπως το *Facebook*, οι εφαρμογές διαμοιρασμού βίντεο, για παράδειγμα το *Youtube*, οι διαδικτυακές υπηρεσίες ταινιών, όπως το *Netflix* και σε κάθε περίπτωση που μπορεί να προταθούν στο χρήστη προϊόντα ή υπηρεσίες που σύμφωνα με το προφίλ του τον αφορούν. Στην παρούσα διατριβή μελετούμε την εφαρμογή των ΣΣ σε Ηλεκτρονικούς Συμβούλους Ψήφου.

1.4 Ηλεκτρονικοί Σύμβουλοι Ψήφου

Παρόλο που το δικαίωμα ψήφου δίνει λόγο και δύναμη στον πολίτη, επιτρέποντας του να αποφασίζει ο ίδιος για τον τρόπο διακυβέρνησης της χώρας του και να γίνεται αναπόσπαστο κομμάτι των πολιτικών εξελίξεων του τόπου του, οι πολίτες τη σημερινή εποχή τείνουν να μη λαμβάνουν μέρος σε εκλογικές διαδικασίες. Η χαμηλή προσέλευση των ψηφοφόρων στις κάλπες είναι ένα φαινόμενο που απειλεί την αντιπροσωπευτική δημοκρατία σε πολλές ανεπτυγμένες χώρες μιας και πολλοί ψηφοφόροι δεν αντιπροσωπεύονται σωστά. Οι Ladner και Pianzola ([37]) αναφέρουν συγκεκριμένα ως παράδειγμα την Ελβετία, όπου η προσέλευση των ψηφοφόρων στις κάλπες δεν έχει ξεπεράσει το 50% από το 1975.

Το μειωμένο ενδιαφέρον των ψηφοφόρων για την άσκηση ψήφου, οφείλεται σε μεγάλο βαθμό στην έλλειψη γνώσης από την πλευρά των πολιτών σχετικά με τα πολιτικά κόμματα και τις πολιτικές εξελίξεις στη χώρα τους καθώς επίσης και στην περίπλοκη και χρονοβόρα διαδικασία λήψης αποφάσεων, την οποία ακολουθούν οι ψηφοφόροι μέχρι να καταλήξουν στον υποψήφιο ή στο κόμμα, που πρόκειται να υποστηρίξουν [38, 37, 39]. Ως εκ τούτου είναι αναγκαία η εύκολη προσέγγιση σε ακριβείς και έγκυρες πληροφορίες σχετικά με τα πολιτικά δρώμενα μιας χώρας και τις θέσεις των κομμάτων σε αυτά, έτσι ώστε οι πολίτες να εξοικειώνονται με τους υποψηφίους και να είναι σωστά και επαρκώς ενημερωμένοι για τις οποιοσδήποτε πολιτικές εξελίξεις στη χώρα τους. Η πρόσβαση αυτή στην ενημέρωση, γίνεται εύκολα και γρήγορα με την λεγόμενη «ηλεκτρονική δημοκρατία», η οποία υποστηρίζει τη διαδικασία συλλογής πληροφοριών και τη διαμόρφωση γνώμης πριν από την ημέρα των εκλογών [40].

1.4.1 Η “ηλεκτρονική δημοκρατία”

Η “ηλεκτρονική δημοκρατία” είναι ένας τρόπος για να χρησιμοποιηθούν οι νέες τεχνολογίες, έτσι ώστε να αυξηθεί η συμμετοχή των πολιτών στις δημοκρατικές διαδικασίες και να προωθηθεί η άμεση

συμμετοχή σε πολιτικές δραστηριότητες [41]. Οι Ηλεκτρονικοί Σύμβουλοι Ψήφου (από το 2007 χρησιμοποιείται ο κοινός όρος Voting Advice Applications (VAAs) [42] είναι εργαλεία της “ηλεκτρονικής δημοκρατίας”, τα οποία ενημερώνουν τους πολίτες σχετικά με τις πολιτικές θέσεις των κομμάτων ή των υποψηφίων, έτσι ώστε να διευκολυνθεί η διαδικασία λήψης αποφάσεων.

Τα VAAs είναι διαδικτυακές εφαρμογές που συγκρίνουν τις πολιτικές απόψεις των χρηστών τους με αυτές των κομμάτων ή υποψηφίων, προτείνοντας, τελικά, στον κάθε χρήστη τον υποψήφιο ή το κόμμα που του ταιριάζει καλύτερα. Έχουν γίνει πολύ δημοφιλή στην Ευρώπη και πέραν αυτής [43, 44, 45, 46]. Ο ρόλος τους, ουσιαστικά, είναι να προσφέρουν ακόμα ένα κανάλι ενημέρωσης προς τους ψηφοφόρους, έτσι ώστε να τους παρέχουν την απαραίτητη πληροφορία που χρειάζονται για να είναι σε θέση να συγκρίνουν τις δικές τους απόψεις με αυτές των υποψηφίων / κομμάτων και να οδηγούνται στο ποιόν υποψήφιο ή πολιτικό κόμμα τελικά θα υποστηρίξουν στις εκλογές [42, 38, 37, 43, 46]. Έπειτα, εφόσον ο πολίτης είναι καλά πληροφορημένος, μπορεί να αντιπαραθέτει τις θέσεις των κομμάτων, για να αντιλαμβάνεται τις διαφορές και ομοιότητες που μπορεί να έχουν τα κόμματα ή υποψήφιοι μεταξύ τους και να μπορεί να κατανοήσει πώς ένα κόμμα μπορεί να είναι περισσότερο “χρήσιμο” από κάποιον άλλο για τον ίδιο αλλά και για τη χώρα στην οποία διαμένει [47].

Τα VAAs συστήνοντας στους χρήστες τον υποψήφιο / κόμμα που τους ταιριάζει καλύτερα και δίνοντας εύκολη πρόσβαση σε πληροφορίες σχετικά με τις θέσεις και αντιλήψεις των κομμάτων για τα τρέχοντα θέματα της χώρας τους, μπορούν να οδηγήσουν στις κάλπες τα άτομα που απέχουν από τις εκλογικές διαδικασίες γιατί δυσκολεύονται να εντοπίσουν τον υποψήφιο με τον οποίο συμφωνούν οι απόψεις τους, τους κοινωνικά απομονωμένους πολίτες που δυσκολεύονται να βρουν πρόσβαση στην πληροφορία και τον κόσμο που δε νιώθει αρμόδιος για να ψηφίσει λόγω της έλλειψης γνώσης σχετικά με τα πολιτικά θέματα [37]. Με τη χρήση των VAAs, λοιπόν, ενισχύονται οι δημοκρατικές διαδικασίες μιας χώρας, αφού ο πολίτης στα δημοκρατικά πολιτεύματα πρέπει να είναι πολιτικά ενεργός, να ενημερώνεται προσεκτικά για τα δρώμενα του τόπου του και τις θέσεις των πολιτικών κομμάτων σε σχέση με αυτά και να ασκεί το δικαίωμα ψήφου που έχει.

1.4.2 Το αντίκτυπο των Ηλεκτρονικών Συμβούλων Ψήφου

Το πρώτο VAA δημιουργήθηκε στην Ολλανδία με το όνομα *Stemwijzer*, που σημαίνει ψηφίζω σοφά [48, 39]. Εμφανίστηκε αρχικά το 1989 ως Διδακτικό Εγχειρίδιο Κοινωνικής Μελέτης για τα γυμνάσια που περιείχε ένα απλό γραπτό τεστ από 60 δηλώσεις των πολιτικών κομμάτων (κυρίως ιδεολογικές) και μία δισκέτα. Μετά από 10 περίπου χρόνια το *Stemwijzer* έγινε διαθέσιμο στο διαδίκτυο παρέχοντας 250.000 συμβουλές ψήφου [42, 39]. Κατά τη διάρκεια των εκλογών του 2006, οι χρήστες του *Stemwijzer* αυξήθηκαν σε 4.7 εκατομμύρια, αριθμός που αντιστοιχεί στο 40% των Ολλανδών ψηφοφόρων (Walgrave et al. [49], p. 52), ενώ παράλληλα παράχθηκαν ακόμη 1.5 εκατομμύρια συμβουλές ψήφου από ένα δεύτερο Ολλανδικό VAA με το όνομα *Kieskompas* [42]. Τα VAAs δεν άργησαν να γίνουν δημοφιλή και να χρησιμοποιηθούν και από πολλές άλλες χώρες στις προεκλογικές περιόδους. Κάποια παραδείγματα των VAAs είναι το *Smartvote* στην Ελβετία, το οποίο ξεκίνησε δίνοντας 255.000 συμβουλές ψήφου το 2003 και κατέληξε να χρησιμοποιείται περισσότερες από 1.2 εκατομμύρια φορές κατά τη διάρκεια της προεκλογικής εκστρατείας των εθνικών εκλογών του 2011 και το *Wahl-o-mat*⁴ στη Γερμανία που πρωτοεμφανίστηκε το 2002 δίνοντας 3.6 εκατομμύρια συμβουλές ψήφου και έφτασε να χρησιμοποιείται περίπου 13.3 εκατομμύρια φορές το 2013. Μερικά ακόμα γνωστά VAAs είναι το *Do the Vote Test* στο Βέλγιο, το *Pick your Party* στην Ιρλανδία, το *Meuvoto* στη Βραζιλία, το *Choose4greece* και *Helpmevote* στην Ελλάδα, το *Peruescoge* στο Περού, το *Choose4cyprus* στην Κύπρο, το *Scottish vote compass* στη Σκωτία και άλλα

⁴<http://www.wahl-o-mat.uni-duesseldorf.de/wahl-o-mat-forschung/en/facts-about-the-wahl-o-mat/>

που εφαρμόστηκαν στη Φινλανδία, Σουηδία, Νορβηγία, Σλοβακία, Τσεχία, Πολωνία, Αυστρία, Ιταλία, Πορτογαλία και τα τελευταία χρόνια με μεγάλη επιτυχία στη Γαλλία, τον Καναδά, και την Τουρκία [50].

Τα VAAs εκτός από τις εθνικές εκλογές, έχουν χρησιμοποιηθεί και στις εκλογές για το Ευρωπαϊκό Κοινοβούλιο υπολογίζοντας την ομοιότητα μεταξύ των χρηστών και των κομμάτων από την Ευρώπη. Το *VoteMatch EU*⁵ είναι ένα εκπαιδευτικό εργαλείο πληροφόρησης που παρέχει πρόσβαση σε πολιτικές πληροφορίες σχετικά με τις ευρωπαϊκές εκλογές. Πρόκειται για ένα διεθνές δίκτυο ανάμεσα σε δεκατέσσερις χώρες της ΕΕ: Γερμανία, Ηνωμένο Βασίλειο, Βέλγιο, Πολωνία, Λετονία, Ελλάδα, Βουλγαρία, Αυστρία, Ιταλία, Ισπανία, Τσεχία, Σλοβακία, Γαλλία και Ολλανδία. Πρωτοεμφανίστηκε στις ευρωπαϊκές εκλογές του 2004 στα πλαίσια του NECE (Network of European Citizenship Education)⁶, μια πρωτοβουλία για τη δικτύωση της εκπαίδευσης της ευρωπαϊκής ιθαγένειας. Ωστόσο, αξίζει να σημειωθεί ότι αυτή η εφαρμογή δεν επεσήμανε στους χρήστες το βαθμό ομοιότητας τους με πραγματικά πολιτικά κόμματα, αλλά μόνο με “ονομαστικές” διακρατικές ομάδες κομμάτων (π.χ. το Ευρωπαϊκό Λαϊκό Κόμμα) του Ευρωπαϊκού Κοινοβουλίου [51, 52].

Στο τέλος του Απριλίου το 2009 λειτούργησε το *EU Profiler*, έργο μιας κοινοπραξίας ερευνητικών ιδρυμάτων στην Ιταλία, την Ολλανδία και την Ελβετία. Το *EU Profiler*, σχεδιάστηκε από το Ευρωπαϊκό Πανεπιστημιακό Ινστιτούτο (EUI) στη Φλωρεντία σε συνεργασία με το *Kieskompas* και το *Smartvote*, επιτρέποντας στους πολίτες να συγκρίνουν τις θέσεις τους με τα κόμματα ολόκληρης της Ευρώπης [53].

Στις εκλογές για το Ευρωπαϊκό Κοινοβούλιο το 2014 εμφανίστηκε το *EUvox*, μια πανευρωπαϊκή εφαρμογή ηλεκτρονικού συμβούλου ψήφου. Σχεδιάστηκε και υλοποιήθηκε από μια κοινοπραξία ερευνητών, με βάση το *Kieskompas* και το *PreferenceMatcher*. Το *PreferenceMatcher* είναι μια ακαδημαϊκή κοινοπραξία που περιλαμβάνει πολιτικούς επιστήμονες, κοινωνικούς ψυχολόγους, επιστήμονες της πληροφορικής, καθώς και ειδικούς της επικοινωνίας από το Πανεπιστήμιο της Ζυρίχης, το Πανεπιστήμιο του Twente και το Τεχνολογικό Πανεπιστήμιο Κύπρου, τα οποία συνεργάζονται στην ανάπτυξη εργαλείων ηλεκτρονικού αλφαριθμητισμού για την ενίσχυση της εκπαίδευσης των ψηφοφόρων. Το *EUvox* δημιουργήθηκε για να ενημερώσει τους πολίτες σχετικά με τις πολιτικές θέσεις των κομμάτων στις εκλογές εκείνες και να τους βοηθήσει να αποφασίσουν ποιο πολιτικό κόμμα ταιριάζει καλύτερα στις δικές τους προτιμήσεις.

Οι παράγοντες που έχουν παίξει ρόλο στην επιτυχία των VAAs μέχρι σήμερα είναι η εξάπλωση του διαδικτύου, η οποία προσφέρει εύκολη πρόσβαση σε αυτά, η ανάπτυξη της τεχνολογίας η οποία κάνει την ανάπτυξη τους πιο εύκολη και τα Μέσα Μαζικής Ενημέρωσης που ενημερώνουν για αυτά και προωθούν τη χρήση τους σε χώρες που δεν τα έχουν εφαρμόσει ακόμη [51]. Το αυξανόμενο ενδιαφέρον για τα VAAs δεν έγκειται μόνο στο γεγονός ότι το σύστημα μπορεί να ταιριάζει τους χρήστες με τους υποψηφίους που βρίσκονται πιο κοντά στις δικές τους απόψεις, αλλά και στο γεγονός ότι οι χρήστες έχουν την ευκαιρία να δουν όλους τους υποψηφίους μαζί με τις πολιτικές τους θέσεις και να καταλάβουν πόσο μακριά ή πόσο κοντά τοποθετούνται οι δικές τους ιδέες με τις ιδέες και απόψεις των υποψηφίων. Αυτό μπορεί να προσφέρει στους χρήστες πρόσθετες πληροφορίες, πνευματική διέγερση και ψυχαγωγία, τα οποία μπορεί να οδηγήσουν σε αλλαγή ή ενίσχυση της ψήφου τους [54]. Ωστόσο, το κατά πόσο χρησιμοποιείται ή όχι ένα VAA εξαρτάται από το επίπεδο μόρφωσης του ατόμου, από το ενδιαφέρον του για την πολιτική και τη δυνατότητα (ή μη) πρόσβασης στο διαδίκτυο [45].

⁵<http://www.votematch.eu/>

⁶<http://www.nece.eu>

1.5 Σχεδιασμός Ηλεκτρονικών Συμβούλων Ψήφου

Σε γενικές γραμμές τα VAAs είναι διαδικτυακά (online) εργαλεία έρευνας, τα οποία συνιστούν έναν υποψήφιο ή ένα πολιτικό κόμμα σε κάποιο χρήστη, ανάλογα με το πόσο ταιριάζουν οι πολιτικές τους πεποιθήσεις [55, 56]. Για να συμβεί αυτό είναι απαραίτητο οι χρήστες και υποψήφιοι να απαντήσουν σε ένα online ερωτηματολόγιο, το οποίο βρίσκεται αναρτημένο στην ιστοσελίδα του VAA. Ο σχεδιασμός του ερωτηματολογίου, η απόφαση για τον αριθμό και τη θεματολογία των ερωτήσεων που θα το απαρτίζουν καθώς και οι επιλογές των απαντήσεων που θα δίνονται στο χρήστη για να εκφράσει την άποψη του είναι τα πρωταρχικά θέματα που απασχολούν στο σχεδιασμό ενός VAA. Ακολουθεί η εκτίμηση των απαντήσεων των πολιτικών κομμάτων στο βασικό ερωτηματολόγιο, η επιλογή των κομμάτων που θα συμπεριληφθούν στο σχεδιασμό του VAA, η δημιουργία της πρότασης ψήφου από το σύστημα και η γραφική αναπαράσταση των αποτελεσμάτων.

1.5.1 Επιλογή ερωτήσεων

Το βασικό ερωτηματολόγιο του VAA αποτελείται από ερωτήσεις που αφορούν σημαντικά πολιτικά, οικονομικά και κοινωνικά ζητήματα, τα οποία επιλέγονται και διατυπώνονται από ομάδες ειδικών. Τα θέματα αυτά είναι διαφορετικά για κάθε χώρα και σχετίζονται με τα τρέχοντα γεγονότα που απασχολούν το έθνος εν καιρώ εκλογών και ενδέχεται να συζητιούνται στις προεκλογικές εκστρατείες των υποψηφίων ή να συνάδουν με τις πολιτικές πεποιθήσεις των κομμάτων [57, 58].

Για την επιλογή και τη διαμόρφωση των ερωτήσεων ενός VAA, οι ειδικοί ασχολούνται με δύο σημαντικά ζητήματα. Το πρώτο είναι το περιεχόμενο των ερωτήσεων το οποίο πρέπει να είναι πλήρες, σαφές και εύκολο στην κατανόηση και απάντηση. Είναι σημαντικό να λαμβάνεται υπόψη ότι οι χρήστες των VAAs δεν χρειάζεται να έχουν εξειδικευμένες γνώσεις για να μπορέσουν να απαντήσουν στις ερωτήσεις [50]. Όπως επίσης, οι ερωτήσεις αυτές χρειάζεται να είναι συγκεκριμένες και όχι ασαφείς, μιας και οι δηλώσεις που είναι υπερβολικά ασαφείς μπορούν να οδηγήσουν σε προκατειλημμένες απαντήσεις λόγω παρερμηνείας και δεν παρέχουν χρήσιμες πληροφορίες στους ψηφοφόρους. Οι ψηφοφόροι μπορεί να έχουν μια γενική αίσθηση για το που τοποθετούνται ιδεολογικά τα κόμματα, αλλά αυτό που τους λείπει είναι η γνώση σχετικά με τις θέσεις των κομμάτων σε συγκεκριμένα πολιτικά ζητήματα. Για να αποφεύγεται η ερμηνεία των ερωτήσεων με διαφορετικό τρόπο από κάθε ψηφοφόρο και κόμμα, είναι εξίσου επιθυμητό οι ερωτήσεις να μην είναι διπλής σημασίας, αλλά να μπορούν να μετρήσουν μόνο ένα πράγμα τη φορά [59].

Επιπρόσθετα πρέπει να αποφεύγονται οι ερωτήσεις με ποσοτικό περιεχόμενο, αφού η χρήση λέξεων που κάνουν αναφορά σε αριθμούς ή ποσοτικά επιρρήματα όπως “αρκετά” και “λίγο”, μπορούν να οδηγήσουν σε προκατειλημμένες και αναξιόπιστες απαντήσεις λόγω παρερμηνείας [60]. Υπό τέτοιες συνθήκες, όταν οι ψηφοφόροι εκφράζουν την άποψη τους για το περιεχόμενο μιας ερώτησης, δεν είναι ξεκάθαρο σε τι ακριβώς αναφέρονται (π.χ. η διαφωνία του χρήστη στην ερώτηση “οι τιμωρίες προς τους κατ’ επανάληψη εγκληματίες πρέπει να είναι περισσότερο αυστηρές;” δεν είναι σαφές εάν σημαίνει πως ο χρήστης υποστηρίζει τα μέτρα της τρέχουσας κυβέρνησης ή πως πιστεύει πως τα μέτρα αυτά είναι πολύ αυστηρά). Έτσι, δεν μπορεί να εξακριβωθεί εάν οι ψηφοφόροι και τα κόμματα ταιριάζουν πραγματικά, ακόμα και όταν δίνουν την ίδια απάντηση. Για τον ίδιο λόγο πρέπει να αποφεύγονται οι επιπρόσθετες πληροφορίες σε μια ερώτηση, που μπορεί να στρέψουν την προσοχή του χρήστη αλλού και να αλλοιώσουν το περιεχόμενο της. Γενικά οι ερωτήσεις στο VAA πρέπει να είναι σαφείς, ακριβείς και συνοπτικές [59].

Το δεύτερο ζήτημα σχετικά με το βασικό ερωτηματολόγιο ενός VAA είναι η ποικιλία των θεμάτων. Τα VAAs τείνουν να περιλαμβάνουν στο ερωτηματολόγιο τους μια ευρεία ποικιλία θεμάτων, κάνοντας

με αυτό τον τρόπο ακόμα και τα δευτερεύοντα ζητήματα να λαμβάνουν την προσοχή του χρήστη αλλά και του υποψήφιου κόμματος. Ωστόσο, φαίνεται να εστιάζουν περισσότερο στις κοινωνικοοικονομικές ερωτήσεις της αριστεράς-δεξιάς κλίμακας, στις περιπτώσεις που η διάσταση αυτή είναι πιο σημαντική στη χώρα τους. Βέβαια, αυτό που τηρείται σε όλες τις περιπτώσεις είναι ο πολύ διακριτικός χαρακτήρας των ερωτήσεων [59].

Γενικά στα VAAs είναι σημαντικό να διασφαλίζεται πως οι ερωτήσεις κατανέμονται σε όλους τους τομείς των ζητημάτων που απασχολούν τη χώρα και ότι καλύπτουν τις τρέχουσες πολιτικές συζητήσεις, ειδικά αυτές που προσελκύουν το ενδιαφέρον των πολιτών. Κατά τη διαδικασία της επιλογής των θεμάτων που θα συμπεριληφθούν σε ένα VAA, ομάδα ειδικών μελετά το δημόσιο λόγο όπως αυτός προκύπτει από τα μέσα ενημέρωσης και συμβουλευεται το ευρύ κοινό, ακαδημαϊκούς ή δημοσιογράφους για να καταλήξει στους σημαντικότερους τομείς της πολιτικής. Το ερωτηματολόγιο ενός VAA συνήθως περιέχει από 20 έως 35 δηλώσεις πολιτικής (αυτό που ονομάσαμε νωρίτερα βασικό ερωτηματολόγιο), στις οποίες οι χρήστες ή τα υποψήφια κόμματα καλούνται να τοποθετηθούν χρησιμοποιώντας μια κλίμακα διαφορών βαθμίδων [46]. Οι δηλώσεις πολιτικής πρέπει να επιλέγονται έτσι ώστε να διαχωρίζουν τις θέσεις των κομμάτων και να προβάλλουν τη διαφορετική τους αντίληψη, και να καθιστούν εφικτή τη μεταξύ τους σύγκριση [58]. Επίσης δεν πρέπει να ευνοούν κάποιο από τα πολιτικά κόμματα.

Τα περισσότερα VAAs συνοδεύουν τις ερωτήσεις τους (δηλώσεις πολιτικής) με μια κλίμακα Likert, από την οποία οι ερωτηθέντες επιλέγουν τη βαθμίδα που εκφράζει καλύτερα τη γνώμη τους. Η προσέγγιση αυτή αναπτύχθηκε από τον Rensis Likert για τη μέτρηση της ψυχολογικής συμπεριφοράς στις αρχές της δεκαετίας του 1930 [60]. Ο αριθμός των βαθμίδων που χρησιμοποιεί το κάθε VAA είναι διαφορετικός. Όταν αποτελείται από πέντε βαθμίδες δίνεται η ευκαιρία στον ερωτηθέντα να εκφράσει καλύτερα την άποψη του, είτε αυτή αναφέρεται σε ουδέτερη, μεσαία ή ακραία συμπεριφορά [61]. Η πενταβάθμια κλίμακα Likert είναι και αυτή που υιοθετούν πολλά από τα VAAs όπως το *EU profiler* και το *EUVoX*, τα VAAs της *Kieskompas*⁷ και της ομάδας του Πανεπιστημίου της Ζυρίχης [50] και όλα τα VAAs που σχεδιάστηκαν από την ερευνητική ομάδα του *PreferenceMatcher*⁸.

Με την πενταβάθμια κλίμακα οι χρήστες και τα κόμματα μπορούν να τοποθετηθούν στις δηλώσεις πολιτικής του VAA ερωτηματολογίου απαντώντας κατά πόσο συμφωνούν ή διαφωνούν με αυτές. Πίσω από κάθε σημείο της κλίμακας βρίσκονται παραλλαγές κατά μήκος των απαντήσεων “διαφωνώ πλήρως”, “τείνω να διαφωνώ”, “ουδέτερος”, “τείνω να συμφωνώ”, “συμφωνώ πλήρως”, δίνοντας με αυτές τις επιλογές την ευκαιρία στο χρήστη να προσδιορίσει την κατεύθυνση (“διαφωνώ” / “συμφωνώ”) και την ένταση (“πλήρως” / “τείνω”) της άποψής του [60]. Το *Smartvote*⁹ της Ελβετίας αν και ακολουθεί τη λογική της πενταβάθμιας κλίμακας, δε δίνει τη δυνατότητα στον ερωτηθέντα να διαλέξει τη στάση “ουδέτερος” αλλά του δίνει τις επιλογές “Ναι”, “Μάλλον Ναι”, “Μάλλον όχι”, “Όχι” ή “Καμία απάντηση” επιτρέποντας του να σταθμίσει την κάθε ερώτηση ξεχωριστά.

Υπάρχουν και VAAs που παραβλέπουν την ένταση στις ερωτήσεις τους και χρησιμοποιούν κλίμακες τριών βαθμίδων. Το *Wahl-O-Mat*¹⁰ της Γερμανίας επιτρέπει στους χρήστες να απαντήσουν στις ερωτήσεις του ερωτηματολογίου επιλέγοντας μια από τις κατηγορίες “Συμφωνώ”, “Ουδέτερος” ή “Διαφωνώ”. Στη συνέχεια ο χρήστης μπορεί να επιλέξει ποια θέματα είναι ιδιαίτερα σημαντικά για εκείνον, έτσι ώστε να προστεθεί βάρος σε αυτές τις δηλώσεις κατά τον υπολογισμό της πρότασης ψήφου από το σύστημα. Στο *Wahlkabine*¹¹ της Αυστρίας, οι επιλογές που δίνονται είναι “Ναι”, “Όχι” ή “Δεν γνωρίζω” και ταυτόχρονα με τις απαντήσεις του ο χρήστης βαθμολογεί το πόσο σημαντικό είναι το κάθε θέμα για τον ίδιο σε κλίμακα εννέα βαθμίδων, η οποία αρχίζει από το “όχι τόσο σημαντικό” και κατευθύνεται προς το “πολύ

⁷<https://home.kieskompas.nl/>

⁸<http://www.preferencematcher.org/>

⁹<http://smartvote.ch/>

¹⁰<http://wahl-o-mat.de/>

¹¹<http://wahlkabine.at/>

The screenshot shows the EUVox 2014 survey interface. At the top, there is a logo for EUVox 2014 and navigation links for 'Σχετικά με εμάς' and 'Συχνές ερωτήσεις'. Below this is a progress bar with a '1' in a blue circle. The main heading is 'ΕΥΡΩΠΑΪΚΗ ΈΝΩΣΗ' followed by the statement 'Η Κύπρος θα πρέπει να εγκαταλείψει το Ευρώ'. Below the statement are five buttons representing the Likert scale: 'Συμφωνώ απόλυτα', 'Συμφωνώ', 'Ούτε συμφωνώ, ούτε διαφωνώ', 'Διαφωνώ', and 'Διαφωνώ απόλυτα'. A sixth button, 'Δεν έχω άποψη', is positioned below the other five.

Σχήμα 1.1: Παράδειγμα δήλωσης πολιτικής από το EUVox 2014

σημαντικό”.

Στην κατηγορία των VAAs που χρησιμοποιούν κλίμακες με τρεις βαθμίδες για την ανταπόκριση του χρήστη, βρίσκεται και το *StemWijzer*¹² της Ολλανδίας με τις επιλογές “Συμφωνώ”, “Διαφωνώ” ή “Ούτε συμφωνώ, ούτε διαφωνώ”. Στην επίσημη ιστοσελίδα του αναφέρει “Το StemWijzer επιλέγει τις συγκεκριμένες κατηγορίες ανταπόκρισης γιατί είναι πιο ξεκάθαρες προς τους χρήστες. Επιπλέον, αυτό σχετίζεται με το πώς λειτουργεί η δημοκρατική διαδικασία λήψης αποφάσεων. Το λίγο πολύ σύντομα ή όχι, δεν είναι σημαντικό. Άλλοι σύμβουλοι ψήφου εξακολουθούν να χρησιμοποιούν μια πενταβάθμια κλίμακα (για παράδειγμα “διαφωνώ πλήρως”). Ως αποτέλεσμα, ο χρήστης πρέπει να δίνει δύο απαντήσεις την ίδια στιγμή: τη γνώμη και την ένταση της γνώμης του. Το *StemWijzer* επιλέγει να διαχωρίσει το ένα από το άλλο. Πόσο σημαντικό είναι ένα συγκεκριμένο θέμα μπορεί να δηλωθεί μετά την απάντηση των ερωτήσεων.”

Εκτός από το βασικό ερωτηματολόγιο που αποσκοπεί στην εύρεση της συνάφειας μεταξύ του χρήστη και του κόμματος, σε κάποιες περιπτώσεις οι χρήστες καλούνται να ανταποκριθούν και σε μια σειρά από συμπληρωματικές ερωτήσεις (όπως είναι τα κριτήρια με τα οποία επιλέγει το κόμμα ή τον υποψήφιο που θα υποστηρίξει στις επερχόμενες εκλογές, το ενδιαφέρον τους για την πολιτική, η πρόθεση ψήφου τους στις επικείμενες εκλογές). Οι απαντήσεις των χρηστών στις συγκεκριμένες ερωτήσεις είναι προαιρετικές και μπορούν να αξιοποιηθούν για ερευνητικούς σκοπούς ή για τις επιπλέον δυνατότητες στην παρουσίαση των αποτελεσμάτων [62].

Στο Σχήμα 1.1 φαίνεται ένα παράδειγμα δήλωσης πολιτικής από το EUVox και η πενταβάθμια κλίμακα απαντήσεων που δόθηκε. Σημειώστε ότι η έκτη επιλογή (“Δεν έχω άποψη”) εμφανίζεται κάτω από τις επιλογές της πενταβάθμιας κλίμακας καθώς δεν εκφράζει ένταση άποψης.

Ένα παράδειγμα συμπληρωματικών ερωτήσεων, κυρίως δημογραφικής υφής, από το EUVox 2014 φαίνεται στο Σχήμα 1.2. Η τελευταία ερώτηση αφορά το ενδιαφέρον για την πολιτική και είναι μια τυπική ερώτηση που χρησιμοποιείται σε δημοσκοπήσεις.

¹²<https://stemwijzer.nl/>

Σχήμα 1.2: Παράδειγμα συμπληρωματικών ερωτήσεων δημογραφικού, κυρίως, περιεχομένου από το EUVox 2014

1.5.2 Κωδικοποίηση των θέσεων των κομμάτων

Ένα ακόμα σημαντικό κομμάτι στο σχεδιασμό του VAA είναι η επιλογή των πολιτικών κομμάτων που θα εντάξει στο σχεδιασμό του και η κωδικοποίηση των θέσεων τους στα ζητήματα του βασικού ερωτηματολογίου. Κάποια VAAs τείνουν να περιλαμβάνουν όλα τα κόμματα που λαμβάνουν μέρος στις επικείμενες εκλογές, όπως γίνεται με το *StemWijzer* και το *Wahl-O-Mat*, ενώ άλλα επιλέγουν μερικά από αυτά [51]. Στην επίσημη ιστοσελίδα του EUvox¹³ σχετικά με τον αριθμό των κομμάτων που περιλαμβάνονται, αναφέρεται το εξής: “Τόσο τεχνικοί λόγοι, όσο και η έλλειψη πληροφοριών σχετικά με κάποια κόμματα περιορίζουν τον αριθμό των κομμάτων που περιλαμβάνονται. Τα κύρια πολιτικά κόμματα κάθε χώρας, συμπεριλαμβανομένων όσων ήδη βρίσκονται εντός της Ευρωβουλής και όσων είναι πιθανό να λάβουν κάποια από τις έδρες του Ευρωκοινοβουλίου στις επερχόμενες εκλογές σύμφωνα με τα γκάλοπ, περιλαμβάνονται στο EUvox”.

Οι θέσεις των κομμάτων ή υποψηφίων στις δηλώσεις πολιτικής του βασικού ερωτηματολογίου εκτιμώνται ρωτώντας απευθείας τα κόμματα ή προσλαμβάνοντας ειδικούς ερευνητές που απαντούν εκ μέρους τους στα θέματα αυτά αφού πρώτα αναλύσουν τους λόγους που έχουν εκφωνήσει ή τα έγγραφα που έχουν δημοσιεύσει κατά καιρούς οι υποψήφιοι ή οι εκπρόσωποι των κομμάτων, έτσι ώστε να είναι σε θέση να προβλέψουν με ακρίβεια τι θα απαντούσαν οι ίδιοι [38, 63, 57].

Όταν τα VAAs απευθύνονται στα κόμματα ή στους υποψηφίους τους για να δηλώσουν οι ίδιοι τις θέσεις τους στο βασικό ερωτηματολόγιο, η απλούστερη διαδικασία που ακολουθείται είναι η αποστολή του ερωτηματολογίου στους εκπροσώπους των κομμάτων - υποψηφίων για τη συμπλήρωση του. Κατά τη συμπλήρωση του ερωτηματολογίου, κάποια από τα VAAs ζητούν από τα κόμματα στοιχεία τεκμηρίωσης για τις απαντήσεις τους (π.χ. αναφορές στο πρόγραμμά τους, σε ομιλίες στελεχών / υποψηφίων, συνεντεύξεις, κοκ), ενώ κάποια άλλα αρκούνται με την απλή δικαιολόγηση της θέσης τους. Η τεκμηρίωση των θέσεων των κομμάτων θεωρείται απαραίτητη, έτσι ώστε να αποφεύγονται οι ανακριβείς απαντήσεις από την πλευρά τους [64, 46]. Αυτή η μεθοδολογία κωδικοποίησης των θέσεων των κομμάτων υιοθετείται από πολλά VAAs, όπως το *StemWijzer* της Ολλανδίας, το *VoteMatch* του Ηνωμένου Βασιλείου, το *Wahl-O-Mat* της Γερμανίας, το *Smartvote* της Ελβετίας, το *Vaalikone* της Φινλανδίας και το *Manobalsas* της

¹³<http://www.euvox.eu/>

Λιθουανίας [64]. Το βασικό της πλεονέκτημα είναι πως τα κόμματα καλούνται να αυτοτοποθετηθούν στα ζητήματα του ερωτηματολογίου με τον ίδιο τρόπο που καλούνται να τοποθετήσουν τον εαυτό τους οι χρήστες μετά την ανάρτηση του συγκεκριμένου ερωτηματολογίου στην ιστοσελίδα του VAA [60].

Ο υπολογισμός των θέσεων των κομμάτων στο βασικό ερωτηματολόγιο του VAA με την άμεση εμπλοκή τους, εμπεριέχει πολλές δυσκολίες. Αρχικά υπάρχουν περιπτώσεις όπου παρατηρείται έλλειψη συνεργασίας από την πλευρά των κομμάτων [62], είτε γιατί δεν ενδιαφέρονται για τα VAAs ή επειδή δε συμφωνούν με τη διατύπωση των ερωτήσεων [60]. Τα κόμματα ενώ είναι πρόθυμα να δηλώσουν τις θέσεις τους για θέματα που είναι προς το συμφέρον τους, είναι λιγότερο πιθανόν να ανταποκριθούν σε αυτά που είναι αμφιλεγόμενα και θεωρούν πως είναι εκλογικά επιβλαβή [64]. Επίσης, πολλά κόμματα συνηθίζουν να τοποθετούνται στη μέση της κλίμακας των απαντήσεων (αποφεύγοντας να πάρουν ξεκάθαρη θέση) έτσι ώστε να συμφωνούν με τους ψηφοφόρους και από τις δύο πλευρές του πολιτικού φάσματος [50], ενώ άλλα παρέχουν στρατηγικές απαντήσεις για να ταιριάζουν με το μέγιστο αριθμό χρηστών και να χειραγωγούν με αυτόν τον τρόπο τις προτάσεις ψήφου που γίνονται από τα VAAs [64, 46].

Για την αντιμετώπιση των πιο πάνω προβλημάτων, όσον αφορά την κωδικοποίηση των θέσεων των πολιτικών κομμάτων στο ερωτηματολόγιο του VAA, πολλές φορές χρησιμοποιείται μια μέθοδος που ουσιαστικά συνδυάζει τις δύο προαναφερθείσες μεθοδολογίες. Στη μέθοδο αυτή το ερωτηματολόγιο αποστέλλεται στα κόμματα για την αυτοτοποθέτηση τους και την τεκμηρίωση των απαντήσεων τους, ενώ παράλληλα μια μικρή ομάδα ειδικών καλείται να τοποθετήσει τα κόμματα βάσει των δημόσιων δηλώσεών τους. Με το συγκεκριμένο τρόπο, οι δύο τοποθετήσεις μπορούν να συγκριθούν μεταξύ τους για να οδηγήσουν σε πιο αξιόπιστα αποτελέσματα. Όταν οι θέσεις των κομμάτων δεν είναι οι ίδιες με αυτές που υπολόγισαν οι ειδικοί, τότε ζητείται από τα κόμματα να εξετάσουν ξανά την αρχική τοποθέτησή τους. Η διαδικασία επαναλαμβάνεται, μέχρι να υπάρξει ένα ικανοποιητικό ποσοστό συμφωνίας μεταξύ ειδικών και κομμάτων. Εάν αυτό δεν επιτευχθεί, τότε υπερισχύει η γνώμη των ειδικών [64].

Ο συνδυασμός μεταξύ της αυτοτοποθέτησης των κομμάτων και της τοποθέτησης τους από μια ομάδα ειδικών, εξακολουθεί να παρουσιάζει δύο σημαντικά προβλήματα. Το πρώτο είναι πως και πάλι απαιτείται η συνεργασία των κομμάτων. Το δεύτερο είναι ο τρόπος με τον οποίο η ομάδα ειδικών φτάνει στη συναίνεση κατά τη διαδικασία εκτίμησης των θέσεων των πολιτικών κομμάτων και οι διαφωνίες που προκύπτουν μεταξύ τους [64]. Για τους λόγους αυτούς κάποια VAAs επιλέγουν να υπολογίσουν τις θέσεις των κομμάτων με ένα άλλο τρόπο, έτσι ώστε να αποφευχθούν τα συγκεκριμένα προβλήματα. Οι ερευνητές της κοινοπραξίας του *PreferenceMatcher* παρουσίασαν μια νέα μέθοδο στην οποία η εκτίμηση των θέσεων των κομμάτων από ειδικούς βάσει τεκμηρίωσης πραγματοποιείται σε πολλαπλούς γύρους κωδικοποίησης [62]. Η μέθοδος αυτή είναι γνωστή ως *Delphi* και έχει εφαρμοστεί στα VAAs της ερευνητικής ομάδας του *PreferenceMatcher*.

Η διαδικασία που ακολουθείται κατά τη μέθοδο *Delphi* στην απλούστερη της μορφή περιγράφεται ως εξής:

- Ένα αρμόδιο άτομο αναλαμβάνει το ρόλο του συντονιστή και επιλέγει μια ομάδα εμπειρογνομόνων για να εκτιμήσουν τις θέσεις των κομμάτων στα διάφορα ζητήματα του ερωτηματολογίου του VAA και να τεκμηριώσουν τις απόψεις τους παρέχοντας σχετικά στοιχεία. Οι ειδικοί αυτοί εργάζονται ανεξάρτητα ο ένας από τον άλλο.
- Όταν ολοκληρώσουν τις εκτιμήσεις τους, ο συντονιστής συλλέγει τους υπολογισμούς και τις σχετικές πληροφορίες από τον καθένα τους και τις τροφοδοτεί ανώνυμα πίσω στην ομάδα για νέο γύρο κωδικοποίησης.
- Οι ειδικοί τώρα καλούνται να ενημερώσουν τις αρχικές τους εκτιμήσεις με βάση τις νέες πληροφορίες και η διαδικασία επαναλαμβάνεται έως ότου να επιτευχθεί επαρκής συναίνεση μεταξύ της

ομάδας.

- Τέλος οι απαντήσεις συναθροίζονται για την αποφυγή τυχαίων σφαλμάτων και καθορίζονται οι τελικές εκτιμήσεις.

Η διαδικασία αυτή είναι αρκετά χρονοβόρα και κοστίζει αρκετά, αφού οι ειδικοί πρέπει να αμείβονται για τη συμμετοχή τους. Αυτό είναι και το κύριο μειονέκτημα της μεθόδου [64].

1.5.3 Δημιουργία της πρότασης ψήφου

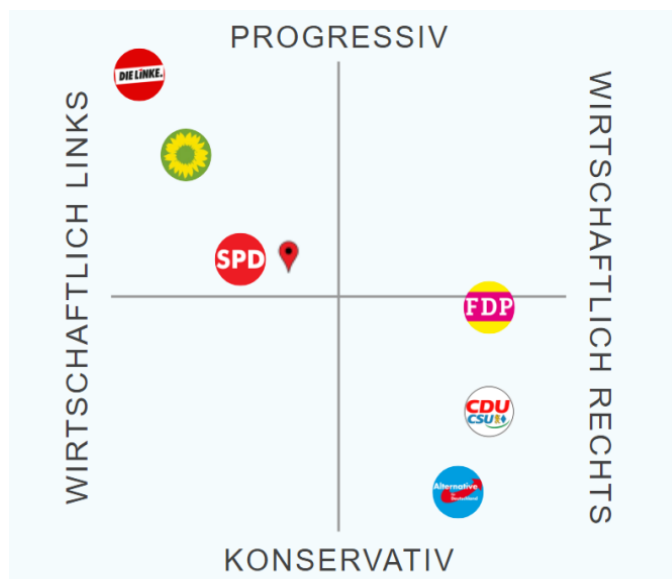
Η δημιουργία της πρότασης ψήφου μέσω ενός VAA στηρίζεται στις απαντήσεις των χρηστών και των κομμάτων στο βασικό ερωτηματολόγιο. Αφού πραγματοποιηθεί η κωδικοποίηση των απαντήσεων των κομμάτων στο εν λόγω ερωτηματολόγιο, με τους τρόπους που έχουν αναφερθεί προηγουμένως, οι πολίτες καλούνται να ανταποκριθούν στα ίδια ζητήματα με την επίσκεψή τους στην ιστοσελίδα του VAA. Όπως ήδη έχει αναφερθεί, οι απαντήσεις στο συγκεκριμένο ερωτηματολόγιο συνήθως δίνονται επιλέγοντας μια από τις ακόλουθες δηλώσεις: “διαφωνώ απόλυτα”, “διαφωνώ”, “ούτε συμφωνώ ούτε διαφωνώ”, “συμφωνώ”, “συμφωνώ απόλυτα”. Επιπρόσθετα ο χρήστης έχει το δικαίωμα να μην δώσει απάντηση σε κάποια ερώτηση διαλέγοντας την επιλογή “δεν έχω άποψη”. Έχοντας τις απαντήσεις των κομμάτων και των χρηστών, το VAA μπορεί να συνεχίσει στη δημιουργία της πρότασης ψήφου.

Αφού ο χρήστης απαντήσει στο βασικό ερωτηματολόγιο του VAA, ακολουθεί η εκτίμηση της συνάφειας του με τα πολιτικά κόμματα ή υποψηφίους που λαμβάνουν μέρος στο σχεδιασμό του. Στα παραδοσιακά VAAs η πρόταση ψήφου γίνεται με τη βοήθεια ενός κατάλληλα σχεδιασμένου αλγόριθμου, ο οποίος συγκρίνει τις απαντήσεις των πολιτών με τις κωδικοποιημένες θέσεις των κομμάτων/υποψηφίων και στο τέλος κάθε χρήστης συνιστάται να δώσει ψήφο στον υποψήφιο ή το κόμμα που έχει τον υψηλότερο βαθμό ταύτισης μαζί του σύμφωνα με κάποια μετρική. Συνήθως ο χρήστης ενημερώνεται, επίσης, με την τοποθέτησή του σε ένα ιδεολογικό χάρτη, ο οποίος περιγράφει πόσο κοντά ή πόσο μακριά βρίσκεται από το κάθε κόμμα / υποψήφιο στη βάση δύο αξόνων πολιτικής [55] (βλέπε Σχήμα 1.3).

Οι τεχνικές που εφαρμόζουν τα VAAs για τον υπολογισμό της συνάφειας μεταξύ του χρήστη και των κομμάτων, μπορούν να χωριστούν σε δύο κατηγορίες [65]. Στην πρώτη η ταύτιση γίνεται χρησιμοποιώντας όλες τις ερωτήσεις (δηλώσεις πολιτικής) του βασικού ερωτηματολογίου και ονομάζεται *ταύτιση υψηλών διαστάσεων*. Στη δεύτερη η συσχέτιση του χρήστη με τα κόμματα γίνεται χρησιμοποιώντας ένα διάγραμμα διασποράς, που συνήθως περιλαμβάνει δύο διαστάσεις: η μια αφορά τον κοινωνικό φιλελευθερισμό έναντι του κοινωνικού συντηρητισμού και η άλλη σχετίζεται με την οικονομική αριστερά έναντι της οικονομικής δεξιάς πολιτικής. Η τεχνική αυτή αναφέρεται ως *ταύτιση χαμηλών διαστάσεων*.

Ταύτιση υψηλών διαστάσεων

Το *StemWijzer* στηρίζεται σε μια πολύ απλή μέθοδο. Κάθε απάντηση του κόμματος στο βασικό ερωτηματολόγιο που είναι όμοια με αυτήν που έδωσε ο χρήστης, παίρνει ένα βαθμό. Εάν η απάντηση αυτή σχετίζεται με τα ζητήματα, στα οποία ο χρήστης επέλεξε να δώσει περισσότερη βαρύτητα, τότε δίνονται δύο βαθμοί. Για όλες τις υπόλοιπες διαφορετικές απαντήσεις μεταξύ των κομμάτων και του χρήστη δε προστίθεται κανένας βαθμός. Το κόμμα που ταιριάζει καλύτερα με το χρήστη είναι αυτό που συγκεντρώνει τους περισσότερους βαθμούς. Στο τέλος τα αποτελέσματα παρουσιάζονται στην οθόνη του χρήστη γραφικά με ραβδόγραμμα, όπου το κόμμα με τη μεγαλύτερη συνάφεια βρίσκεται στην κορυφή και τα υπόλοιπα ακολουθούν σε φθίνουσα σειρά. Ένα παράδειγμα για το πώς εμφανίζονται τα αποτελέσματα του *StemWijzer* στην οθόνη του χρήστη, παρατίθεται στο Σχήμα 1.4.



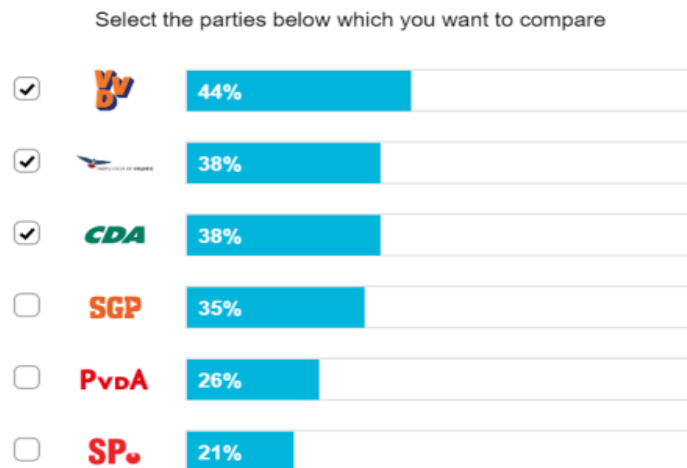
Σχήμα 1.3: Παράδειγμα της γραφικής απεικόνισης των αποτελεσμάτων από το *Kieskompas* για τις ομοσπονδιακές εκλογές της 24ης Σεπτεμβρίου 2017 στη Γερμανία. Η σταγόνα με την ενσωματωμένη κουκκίδα αναπαριστά την τοποθέτηση του χρήστη από το VAA με βάση τις δηλώσεις του στις δηλώσεις πολιτικής

Η μέθοδος ταύτισης που ακολουθείται από το *StemWijzer* χαρακτηρίζεται από απλότητα και ευκολία στην κατανόηση αλλά ο τρόπος υπολογισμού της συνάφειας χρήστη - κόμματος είναι ευριστικός και δεν παραπέμπει σε κάποια μετρική [66]. Ο Mendez ([1]) διέκρινε τέσσερις τρόπους με τους οποίους μπορεί να υπολογιστεί η συνάφεια μεταξύ του χρήστη και των κομμάτων, όταν εφαρμόζεται ταύτιση υψηλών διαστάσεων. Στους Πίνακες 1.1 – 1.4 φαίνονται οι τρόποι αυτοί, όπως εμφανίζονται στην εργασία του Mendez [1]. Οι συντομογραφίες ΔΠ, Δ, ΟΟ, Σ, ΣΠ που βρίσκονται στις γραμμές και τις στήλες αναφέρονται στις αντίστοιχες τοποθετήσεις “Διαφωνώ Πλήρως”, “Διαφωνώ”, “Ούτε Διαφωνώ Ούτε Συμφωνώ”, “Συμφωνώ”, “Συμφωνώ Πλήρως” που δίνονται από το χρήστη και το κόμμα στις δηλώσεις πολιτικής του βασικού ερωτηματολογίου. Τα κελιά δείχνουν το βάρος με το οποίο σταθμίζεται η κάθε περίπτωση που μπορεί να προκύψει κατά τη σύγκριση μεταξύ των τοποθετήσεων του χρήστη και του κόμματος.

Πίνακας 1.1: Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει της City-block (Manhattan) απόστασης

Τοποθέτηση χρήστη Τοποθέτηση κόμματος	ΔΠ	Δ	ΟΟ	Σ	ΣΠ
ΔΠ	1	0.5	0	-0.5	-1
Δ	0.5	1	0.5	0	-0.5
ΟΟ	0	0.5	1	0.5	0
Σ	-0.5	0	0.5	1	0.5
ΣΠ	-1	-0.5	0	0.5	1

Τα περισσότερα VAAs υιοθετούν το αντίθετο της City-block (Manhattan) απόστασης (βλέπε Πίνακα 1.1) ως μετρική για τον υπολογισμό της συνάφειας μεταξύ χρήστη και κόμματος. Κάποια άλλα χρησιμοποιούν το αντίθετο της Ευκλείδειας απόστασης (βλέπε Πίνακα 1.2) [65]. Ο υπολογισμός των τιμών σε κάθε κελί των Πινάκων 1.1 – 1.4 έγινε στη βάση μιας κλίμακας Likert πέντε σημείων με τις ακόλουθες τιμές ΔΠ = 1, Δ = 2, ΟΟ = 3, Σ = 4 και ΣΠ = 5. Η απόσταση μεταξύ των τοποθετήσεων του χρήστη και του κόμματος,



Σχήμα 1.4: Παράδειγμα της γραφικής απεικόνισης των αποτελεσμάτων του *StemWijzer* για τις κοινοβουλευτικές εκλογές της 15ης Μαρτίου 2017 στην Ολλανδία

Πίνακας 1.2: Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει της Ευκλείδειας απόστασης

Τοποθέτηση χρήστη Τοποθέτηση κόμματος	ΔΠ	Δ	ΟΟ	Σ	ΣΠ
ΔΠ	1	0.5	0	-0.5	-1
Δ	0.5	0.25	0	-0.25	-0.5
ΟΟ	0	0	0	0	0
Σ	-0.5	-0.25	0	0.25	0.5
ΣΠ	-1	-0.5	0	0.5	1

στην περίπτωση της City-block απόστασης, υπολογίστηκε ως η απόλυτη τιμή της αφαίρεσης των δύο σημείων (για παράδειγμα η διαφορά μεταξύ της ΔΠ και Σ είναι $4 - 1 = 3$), ενώ κατά την Ευκλείδεια απόσταση, εκτιμήθηκε το τετράγωνο της κάθε απόστασης (για παράδειγμα η διαφορά μεταξύ της ΔΠ και Σ τώρα είναι το τετράγωνο της τιμής $(4 - 1)^2 = 3^2 = 9$). Στη συνέχεια οι τιμές αυτές αναπροσαρμόστηκαν στους Πίνακες ώστε να κυμαίνονται στο διάστημα $[-1, 1]$ [1].

Όπως αναφέρθηκε ήδη, στον Πίνακα 1.1 η απόσταση μεταξύ των απαντήσεων ενός χρήστη και ενός κόμματος υπολογίζεται με τη City-block (Manhattan) απόσταση¹⁴. Όταν για παράδειγμα ο χρήστης τοποθετείται σε μια ερώτηση με το “Διαφωνώ Πλήρως” ενώ το κόμμα δηλώνει “Συμφωνώ”, τότε αφαιρούνται 0.5 μονάδες (βλέπε γραμμή 4 - στήλη 1 στον Πίνακα 1.1). Στην περίπτωση που ο χρήστης δηλώνει “Διαφωνώ Πλήρως” και το κόμμα “Διαφωνώ” (βλέπε γραμμή 4 - στήλη 3 στον Πίνακα 1.1), τότε προστίθεται 0.5 μονάδες. Όταν ο χρήστης και το κόμμα επιλέγουν ακριβώς την ίδια τοποθέτηση τότε προστίθεται μία μονάδα. Η διαδικασία συνεχίζεται μέχρι να σταθμιστούν όλες οι απαντήσεις του χρήστη σύμφωνα με τις απαντήσεις του κόμματος, με βάση τις τιμές του πίνακα. Ο συντελεστής συνάφειας για την κάθε ερώτηση κυμαίνεται από το -1 (όταν, δηλαδή, υπάρχει πλήρης διαφωνία μεταξύ χρήστη και κόμματος) μέχρι το $+1$ (όταν, δηλαδή, υπάρχει πλήρης συμφωνία μεταξύ χρήστη και κόμματος). Ο συνολικός δείκτης συνάφειας μεταξύ χρήστη και κόμματος προκύπτει ως ο μέσος όρος της συνάφειας για όλες τις

¹⁴https://en.wikipedia.org/wiki/City_block

δηλώσεις πολιτικής του ερωτηματολογίου (άθροισμα βαθμών συνάφειας διά του συνολικού αριθμού των δηλώσεων πολιτικής).

Πίνακας 1.3: Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει της θεωρίας της εκλογικής συμπεριφοράς των Rabinowitz & MacDonald [2]

Τοποθέτηση χρήστη Τοποθέτηση κόμματος	ΔΠ	Δ	ΟΟ	Σ	ΣΠ
ΔΠ	1	0.5	0	-0.5	-1
Δ	0.5	0.25	0	-0.25	-0.5
ΟΟ	0	0	0	0	0
Σ	-0.5	-0.25	0	0.25	0.5
ΣΠ	-1	-0.5	0	0.5	1

Πίνακας 1.4: Υπολογισμός της συνάφειας μεταξύ χρήστη και κομμάτων όπως παρουσιάζονται στη δουλειά του Mendez [1] βάσει του συνδυασμού της City-block (Manhattan) απόστασης της και της θεωρίας της εκλογικής συμπεριφοράς των Rabinowitz & MacDonald [2]

Τοποθέτηση χρήστη Τοποθέτηση κόμματος	ΔΠ	Δ	ΟΟ	Σ	ΣΠ
ΔΠ	1	0.5	0	-0.5	-1
Δ	0.5	0.625	0.25	-0.125	-0.5
ΟΟ	0	0.25	0.5	0.25	0
Σ	-0.5	-0.125	0.25	0.625	0.5
ΣΠ	-1	-0.5	0	0.5	1

Τόσο η Ευκλείδεια απόσταση¹⁵ όσο και η απόσταση Manhattan είναι συνηθισμένες μετρικές εύρεσης της απόστασης δύο σημείων σε έναν πολυδιάστατο χώρο. Ιδιαίτερο ενδιαφέρον παρουσιάζουν τεχνικές συνάφειας οι οποίες είναι εστιασμένες στις ιδιαιτερότητες των Ηλεκτρονικών Συμβούλων Ψήφου. Στον Πίνακα 1.3 οι συντελεστές συνάφειας μεταξύ χρήστη και κόμματος, για κάθε δήλωση πολιτικής, εκτιμήθηκαν βάσει της θεωρίας της εκλογικής συμπεριφοράς των Rabinowitz & MacDonald [2]. Οι τιμές σε αυτήν την περίπτωση δεν βασίζονται στην απόσταση μεταξύ χρήστη και κόμματος αλλά στο κατά πόσο τοποθετούνται στην ίδια πλευρά της πολιτικής τοποθέτησης [1, 65]. Έτσι εάν ο χρήστης και το κόμμα δηλώνουν και οι δύο πως διαφωνούν ή συμφωνούν σε μια απάντηση, τότε ο συντελεστής συνάφειας παίρνει θετική τιμή. Στην αντίθετη περίπτωση που ο ένας από τους δύο δηλώνει πως συμφωνεί ενώ ο άλλος διαφωνεί στην ίδια ερώτηση, τότε ο συντελεστής συνάφειας γίνεται αρνητικός. Εάν τουλάχιστον ένας από τους δύο απαντήσει “Ούτε Διαφωνώ Ούτε Συμφωνώ”, ο συντελεστής ομοιότητας παίρνει την τιμή μηδέν.

Ο Πίνακας 1.4 παρουσιάζει μια εναλλακτική πρόταση υπολογισμού της συνάφειας όπως προτάθηκε από τον Mendez [1] στην οποία συνδυάζονται οι τεχνική υπολογισμού συνάφειας με βάση την απόσταση City-block και τη θεωρία της εκλογικής συμπεριφοράς. Οι τιμές στα κελιά του Πίνακα 1.4 προκύπτουν από το μέσο όρο των αντίστοιχων τιμών στα κελιά των Πινάκων 1.1 και 1.3. Με τον τρόπο αυτό λαμβάνεται υπόψη, έστω και με χαμηλή βαρύτητα, τόσο η ταύτιση στη μεσαία κατηγορία όσο και η ένταση των

¹⁵https://en.wikipedia.org/wiki/Euclidean_distance

τοποθετήσεων. Όσο πιο έντονη είναι η τοποθέτηση του χρήστη σε μια δήλωση πολιτικής (εάν δηλαδή συμφωνεί ή διαφωνεί πλήρως), τόσο μεγαλύτερη είναι και η τιμή του συντελεστή συνάφειας για αυτή.

Ταύτιση χαμηλών διαστάσεων

Πολλά VAAs επιλέγουν μια διαφορετική τεχνική από αυτήν της ταύτισης υψηλών διαστάσεων, η οποία αναφέρεται ως ταύτιση χαμηλών διαστάσεων και ουσιαστικά πρόκειται για την τοποθέτηση του χρήστη και των κομμάτων σε ένα πολιτικό χάρτη [43, 65]. Το *Kieskompas*, που ακολουθεί τη συγκεκριμένη μέθοδο, τοποθετεί το χρήστη και τα κόμματα σε ένα δισδιάστατο πολιτικό χάρτη, στον οποίο το κόμμα που βρίσκεται πιο κοντά στο χρήστη είναι και αυτό που του ταιριάζει καλύτερα. Ο πολιτικός αυτός χάρτης αποτελείται από ένα οριζόντιο άξονα που αναφέρεται στην αριστερή-δεξιά πολιτική τοποθέτηση του χρήστη και του κόμματος σε οικονομικά ζητήματα, έναντι ενός κατακόρυφου άξονα που αντιστοιχεί στην προοδευτική-συντηρητική τοποθέτηση τους σε θέματα που αφορούν την πολιτική-κουλτούρα. Ένα παράδειγμα της παρουσίασης των αποτελεσμάτων από το *Kieskompas* δίνεται στο Σχήμα 1.3. Με αυτόν τον τρόπο δίνεται η δυνατότητα του χρήστη να αντιληφθεί τις πολιτικές θέσεις των κομμάτων και τις διαφορές μεταξύ τους, σε σχέση με ένα δισδιάστατο χώρο πολιτικής αντιπαράθεσης. Σε αυτές τις περιπτώσεις η συνάφεια μεταξύ του χρήστη και των κομμάτων μπορεί να υπολογιστεί με την Ευκλείδεια απόσταση [62].

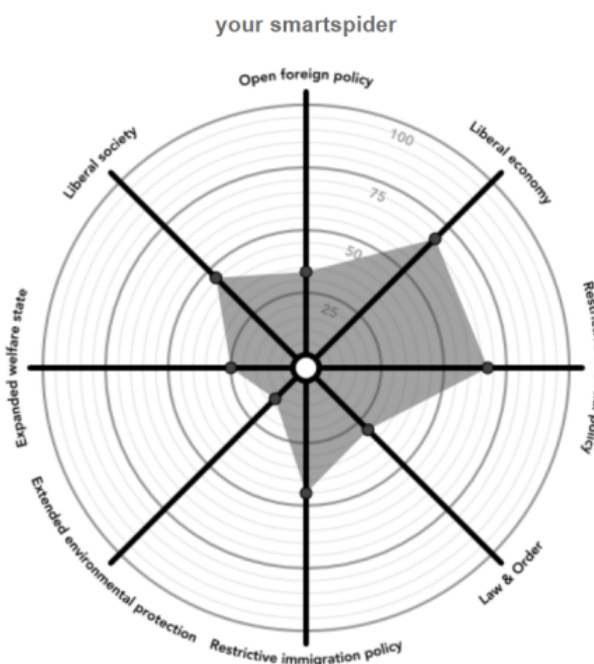
Ο υπολογισμός των συντεταγμένων, όταν εφαρμόζεται η ταύτιση χαμηλών διαστάσεων στηρίζεται στη λογική της *παραγοντικής ανάλυσης*¹⁶ με βάση όμως προεπιλεγμένους άξονες. Η διαδικασία συνοψίζεται ως εξής [65]: Οι δηλώσεις πολιτικής από το βασικό ερωτηματολόγιο του VAA τοποθετούνται σε δύο άξονες x και y . Συνήθως ο άξονας x αντιπροσωπεύει τα οικονομικά ζητήματα και περιλαμβάνει τιμές από 0 μέχρι 1, όπου με 0 υποδεικνύεται η αριστερή πολιτική τοποθέτηση και με 1 η δεξιά τοποθέτηση. Η ίδια διαδικασία ακολουθείται και για τον άξονα y , ο οποίος διαχωρίζει την προοδευτική τοποθέτηση από τη συντηρητική. Η βαρύτητα κάθε δήλωση πολιτικής σε κάθε ένα από τους δύο άξονες υπολογίζεται με παραγοντική ανάλυση και συμμετέχει με αντίστοιχη σημαντικότητα στη διαμόρφωση των τελικών συντεταγμένων (με βάση όλες τις δηλώσεις πολιτικής) για την τοποθέτηση του χρήστη και των κομμάτων στο συγκεκριμένο γράφημα διασποράς.

Συνδυασμός μεθόδων ταύτισης χρήστη - κομμάτων

Αρκετά VAAs επιλέγουν να χρησιμοποιήσουν και τις δύο μεθόδους που αναφέρθηκαν πιο πάνω για τον υπολογισμό της συνάφειας μεταξύ του χρήστη και των κομμάτων, όπως συμβαίνει με το *HelpMeVote*, το *Smartvote*, το *EU Profiler*, το *EuandI* και τα VAAs της ερευνητικής ομάδας του *PreferenceMatcher* [62]. Τα VAAs που αναφέρονται στις Ευρωπαϊκές εκλογές, επιπλέον των δύο διαστάσεων που αναπαριστούν την αριστερή-δεξιά πολιτική τοποθέτηση του χρήστη έναντι της προοδευτικής-συντηρητικής του αντίληψης, χρησιμοποιούν και μια τρίτη διάσταση που αφορά την Ευρωπαϊκή Ένωση και τα ζητήματα που δημιουργούνται γύρω από αυτήν. Επιπρόσθετα σε κάποια από αυτά, όπως το *Smartvote*, το *EU Profiler* και αρκετές από τις εφαρμογές του *PreferenceMatcher* γίνεται σύγκριση μεταξύ του χρήστη και των κομμάτων σε περισσότερες από τρεις διαστάσεις [62, 43].

Για παράδειγμα στο *Smartvote* κάθε χρήστης και κόμμα τοποθετούνται σε ένα “διάγραμμα αράχνης” (spider diagram) που περιλαμβάνει οκτώ άξονες με τιμές που κυμαίνονται μεταξύ του 0 και του 100. Στο Σχήμα 1.5 φαίνεται ένα παράδειγμα από τη γραφική απεικόνιση των αποτελεσμάτων του *Smartvote* με τη χρήση του “διαγράμματος αράχνης”. Πρόκειται για ένα κυκλικό πεδίο με οκτώ άξονες, που ανα-

¹⁶https://en.wikipedia.org/wiki/Factor_analysis



Σχήμα 1.5: Παράδειγμα της γραφικής απεικόνισης των αποτελεσμάτων από το *Smartvote* με τη χρήση του “διαγράμματος αράχνης” για τις κοινοτικές εκλογές της 24ης Σεπτεμβρίου 2017 στο Κόνιζ της Ελβετίας

φέρονται στα ακόλουθα θέματα: “ανοιχτή εξωτερική πολιτική” (open foreign policy), “ελεύθερη οικονομία” (liberal economy), “περιοριστική χρηματοοικονομική πολιτική” (restrictive financial policy), “νόμος και τάξη” (law & order), “περιοριστική πολιτική μετανάστευσης” (restrictive immigration policy), “επέκταση της προστασίας του περιβάλλοντος” (extended environmental protection), “ενισχυμένο κράτος πρόνοιας” (expanded welfare state) και “ελεύθερη κοινωνία” (liberal society).

Εκτός από τους παραδοσιακούς τρόπους σύστασης, κατά τους οποίους υπολογίζεται η συνάφεια μεταξύ του χρήστη και των κομμάτων, η ομάδα του *PreferenceMatcher* εφαρμόζει και τη λεγόμενη “κοινωνική σύσταση”, κατά την οποία υπολογίζεται η ομοιότητα μεταξύ του τρέχοντος χρήστη και των χρηστών που δήλωσαν ότι υποστηρίζουν τα κόμματα που περιλαμβάνονται στο VAA. Όταν ακολουθείται η συγκεκριμένη προσέγγιση, οι απαντήσεις των χρηστών στις δηλώσεις πολιτικής συγκρίνονται μεταξύ τους για να βρεθούν ποιοι χρήστες έχουν παρόμοιες τοποθετήσεις. Στη συνέχεια εμφανίζονται σε ένα χρήστη τα κόμματα ή υποψήφιοι που προτιθενται να ψηφίσουν οι παρόμοιοι χρήστες με αυτόν [62]. Ουσιαστικά ο τρόπος με τον οποίο λειτουργεί η κοινωνική προσέγγιση του VAA (Social VAA-SVAA) [57], στηρίζεται σε προσεγγίσεις του συνεργατικού φιλτραρίσματος και περιγράφεται ως εξής:

Οι χρήστες καλούνται να απαντήσουν σε μια σειρά από συμπληρωματικές ερωτήσεις, που είναι επιπρόσθετες των δηλώσεων πολιτικής του βασικού ερωτηματολογίου. Μία από τις συμπληρωματικές ερωτήσεις αφορά την πρόθεση ψήφου του χρήστη (δηλαδή ποιο κόμμα ο χρήστης προτίθεται να ψηφίσει στις επερχόμενες εκλογές). Ένα παράδειγμα του πώς εμφανίζονται οι συμπληρωματικές ερωτήσεις στο χρήστη φαίνεται στο Σχήμα 1.6. Η πρόθεση ψήφου για τις Ευρωπαϊκές Εκλογές 2014 δίνεται στην ερώτηση 2. Ακολούθως οι χρήστες που δήλωσαν την ίδια πρόθεση ψήφου ομαδοποιούνται μεταξύ τους και μοντελοποιείται ο τρόπος με τον οποίο τοποθετήθηκαν στο βασικό ερωτηματολόγιο, εφαρμόζοντας στατιστικές μεθόδους ή προσεγγίσεις από τη μηχανική μάθηση. Στη συνέχεια, τα μοντέλα αυτά μπορούν να αξιοποιηθούν για την παροχή σύστασης υιοθετώντας προσεγγίσεις του συνεργατικού φιλτραρίσματος [63, 57].

Η ομάδα του *PreferenceMatcher*, δίνει τη δυνατότητα στους χρήστες να αξιολογήσουν κατά πόσο

Μπορείτε να βοηθήσετε την έρευνα μας απαντώντας και στα παρακάτω σύντομα ερωτήματα. Εάν δεν επιθυμείτε να απαντήσετε, κάντε κλικ στο βέλος που αναγράφει «Συνέχεια»


Ποιό είναι το κόμμα στο οποίο νιώθετε πιο κοντά;

Στις Ευρωεκλογές, ποιό κόμμα προτιμάτε να ψηφίσετε;

Ποιός είναι ο κύριος λόγος για τον οποίο θα ψηφίσετε κάποιο κόμμα;

Στις τελευταίες εθνικές εκλογές του 2011, ποιό κόμμα ψηφίσατε;

Αν πραγματοποιούνταν αύριο εθνικές εκλογές, ποιό κόμμα θα ψηφίζατε;



Σχήμα 1.6: Η συμπληρωματική ερώτηση για την πρόθεση ψήφου όπως εμφανίζεται στο *EUvox 2014*

τα αποτελέσματα των πιο πάνω μεθόδων τους άφησαν ικανοποιημένους, χρησιμοποιώντας διαδικτυακά σύμβολα έκφρασης συναισθημάτων (emoticons). Ακόμη, μπορούν να μοιραστούν τα αποτελέσματα αυτά μέσω κοινωνικών δικτύων, όπως και να καλέσουν τους φίλους τους για να χρησιμοποιήσουν την εν λόγω εφαρμογή [57].

1.6 Η ποιότητα και ο “καθαρισμός” των δεδομένων των VAAs

Η ποιότητα των δεδομένων στα VAAs επηρεάζεται από δύο σημαντικά ζητήματα. Το πρώτο είναι η έλλειψη αντιπροσωπευτικότητας από το δείγμα λόγω μη στρωματοποιημένης κάλυψης του πληθυσμού. Για τη χρήση ενός VAA είναι απαραίτητη η πρόσβαση του χρήστη στο διαδίκτυο. Ένα μέρος του πληθυσμού όμως δεν έχει πρόσβαση στο διαδίκτυο ή επιλέγει να μην το χρησιμοποιεί ποτέ. Επιπλέον, όσοι το χρησιμοποιούν δεν επισκέπτονται όλοι την ιστοσελίδα των VAAs, αλλά είναι πιθανότερο να το πράξουν αυτοί που ενστερνίζονται το διαδίκτυο ως πηγή πληροφόρησης παρά εκείνοι που το θεωρούν ως μια πηγή διασκέδασης [67].

Το δεύτερο ζήτημα σχετικά με την ποιότητα των δεδομένων στα VAAs είναι το σφάλμα μέτρησης που παρατηρείται από εσκεμμένα λανθασμένες καταχωρήσεις των χρηστών. Ορισμένοι χρήστες προβαίνουν στη συμπλήρωση του ερωτηματολογίου του VAA περισσότερες από μία φορές. Την πρώτη φορά τείνουν να δίνουν τις αληθινές τους απαντήσεις, ενώ τις επόμενες φορές ανταποκρίνονται διαφορετικά για να παρατηρήσουν κατά πόσο αλλάζει η πρόταση ψήφου που δημιουργείται από το σύστημα με τη διαφοροποίηση των απαντήσεων τους. Υπάρχουν, όμως και οι χρήστες που επιθυμούν να διερευνήσουν το εργαλείο την πρώτη φορά που το χρησιμοποιούν δίνοντας τυχαίες απαντήσεις, ενώ στη συνέχεια ανταποκρίνονται με τις πραγματικές του θέσεις προς αυτό [67].

Για τον “καθαρισμό” των δεδομένων από ανεπιθύμητες καταχωρήσεις μπορούν να εφαρμοστούν πολλές τεχνικές [68]. Μια από αυτές είναι η διαγραφή όσων καταχωρήσεων προέρχονται από τον ίδιο υπολογιστή (IP Address), αλλά και αυτών που αφορούν χρήστες με διαφορετική γεωγραφική τοποθεσία από αυτήν της χώρας, στην οποία απευθύνεται το VAA. Οι δύο αυτές προσεγγίσεις, όμως, παρά την απλότητα στην εφαρμογή τους, αντιμετωπίζουν διάφορα προβλήματα. Για παράδειγμα ο ίδιος χρήστης μπορεί να χρησιμοποιήσει διαφορετικούς υπολογιστές για να αποκτήσει πρόσβαση στο VAA. Ή ακόμη οι καταχωρήσεις που προέρχονται από τον ίδιο υπολογιστή, μπορεί να μην αφορούν μόνο ένα άτομο αλλά

διαφορετικά άτομα που χρησιμοποιούν τον ίδιο υπολογιστή για να απαντήσουν στο ερωτηματολόγιο του VAA. Επιπλέον στις εκλογές για το Ευρωπαϊκό Κοινοβούλιο όπου οι πολίτες μπορούν να ψηφίσουν σε ένα κράτος μέλος διαφορετικό από εκείνο που αναφέρεται στο διαβατήριό τους, δεν μπορεί να εφαρμοστεί το μέτρο για τη διαφορετική γεωγραφική τοποθεσία.

Μια άλλη μέθοδος που ακολουθείται για τον καθαρισμό των δεδομένων του VAA είναι η εφαρμογή κατωφλιού στο χρόνο ανταπόκρισης των χρηστών, τόσο σε ολόκληρο το βασικό ερωτηματολόγιο όσο και στο χρόνο που αφιερώνει ο χρήστης για να απαντήσει σε κάθε ερώτηση ξεχωριστά. Υπό αυτές τις συνθήκες μπορούν να διαγράφονται από το δείγμα όσες καταχωρήσεις δε βρίσκονται μέσα στα προκαθορισμένα χρονικά όρια [67]. Επιπλέον οι καταχωρήσεις με μηδενικές τιμές ή και αυτές με ένα μεγάλο αριθμό διαδοχικών πανομοιότυπων απαντήσεων μπορούν να αγνοηθούν από το δείγμα.

Οι Djouvas et al. [69] πρότειναν τη χρήση τεχνικών εξόρυξης δεδομένων με μη επιβλεπόμενη μάθηση καθώς και ψυχομετρικές προσεγγίσεις, για τον εντοπισμό των χρηστών που παρουσιάζουν “παράδοξη συμπεριφορά” σχετικά με το ιδεολογικό προφίλ τους. Η εφαρμογή των συγκεκριμένων τεχνικών βασίζεται στη λογική πως οι χρήστες που απαντούν στο βασικό ερωτηματολόγιο παρουσιάζουν ένα βαθμό ιδεολογικής συνέπειας σε όλες τις απαντήσεις. Αυτό σημαίνει πως ακολουθούν ένα συγκεκριμένο τρόπο ανταπόκρισης, ανάλογα με τις ιδέες και τις απόψεις τους γύρω από τα ζητήματα πολιτικής του VAA. Έτσι, οποιοσδήποτε καταχωρήσεις παρουσιάζουν ακανόνιστη συμπεριφορά και αποκλίνουν πολύ από τον ιδεολογικά δομημένο τρόπο ανταπόκρισης θεωρούνται “παράδοξες” και αφαιρούνται από το δείγμα. Οι τεχνικές αυτές μπορούν να χρησιμοποιηθούν ως συμπληρωματικές των μεθόδων που αναφέρθηκαν στις προηγούμενες παραγράφους.

Για τον καθαρισμό των δεδομένων από την εφαρμογή Euvox, στην εργασία των Djouvas et al. [69], εξαιρέθηκαν όσες καταχωρήσεις χρηστών παρουσίασαν τα ακόλουθα χαρακτηριστικά:

1. Απάντησαν σε τουλάχιστον μια δήλωση πολιτικής του βασικού ερωτηματολογίου σε λιγότερο από δύο δευτερόλεπτα.
2. Ο συνολικός χρόνος που χρειάστηκε για να απαντήσουν και στις 30 δηλώσεις πολιτικής του βασικού ερωτηματολογίου δεν ξεπέρασε τα 120 δευτερόλεπτα.
3. Απάντησαν σε 20 ή περισσότερες δηλώσεις πολιτικής επιλέγοντας τη θέση “Δεν έχω άποψη”.
4. Απάντησαν σε 10 συνεχόμενες δηλώσεις πολιτικής επιλέγοντας την ίδια τοποθέτηση (για παράδειγμα απάντησαν σε 10 συνεχόμενες δηλώσεις πολιτικής επιλέγοντας τη θέση “Συμφωνώ”). Το συγκεκριμένο μέτρο υιοθετείται γιατί το ερωτηματολόγιο του Euvox είναι σχεδιασμένο έτσι ώστε διαδοχικές δηλώσεις πολιτικής να εμπεριέχουν δύο αντιτιθέμενες απόψεις. Έτσι ο χρήστης δεν μπορεί να έχει την ίδια άποψη σε δηλώσεις πολιτικής αντίθετης σημασίας.

1.7 Μεθοδολογία Έρευνας

Η παρούσα διατριβή καταπιάνεται με τη χρήση τεχνικών και μεθοδολογιών από τα Συστήματα Συστάσεων στην πανευρωπαϊκή εφαρμογή ηλεκτρονικού συμβούλου ψήφου Euvox στην οποία συμμετείχαν ενεργά και δήλωσαν πρόθεση ψήφου περισσότεροι από 1.2 εκατομμύρια Ευρωπαίοι πολίτες. Γίνεται διερεύνηση της συμπεριφοράς του Ευρωπαίου χρήστη στο περιβάλλον των VAAs, ενώ ταυτόχρονα αναλύονται τρόποι για τη βελτιστοποίηση της σύστασης ψήφου. Η ανάλυση είναι πρωτευόντως ποσοτική και επεξεργάζεται τα δεδομένα από τις απαντήσεις που έδωσαν οι χρήστες του Euvox, ως εκ τούτου

δεν υπήρξε πρωτογενής συλλογή δεδομένων. Για όλα τα πειράματα και την επεξεργασία των δεδομένων χρησιμοποιήθηκε το λογισμικό πακέτο MATLAB¹⁷.

Αρχικά τα δεδομένα επεξεργάστηκαν έτσι ώστε να αφαιρεθούν από το δείγμα όσοι χρήστες δε δήλωσαν την πρόθεση ψήφου τους. Στη συνέχεια χρησιμοποιήθηκαν διάφορες τεχνικές μηχανικής μάθησης για να προβλέψουν την πρόθεση ψήφου του κάθε χρήστη. Η τεχνική με την καλύτερη απόδοση, αυτή δηλαδή που κατάφερε να προβλέψει σωστά την πρόθεση ψήφου των περισσότερων χρηστών, υιοθετήθηκε σχεδόν σε όλα τα πειράματα που ακολούθησαν. Τα πειράματα αυτά διερευνήσαν μια σειρά από ερευνητικά ερωτήματα, τα όποια συνοψίζονται ως ακολούθως:

1. **Πόσο καλά μπορεί να προβλεφθεί η πρόθεση ψήφου ενός χρήστη του VAA, γνωρίζοντας μόνο τις απαντήσεις που έδωσε στις δηλώσεις πολιτικής; Σύγκριση σε πανευρωπαϊκό επίπεδο.** Εδώ μετρήθηκε η απόδοση των τεχνικών μηχανικής μάθησης που εφαρμόστηκαν σε κάθε χώρα ξεχωριστά. Η κάθε τεχνική προσπάθησε να υπολογίσει την πρόθεση ψήφου του χρήστη λαμβάνοντας υπόψη μόνο τις απαντήσεις του στις δηλώσεις πολιτικής. Τα αποτελέσματα διέφεραν από χώρα σε χώρα και από κόμμα σε κόμμα. Γενικά, όμως, τα αποτελέσματα που προέκυψαν ήταν ικανοποιητικά σε σχέση με άλλες μελέτες που διεξήχθησαν στο παρελθόν, καθώς επίσης για ακόμη μια φορά αποδείχτηκε πως ο κοινωνικός τρόπος σύστασης είναι αποτελεσματικότερος από τον παραδοσιακό. Επιπρόσθετα οι μη γραμμικές τεχνικές έδωσαν καλύτερα αποτελέσματα από τις άλλες τεχνικές, υποδηλώνοντας πως υπάρχει πολυπλοκότητα στον τρόπο που απάντησαν οι χρήστες στις δηλώσεις πολιτικής.
2. **Μπορεί να ενισχυθεί η απόδοση της πρόβλεψης ψήφου, όταν μαζί με τις δηλώσεις πολιτικής λαμβάνονται υπόψη και οι πιο κάτω συμπληρωματικές ερωτήσεις:**
 - τα κριτήρια με τα οποία ψηφίζει ο χρήστης,
 - την αυτοτοποθέτηση του στον πολιτικό χάρτη,
 - το ενδιαφέρον του για την πολιτική,
 - τα δημογραφικά του στοιχεία.

Σε αυτό το ερώτημα υιοθετήθηκε η τεχνική που έδωσε τα καλύτερα αποτελέσματα στο προηγούμενο ερώτημα, οι Μηχανές Διανυσμάτων Υποστήριξης (η περιγραφή αυτής της τεχνικής καθώς και όλων των τεχνικών που χρησιμοποιήθηκαν βρίσκονται στην υποενότητα 1.7.1). Τώρα όμως για την πρόβλεψη ψήφου, λάβαμε υπόψη μαζί με τις δηλώσεις πολιτικής και τις πιο πάνω συμπληρωματικές ερωτήσεις, στην προσπάθειά μας να εξετάσουμε εάν τα αποτελέσματα από το προηγούμενο ερώτημα μπορούν να βελτιωθούν. Επαναλάβαμε την πειραματική διαδικασία με τις Μηχανές Διανυσμάτων Υποστήριξης 24 φορές, προσπαθώντας να υπολογίσουμε την πρόθεση ψήφου του χρήστη κάθε φορά λαμβάνοντας υπόψη τις απαντήσεις του στις δηλώσεις πολιτικής μαζί με διαφορετικό συνδυασμό συμπληρωματικών ερωτήσεων. Για παράδειγμα την πρώτη φορά λάβαμε υπόψη τις δηλώσεις πολιτικής μαζί με τα κριτήρια ψήφου του χρήστη, τη δεύτερη φορά τις δηλώσεις πολιτικής μαζί με τα κριτήρια ψήφου και το φύλο του χρήστη, την τρίτη φορά τις δηλώσεις πολιτικής μαζί με τα κριτήρια ψήφου, το φύλο και την ηλικία του χρήστη και η διαδικασία συνεχίστηκε με αυτό το ρυθμό μέχρι να συνδυαστούν με 24 διαφορετικούς τρόπους οι δηλώσεις πολιτικής μαζί με τις συμπληρωματικές ερωτήσεις.

Στη συνέχεια παρατηρήθηκε ο συνδυασμός που απέδωσε καλύτερα σε κάθε χώρα ξεχωριστά, ο οποίος ήταν διαφορετικός από χώρα σε χώρα. Ωστόσο σε όλες τις χώρες οι συμπληρωματικές ερωτήσεις βελτίωσαν την απόδοση της πρόβλεψης ψήφου κατά μέσο όρο πέντε μονάδες. Οι ερωτήσεις

¹⁷<https://www.mathworks.com/products/matlab.html>

που αφορούσαν το ενδιαφέρον του χρήστη για την πολιτική και τα κριτήρια ψήφου του, ενίσχυσαν την απόδοση του VAA κατά μέσο όρο έξι μονάδες.

Να αναφέρουμε πως για το ερώτημα αυτό, αφαιρέθηκαν από το δείγμα όσοι χρήστες δεν είχαν απαντήσει σε όλες τις συμπληρωματικές ερωτήσεις. Έτσι θεωρήθηκε αναγκαίο να βρεθούν τρόποι για την εκτίμηση των απαντήσεων στις συμπληρωματικές ερωτήσεις που δεν είχαν απαντηθεί και να επαναληφθεί η πειραματική διαδικασία σε ολόκληρο το δείγμα. Ο τρόπος που ακολουθήθηκε για την εκτίμηση των ελλειπουσών τιμών καθώς και τα αποτελέσματα που προέκυψαν απαντώνται στο επόμενο ερευνητικό ερώτημα.

3. Μπορεί να γίνει εκτίμηση των απαντήσεων στις συμπληρωματικές ερωτήσεις του VAA, χρησιμοποιώντας τις δηλώσεις πολιτικής; Σε αυτό το ερώτημα εκτιμήσαμε τις συμπληρωματικές ερωτήσεις που δεν είχαν απαντηθεί από τον κάθε χρήστη, χρησιμοποιώντας τις απαντήσεις του στις δηλώσεις πολιτικής. Στη συνέχεια προσθέσαμε τις εκτιμημένες τιμές στο δείγμα και με τη βοήθεια των Μηχανών Διανυσμάτων Υποστήριξης, επαναλάβαμε την πειραματική διαδικασία του προηγούμενου ερωτήματος. Τα αποτελέσματα ήταν παρόμοια με αυτά που προέκυψαν στο προηγούμενο ερευνητικό ερώτημα, όπου το δείγμα ήταν μειωμένο. Έτσι καταλήξαμε στο συμπέρασμα πως μπορούν να χρησιμοποιηθούν αποτελεσματικά τεχνικές παραγοντοποίησης πινάκων (matrix factorization) και συνεργατικού φιλτραρίσματος για την εκτίμηση των ελλειπουσών τιμών. Βέβαια οι χρήστες του VAA που δεν απάντησαν σε όλες τις συμπληρωματικές ερωτήσεις ήταν ελάχιστοι, έτσι ακόμα και στην περίπτωση που επιλέχθηκε να αφαιρεθούν από το δείγμα αντί να εκτιμηθούν οι απαντήσεις τους, το VAA είχε παρόμοια συμπεριφορά.

4. Ιδεολογική ταύτιση κόμματος και χρήστη. Οι τεχνικές μηχανικής μάθησης που εφαρμόστηκαν στο πρώτο ερευνητικό ερώτημα, είναι αποδεδειγμένο από προηγούμενες έρευνες [96, 81, 24] πως επηρεάζονται από την πολυμορφία εντός των κομμάτων. Ως εκ τούτου τα αποτελέσματα που παρουσίασε η κάθε τεχνική μάθησης μπορούν να δικαιολογηθούν σε μεγάλο βαθμό εάν γνωρίζουμε κατά πόσο υπάρχει ανομοιογένεια μεταξύ των χρηστών που δηλώνουν πρόθεση ψήφου για το ίδιο κόμμα. Έτσι σε αυτό το ερευνητικό ερώτημα διερευνήθηκε η ιδεολογική ταύτιση μεταξύ των κομμάτων και των χρηστών που τα υποστηρίζουν και εξετάστηκε η πολυμορφία στα κόμματα της κάθε χώρας. Για να συμβεί αυτό αρχικά χωρίστηκαν οι χρήστες της κάθε χώρας σε ομάδες, ανάλογα με την ομοιότητα τους στις δηλώσεις πολιτικής. Στη συνέχεια αναλύθηκε το περιεχόμενο της κάθε καινούργιας ομάδας για να απαντηθούν ερωτήματα όπως αυτά που ακολουθούν:

- Από ποια κόμματα προέρχονται οι χρήστες που βρίσκονται στην ίδια ομάδα;
- Υπάρχουν χρήστες στην ίδια ομάδα που προέρχονται από το ίδιο πολιτικό κόμμα;
- Σε πόσες ομάδες ομοίων χρηστών χωρίζονται οι χρήστες του κάθε κόμματος;

Τα αποτελέσματα έδειξαν πως πολλοί χρήστες έχουν παρόμοια ιδεολογία με χρήστες που υποστηρίζουν κάποιο διαφορετικό κόμμα. Αυτό μπορεί να φανερώνει ότι οι χρήστες δε γνωρίζουν πραγματικά τις πολιτικές θέσεις των κομμάτων που υποστηρίζουν, οι πολιτικές θέσεις των κομμάτων δεν είναι ξεκάθαρες προς τους χρήστες ή/και οι χρήστες δε δηλώνουν ειλικρινά την πρόθεση ψήφου τους μπερδεύοντας το σύστημα κατά την κατηγοριοποίηση τους. Επιπρόσθετα παρατηρήθηκε πολυμορφία εντός των κομμάτων της κάθε χώρας, με τα περισσότερα ομοιογενή κόμματα να έχουν καλύτερες προβλέψεις ψήφου.

Ως απάντηση στο εξεταζόμενο ερευνητικό ερώτημα δόθηκε πως υπάρχει ιδεολογική ταύτιση μεταξύ των χρηστών και των κομμάτων, αλλά σε διαφορετικό βαθμό από χώρα σε χώρα και από κόμμα σε κόμμα.

5. **Μπορούν να εξαχθούν συμπεράσματα για το ενδιαφέρον του Ευρωπαίου χρήστη του VAA για την πολιτική, ανάλογα με τα κριτήρια ψήφου και τα δημογραφικά του στοιχεία;** Για το συγκεκριμένο ερώτημα εφαρμόστηκε η μέθοδος της γραμμικής παλινδρόμησης για να εξεταστεί η σχέση μεταξύ του ενδιαφέροντος για την πολιτική και των κριτηρίων ψήφου, καθώς και η σχέση μεταξύ του ενδιαφέροντος για την πολιτική και των δημογραφικών στοιχείων του χρήστη. Τα αποτελέσματα από τα πειράματα που διεξήχθησαν έδειξαν πως οι άντρες χρήστες με υψηλή μόρφωση τείνουν να ενδιαφέρονται περισσότερο για την πολιτική, καθώς επίσης το ενδιαφέρον των αντρών για την πολιτική είναι μεγαλύτερο από αυτό των γυναικών. Η ηλικία φαίνεται να μην επηρεάζει σε μεγάλο βαθμό το ενδιαφέρον του χρήστη, ενώ σε όλες σχεδόν τις χώρες είναι τουλάχιστον 99% σίγουρο πως οι χρήστες που προτιμούν ένα κόμμα εξαιτίας της οικογένειας ή των φίλων τους ενδιαφέρονται λιγότερο για την πολιτική, από αυτούς που δήλωσαν πως ψηφίζουν ένα κόμμα εξαιτίας της ιδεολογίας του. Στις περισσότερες περιπτώσεις οι χρήστες που επιλέγουν να ψηφίσουν ένα κόμμα γιατί είναι ευχαριστημένοι με την ηγεσία του κόμματος ή επειδή το κόμμα αυτό είναι ικανότερο από τα υπόλοιπα ενδιαφέρονται από αρκετά έως πολύ για την πολιτική.
6. **Μπορούν οι απαντήσεις στο VAA ερωτηματολόγιο να χωριστούν σε κατηγορίες - κριτήρια, έτσι ώστε για την πρόβλεψη ψήφου να χρησιμοποιούνται μόνο οι συνολικές βαθμολογίες που προκύπτουν από το άθροισμα των ερωτήσεων σε κάθε κατηγορία - κριτήριο;** Το συγκεκριμένο ερώτημα ασχολήθηκε με το εάν μπορεί να θεωρηθεί το πρόβλημα ενός VAA ως ένα πρόβλημα Συστημάτων Συστάσεων πολλαπλών κριτηρίων, λαμβάνοντας υπόψη τις τρεις θεματικές ενότητες, στις οποίες ανήκουν οι δηλώσεις πολιτικής (Ευρωπαϊκή Ένωση, Οικονομία, Πολιτική -Κουλτούρα). Οι απαντήσεις των χρηστών στις δηλώσεις πολιτικής αθροίστηκαν κάτω από την ενότητα που ανήκαν και οι Μηχανές Διαनुσμάτων Υποστήριξης λάμβαναν υπόψη μόνο τις τρεις συνολικές βαθμολογίες για την πρόβλεψη της πρόθεσης ψήφου. Μέσα από τα αποτελέσματα καταλήξαμε στο συμπέρασμα πως η ομαδοποίηση των ερωτήσεων φαίνεται να μην λειτουργεί στην περίπτωση του VAA.

Να σημειώσουμε πως τα πιο πάνω αναφέρονται στη γενική εικόνα πίσω από τα αποτελέσματα που προέκυψαν μέσα από τη μελέτη των ερευνητικών ερωτημάτων. Η ανάλυση των αποτελεσμάτων για κάθε χώρα γίνεται στα επόμενα κεφάλαια.

1.7.1 Μαθηματικό Υπόβαθρο

Η παραδοσιακή προσέγγιση σύστασης ενός VAA μπορεί να περιγραφεί ως εξής: Ο βασικός στόχος του VAA είναι να συστήνει έναν υποψήφιο/κόμμα για κάθε χρήστη. Σε μια τέτοια περίπτωση, υπάρχει ένα σύνολο από N χρήστες $\mathcal{X} = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N\}$, ένα σύνολο U ερωτήσεων $\mathcal{Q} = \{q_1, q_2, \dots, q_U\}$ και ένα σύνολο D πολιτικών κομμάτων (ή υποψηφίων) $\mathcal{P} = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_D\}$. Κάθε χρήστης $\vec{x}_j \in \mathcal{X}$ και κάθε πολιτικό κόμμα $\vec{p}_i \in \mathcal{P}$, έχει απαντήσει σε κάθε ερώτηση $q_k \in \mathcal{Q}$. Με βάση τις απαντήσεις τους, κάθε πολιτικό κόμμα και κάθε χρήστης αναπαρίστανται σε ένα μοντέλο διανυσματικού χώρου (Vector Space Model)¹⁸:

$$\vec{x}_j = \{x(j, 1), x(j, 2), \dots, x(j, k), \dots, x(j, U)\} \quad (1.1)$$

$$\vec{p}_i = \{p(i, 1), p(i, 2), \dots, p(i, k), \dots, p(i, U)\} \quad (1.2)$$

¹⁸https://en.wikipedia.org/wiki/Vector_space_model

όπου $x(j, k), p(i, k) \in \mathcal{L}$ είναι οι απαντήσεις του j -οστού χρήστη και i -οστού κόμματος, αντίστοιχα, στο k -οστό ερώτημα. Ένα τυπικό σύνολο απαντήσεων στην κλίμακα Likert διατυπώνεται ως ακολούθως: $\mathcal{L} = \{1$ (Συμφωνώ πλήρως), 2 (Συμφωνώ), 3 (Ούτε συμφωνώ ούτε διαφωνώ), 4 (Διαφωνώ), 5 (Διαφωνώ πλήρως), 6 (Δεν έχω άποψη)}.

Σε πολλές περιπτώσεις και στην πλειονότητα των μεθόδων που έχουν προταθεί μέχρι στιγμής για τα SVAA, το έκτο σημείο δε λαμβάνεται υπόψη, δεδομένου ότι δεν αντιστοιχεί σε συγκεκριμένη στάση και συνήθως αντικαθίσταται με το τρίτο σημείο, δηλαδή την επιλογή “ούτε συμφωνώ ούτε διαφωνώ”. Στην παρούσα εργασία δεν έγινε η αντικατάσταση του έκτου σημείου με το τρίτο καθώς οι πολιτικοί επιστήμονες υποστηρίζουν έντονα τη διαφορά μεταξύ των δύο αυτών κατηγοριών [50]. Ως εκ τούτου το σύνολο \mathcal{L} να παραμένει ως έχει: $\mathcal{L} = \{1, 2, 3, 4, 5, 6\}$.

Το πρόβλημα σύστασης που αντιμετωπίζει το VAA προσπαθεί να προσεγγίσει αρχικά την άγνωστη συνάφεια $h(j, i)$ του χρήστη j για το κόμμα i , δεδομένων των απαντήσεων του \vec{x}_j στις δηλώσεις πολιτικής του VAA, και στη συνέχεια να προτείνει την ταξινόμηση των πολιτικών κομμάτων με βάση τη συνάφεια προς το χρήστη εκάστου κόμματος. Στην ορολογία της μηχανικής μάθησης, ο στόχος είναι η προσέγγιση της κρυφής συνάρτησης $h(j, i)$ με μια συνάρτηση $\hat{h}(\vec{x}_j, \vec{p}_i): \mathbb{R}^U \times \mathbb{R}^U \rightarrow \mathbb{R}$, όπου $\hat{h}(\vec{x}_j, \vec{p}_i)$ είναι η εκτίμηση της συνάφειας του χρήστη j με το κόμμα i . Χωρίς περιορισμό της γενίκευσης θεωρούμε ότι ισχύει $\hat{h}(\vec{x}, \vec{p}) \in [0, 1]$. Σε κάθε περίπτωση, η κορυφαία πρόταση (σύσταση) p_q^j για το χρήστη j θα πρέπει να είναι:

$$p_q^j = \underset{i}{\operatorname{argmax}} \left\{ \hat{h}(\vec{x}_j, \vec{p}_i) \right\} \quad (1.3)$$

Για να παραχθεί η σύσταση ψήφου, η πιο απλή προσέγγιση είναι να εκτιμηθεί η $\hat{h}(\vec{x}, \vec{p})$ με τη βοήθεια ενός μέτρου απόστασης $d(\vec{x}, \vec{p})$. Μέτρα που υπολογίζουν την ομοιότητα / απόσταση αναφέρονται στις επόμενες ενότητες. Στην περίπτωση αυτή, η κορυφαία πρόταση p_q^j για το χρήστη j δίνεται από τη σχέση:

$$p_q^j = \underset{i}{\operatorname{argmin}} \left\{ d(\vec{x}_j, \vec{p}_i) \right\} \quad (1.4)$$

Όπως ήδη αναφέρθηκε, το κοινωνικό VAA (social VAA - SVAA) αντί να συγκρίνει τις απαντήσεις των χρηστών με αυτές των κομμάτων, ομαδοποιεί του χρήστες που δήλωσαν ως πρόθεση ψήφου v_j το ίδιο κόμμα και βάσει των απαντήσεων τους, δημιουργεί το μοντέλο των ψηφοφόρων του κάθε κόμματος. Στη συνέχεια υπολογίζεται ο βαθμός ταύτισης των απαντήσεων ενός καινούργιου χρήστη με κάθε ένα από τα μοντέλα αυτά και προτείνεται στον χρήστη το κόμμα το μοντέλο του οποίου δίνει την μεγαλύτερη ταύτιση με τις απαντήσεις του χρήστη.

Προκειμένου να λυθεί το πρόβλημα του SVAA, από πλευράς μηχανικής μάθησης, μπορούν να χρησιμοποιηθούν διάφοροι ταξινομητές μέσω συναρτήσεων βαθμολογίας $f(\vec{a}, b)$, οι οποίες δείχνουν την πιθανότητα το διάνυσμα \vec{a} (στην περίπτωση μας οι απαντήσεις ενός χρήστη στις δηλώσεις πολιτικής) να ανήκει στην κατηγορία b (κάποιο από τα κόμματα). Ως εκ τούτου μπορεί να τεθεί η $\hat{h}(\vec{x}, \vec{p}) \equiv f(\vec{a}, b)$ και να χρησιμοποιηθεί ως συνάρτηση f κάποιος ταξινομητής, όπως για παράδειγμα ο ταξινομητής Bayes, οι Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ) (support vector machines), τα Νευρωνικά Δίκτυα, κ.κ. Στην ουσία αυτό που επιτυγχάνεται με τη μηχανική μάθηση είναι η μοντελοποίηση των κομμάτων βάσει των πολιτικών προτιμήσεων των χρηστών που δήλωσαν πως προτίθενται να τα ψηφίσουν στις επερχόμενες εκλογές, έτσι ώστε όταν ένας χρήστης ταξινομείται σε κάποιο κόμμα, αυτό να σημαίνει πώς μοιράζεται τις ίδιες πολιτικές θέσεις με τα άλλα άτομα που το υποστηρίζουν.

Στο πλαίσιο της παρούσας διατριβής οι ταξινομητές που χρησιμοποιήθηκαν είναι τα Κρυμμένα Μο-

ντέλα Markov, ο ταξινομητής Bayes, οι Μηχανές Διανυσμάτων Υποστήριξης, τα Τεχνητά Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης, ο ταξινομητής Ελάχιστης Απόστασης Mahalanobis και ο αλγόριθμος των k -Πλησιέστερων Γειτόνων.

Κρυμμένα Μοντέλα Μαρκόβ

Τα Κρυμμένα Μοντέλα Markov (Hidden Markov Models -HMMs) είναι μια διπλή στοχαστική διαδικασία που χρησιμοποιείται για τη μοντελοποίηση διεργασιών οι οποίες εξελίσσονται στο χρόνο. Είναι δηλαδή μια στοχαστική διαδικασία που καθορίζεται από δύο αλληλένδετους μηχανισμούς, (α) μία λανθάνουσα αλυσίδα Markov διακριτού χρόνου που αποτελείται από ένα πεπερασμένο αριθμό καταστάσεων και (β) μια σειρά από παρατηρήσεις για κάθε κατάσταση (state) της εσωτερικής αλυσίδας Markov. Κάθε χρονική στιγμή το σύστημα μεταβαίνει από μια κατάσταση σε κάποια άλλη, η οποία μπορεί να είναι ίδια με την αρχική (προηγούμενη). Μετά από κάθε μετάβαση παράγεται ένα σύμβολο εξόδου, βασισμένο σε μια κατανομή πιθανότητας που αντιστοιχεί στην τρέχουσα κατάσταση [70, 71]. Κάθε μία από τις μεταβάσεις συμβαίνει με μια συσχετισμένη πιθανότητα. Η επόμενη κατάσταση εξαρτάται αποκλειστικά και μόνο από την τρέχουσα κατάσταση και τις πιθανότητες μετάβασης [27]. Στα HMMs, οι καταστάσεις δεν είναι παρατηρήσιμες- είναι 'κρυφές', αλλά όταν το μοντέλο επισκέπτεται μια κατάσταση τότε μια παρατήρηση καταγράφεται ως η πιθανοτική συνάρτηση της κατάστασης αυτής [26].

Ένα HMM χαρακτηρίζεται από τα ακόλουθα στοιχεία [26]:

- Ένα σύνολο W διακριτών καταστάσεων $\mathcal{S} = \{S_1, S_2, S_3, \dots, S_W\}$ με $\mathcal{G} = \{g_1 g_2 \dots g_T\}$ μια ακολουθία καταστάσεων τέτοια ώστε $g_k \in \mathcal{S} \forall k$.
- Ένα σύνολο E συμβόλων (παρατηρήσεων) $\mathcal{V} = \{v_1, v_2, v_3, \dots, v_E\}$ με $\mathcal{O} = \{O_1 O_2 \dots O_T\}$ μια ακολουθία παρατηρήσεων, $O_k \in \mathcal{V} \forall k$, που αντιστοιχούν στις καταστάσεις \mathcal{G} .
- Έναν πίνακα πιθανοτήτων μετάβασης μεταξύ των καταστάσεων $A: a_{ij} \equiv Pr(g_{t+1} = S_j | g_t = S_i)$.
- Έναν πίνακα κατανομής πιθανοτήτων παρατήρησης $B: b_j(e) \equiv Pr(O_t = v_e | g_t = S_j)$.
- Την κατανομή πιθανότητας των αρχικών καταστάσεων $\pi: \pi_i \equiv Pr[g_i = S_i]$.

Ο αλγόριθμος Baum-Welch (BW) : Ένα HMM περιγράφεται πλήρως από τρεις παραμέτρους: $\lambda = (A, B, \pi)$. Οι τεχνικές που εφαρμόζονται για την εκτίμηση των παραμέτρων του μοντέλου γίνονται συνήθως με μάθηση παρτίδας (batch learning) είτε με βάση τεχνικές Expectation Maximization (EM) [72], όπως αλγόριθμος Baum-Welch [73], είτε με αριθμητικές τεχνικές βελτιστοποίησης, όπως ο αλγόριθμος Gradient Descent¹⁹.

Ο αλγόριθμος Baum-Welch (BW) είναι μια ειδική περίπτωση του αλγόριθμου Expectation Maximization (EM)²⁰ και χρησιμοποιείται, όπως ήδη αναφέρθηκε, για την εκτίμηση των παραμέτρων ενός HMM. Ο αλγόριθμος EM είναι μια γενική επαναληπτική μέθοδος, η οποία υπολογίζει το κριτήριο της μέγιστης πιθανοφάνειας (MLE) των παραμέτρων μιας υποκείμενης κατανομής σε ένα σύνολο δεδομένων, όταν τα δεδομένα είναι ελλιπή. Ο EM λειτουργεί επαναληπτικά σε δύο βήματα.

1. Το πρώτο, που ονομάζεται Expectation, (E-βήμα), υπολογίζει την κατανομή των κρυμμένων μεταβλητών σύμφωνα με τις παρατηρήσεις και τις καταστάσεις.

¹⁹https://en.wikipedia.org/wiki/Gradient_descent

²⁰https://en.wikipedia.org/wiki/Expectation-maximization_algorithm

2. Το δεύτερο βήμα είναι το Maximization (M- βήμα) και περιλαμβάνει τη μεγιστοποίηση της αναμενόμενης πιθανότητας που βρέθηκε στο E-βήμα.

Οι παράμετροι που υπολογίζονται στο M-βήμα δημιουργούν ένα αρχικό μοντέλο λ_0 και στη συνέχεια χρησιμοποιούνται ως αρχικές συνθήκες στην επόμενη επανάληψη που θα εκτιμήσει τις παραμέτρους του μοντέλου λ_1 , κάτι που θα επαναλαμβάνεται μέχρι να γίνει σύγκλιση σε ένα σταθερό σημείο προεπιλεγμένης πιθανότητας [26, 73, 72, 70]. Ο αλγόριθμος Baum-Welch στο E-βήμα εκτιμά τις πιθανότητες ξ και γ και στο M-βήμα υπολογίζει επαναληπτικά τις πιθανότητες της αρχικής κατάστασης $\hat{\pi}_i$, τις πιθανότητες μετάβασης \hat{a}_{ij} και τις πιθανότητες δημιουργίας συμβόλων από κάθε μια από τις καταστάσεις $\hat{b}_j(k)$, όπου:

- $\xi_t(i, j) \equiv Pr(g_t = S_i, g_{t+1} = S_j | O, \lambda)$ είναι η πιθανότητα να παρατηρηθεί η κατάσταση S_i τη στιγμή t και η κατάσταση S_j τη χρονική στιγμή $t + 1$, δεδομένων των λ και O
- $\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j)$ είναι η πιθανότητα να παρατηρηθεί η κατάσταση S_i τη στιγμή t
- $\hat{\pi}_i = \gamma_1(i)$ είναι η συχνότητα της κατάστασης S_i τη χρονική στιγμή $t=1$
- $\hat{a}_{ij} = \frac{\sum_t \xi_t(i, j)}{\sum_t \gamma_t(i)}$ είναι η αναλογία μεταξύ των μεταβάσεων από την κατάσταση S_i στην S_j , προς τον συνολικό αριθμό των μεταβάσεων από την κατάσταση S_i
- $\hat{b}_j(k) = \frac{\sum_{t, O_t=k} \gamma_t(j)}{\sum_t \gamma_t(j)}$ είναι το πηλίκο του πλήθους των παρατηρήσεων του συμβόλου v_k στην κατάσταση S_j , προς το πλήθος των εμφανίσεων της κατάστασης S_j .

Στην παρούσα διατριβή χρησιμοποιήθηκε το HMM Toolbox της Matlab, το οποίο κατασκευάστηκε από τον Kevin Murphy το 1998, χρησιμοποιώντας τον αλγόριθμο Baum-Welch (BW) για την εκτίμηση των παραμέτρων του HMM με διακριτές εξόδους. Θεωρήσαμε πως υπάρχουν τρεις καταστάσεις των HMMs, δηλαδή $W = 3$ και $S = \{S_1, S_2, S_3\}$. Με την κατάσταση S_1 εκφράζεται η θετική στάση του χρήστη, με S_2 η ουδέτερη στάση και με S_3 η αρνητική στάση. Κάθε κατάσταση αντιστοιχεί στις επιλογές των απαντήσεων $S_1 = \{\text{‘Συμφωνώ πλήρως’}, \text{‘Συμφωνώ’}\}$, $S_2 = \{\text{‘Ούτε συμφωνώ ούτε διαφωνώ’}, \text{‘Δεν έχω άποψη’}\}$ και $S_3 = \{\text{‘Διαφωνώ’}, \text{‘Διαφωνώ πλήρως’}\}$ που θα μπορούσε να δώσει ένας χρήστης στις δηλώσεις πολιτικής του VAA ερωτηματολογίου. Κάθε κατάσταση ακολουθίας \mathcal{G} έχει μήκος ίσο με τον αριθμό των δηλώσεων πολιτικής, δηλ. $T = U = 30$ ενώ η χαρτογράφηση (προβολή) από το προφίλ ενός χρήστη \vec{x}_j (βλ. επίσης εξίσωση 1.1) σε μια ακολουθία συμβόλων $\mathcal{V}^j = \{v_1^j, v_2^j, v_3^j, \dots, v_T^j\}$ εκφράζεται ως ακολούθως:

$$v_q^j = x(j, q) + |\mathcal{L}| \cdot (q - 1) \quad (1.5)$$

όπου $x(j, q)$ είναι η επιλογή απάντησης του χρήστη j στην δήλωση πολιτικής q ($q = 1, 2, \dots, U$), \mathcal{L} είναι το σύνολο των επιλογών απάντησης που έχει ο χρήστης και $|\mathcal{L}|$ ο αντίστοιχος πληθάρθμος (σε αυτήν την περίπτωση $|\mathcal{L}| = 6$).

Η πιθανοτική μέθοδος Naïve Bayes

Η ταξινόμηση κατά το Naïve Bayes βασίζεται στο θεώρημα του Bayes με την υπόθεση πως η πιθανότητα ενός χαρακτηριστικού να ανήκει σε μια συγκεκριμένη κατηγορία (και στην περίπτωση της παρούσας εργασίας η πιθανότητα ενός χρήστη να ψηφίζει (ανήκει σε) ένα συγκεκριμένο κόμμα) δε σχετίζεται με την ύπαρξη των άλλων χαρακτηριστικών σε αυτή την κατηγορία. Αυτό σπάνια συμβαίνει στην πραγματικότητα, εξού και το όνομα Naïve (naïve = αφελής).

Το θεώρημα Bayes, προσαρμοσμένο στο πρόβλημα των VAAs, συνοψίζεται στη εξίσωση 1.6:

$$Pr(p_i|\vec{x}_j) = \frac{Pr(\vec{x}_j|p_i) \cdot Pr(p_i)}{Pr(\vec{x}_j)} \quad (1.6)$$

όπου:

- $Pr(p_i|\vec{x}_j)$ είναι η πιθανότητα να επιλεγεί το κόμμα p_i με δεδομένο το προφίλ του χρήστη \vec{x}_j και ονομάζεται *εκ των υστέρων πιθανότητα (posterior probability) του κόμματος p_i*
- $Pr(\vec{x}_j|p_i)$ είναι η πιθανότητα να παρατηρηθεί το προφίλ \vec{x}_j δεδομένου ότι ο χρήστης ψηφίζει (ανήκει στο) κόμμα p_i και αποκαλείται *πιθανοφάνεια (likelihood) του προφίλ χρήστη \vec{x}_j δεδομένου του κόμματος p_i*
- $Pr(p_i)$ είναι η πιθανότητα να επιλεγεί το κόμμα p_i και λέγεται *εκ των προτέρων πιθανότητα (prior probability) του κόμματος p_i*
- $Pr(\vec{x}_j)$ είναι η πιθανότητα της δημιουργία του προφίλ χρήστη \vec{x}_j , ανεξαρτήτως του κόμματος που ο χρήστης υποστηρίζει και ονομάζεται *εκ των προτέρων πιθανότητα του προφίλ χρήστη \vec{x}_j* .

Ο ταξινομητής Naive Bayes ταξινομεί τον χρήστη j στο κόμμα p_i , του οποίου η τιμή μεγιστοποιεί την πιθανότητα $Pr(p_i|\vec{x}_j)$.

Ο αλγόριθμος των k -κοντινότερων γειτόνων

Η μέθοδος των k πλησιέστερων γειτόνων (k -Nearest Neighbors -KNN) ταξινομεί τους χρήστες με βάση την ομοιότητα μεταξύ των απαντήσεών τους. Είναι ένας από τους απλούστερους αλγόριθμους μηχανικής μάθησης. Μετρά την ομοιότητα μεταξύ των απαντήσεων των χρηστών (προφίλ χρηστών) χρησιμοποιώντας κάποιο μέτρο απόστασης (η συγκεκριμένη εργασία βασίζεται στην Ευκλείδεια απόσταση) και βρίσκει τους k κοντινότερους γείτονες προς τον υπό “διερεύνηση” χρήστη. Ο χρήστης αυτός θα ταξινομηθεί στην κατηγορία που πλειοψηφεί ανάμεσα τους k κοντινότερους γείτονες. Η τιμή του k περιορίζει τον αριθμό των “όμοιων” χρηστών και αντιστοιχεί σε κάποιο θετικό ακέραιο αριθμό μικρού μέτρου [74]. Το πιο συνηθισμένο πρόβλημα με τη μέθοδο των k πλησιέστερων γειτόνων είναι ότι το κατάλληλο k δεν είναι πάντα εύκολο να προσδιοριστεί.

Ο KNN ταξινομητής στο VAA εφαρμόζεται ως ακολούθως:

1. Επιλέγεται το προφίλ ενός χρήστη \vec{x}_j
2. Εντοπίζονται (με τη βοήθεια κάποιας μετρικής απόστασης, συνήθως της Ευκλείδειας) οι k χρήστες-γειτόνες που το προφίλ τους (δηλαδή οι απαντήσεις τους στις δηλώσεις πολιτικής) βρίσκεται πιο κοντά στο προφίλ του χρήστη \vec{x}_j
3. Ο χρήστης j ταξινομείται στην ομάδα - κόμμα που έχει τη μεγαλύτερη πιθανότητα να βρίσκεται μεταξύ των ομάδων - κομμάτων που ανήκουν οι κοντινότεροι του γείτονες

Ο ταξινομητής Mahalanobis

Η ταξινόμηση των χρηστών σε κάποιο κόμμα μπορεί να γίνει και με τη χρήση της απόστασης Mahalanobis [75]. Ο ταξινομητής Mahalanobis υποθέτει πως ο χρήστης έχει μεγάλη πιθανότητα να ανήκει στην κατηγορία

από την οποία έχει τη μικρότερη απόσταση. Για να συμβεί αυτό πρέπει οι κατηγορίες στο n -διάστατο χώρο να είναι διαχωρίσιμες, έτσι ώστε τα όρια τους να είναι ευδιάκριτα και η κατηγοριοποίηση να γίνεται με μεγαλύτερη επιτυχία. Η άγνωστη συνάρτηση $h(j, i)$ του χρήστη j για το κόμμα i με τη χρήση της Mahalanobis απόστασης μπορεί να υπολογιστεί ως ακολούθως:

$$L(j, i | \vec{x}_j) = e^{-\frac{1}{2}(x_j - \vec{m}_i)^T \cdot \mathbf{C}_i^{-1} \cdot (x_j - \vec{m}_i)} \quad (1.7)$$

όπου \vec{m}_i είναι ένα διάνυσμα ($U \times 1$ στοιχείων, όπου U είναι ο αριθμός των ερωτήσεων στο ερωτηματολόγιο του VAA) που αποτελείται από τις μέσες απαντήσεις των χρηστών που εξέφρασαν την πρόθεση ψήφου για το κόμμα i , και \mathbf{C}_i^{-1} είναι ο αντίστροφος του πίνακα συνδιασποράς ($U \times U$ στοιχείων) των ερωτηματολογίων που απάντησαν οι χρήστες που εξέφρασαν την πρόθεση ψήφου για το κόμμα i . Σε ένα VAA δεν εκφράζουν όλοι οι χρήστες την πρόθεση ψήφου τους. Αυτό έχει ως αποτέλεσμα, σε κάποιες περιπτώσεις, τα ερωτηματολόγια των χρηστών που εκφράζουν την πρόθεση ψήφου για ένα συγκεκριμένο κόμμα i να είναι αρκετά. Υπό αυτές τις συνθήκες ο \mathbf{C}_i αρχικοποιείται ως μοναδιαίος πίνακας (identity matrix) και ενημερώνεται διαδοχικά σύμφωνα με τις ακόλουθες εξισώσεις:

$$\mathbf{C}_i^0 = \mathbf{C}_i^1 = \mathbf{I}, \quad \vec{m}_i^0 = \vec{0}, \quad \vec{m}_i^1 = \vec{x}_1 \quad (1.8)$$

$$\mathbf{C}_i^n = \frac{n-1}{n} \mathbf{C}_i^{n-1} + \frac{1}{n} \vec{x}_n \vec{x}_n^T, \quad n > 1 \quad (1.9)$$

$$\vec{m}_i^n = \frac{n-1}{n} \vec{m}_i^{n-1} + \frac{1}{n} \vec{x}_n, \quad n > 1 \quad (1.10)$$

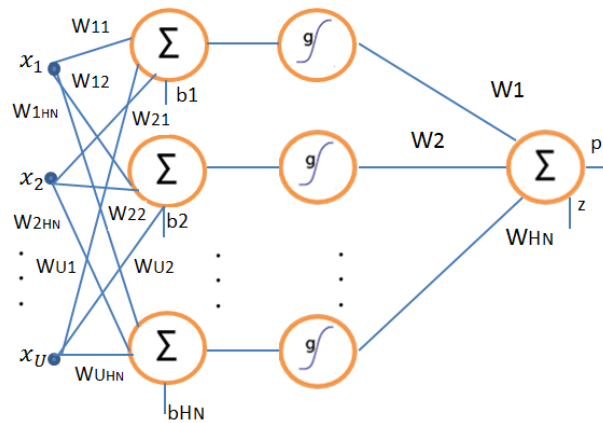
όπου \mathbf{C}_i^n και \vec{m}_i^n είναι ο πίνακας συνδιασποράς και το διάνυσμα της μέσης ανταπόκρισης αντίστοιχα για το κόμμα i στην επανάληψη n , και \vec{x}_n είναι το διάνυσμα των απαντήσεων του n -οστού χρήστη στο ερωτηματολόγιο για το κόμμα i .

Τεχνητά Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης

Τα τεχνητά Νευρωνικά Δίκτυα είναι αρχιτεκτονικές δομές (δίκτυα) που εμπνέονται από τον τρόπο λειτουργίας του εγκεφάλου για τη δημιουργία καλύτερων υπολογιστικών συστημάτων. Ο ανθρώπινος εγκέφαλος είναι κατασκευασμένος από ένα μεγάλο πλήθος τεχνητών νευρώνων, οι οποίοι λειτουργούν παράλληλα και ουσιαστικά παριστάνουν τις διασυνδεδεμένες μονάδες επεξεργασίας. Οι μονάδες αυτές επεξεργάζονται τις εισόδους για να δημιουργήσουν το αποτέλεσμα (την έξοδο) και να το μεταβιβάσουν στην επόμενη συνδεδεμένη μονάδα επεξεργασίας. Η μνήμη βρίσκεται στις συνάψεις μεταξύ των νευρώνων [26].

Στην παρούσα εργασία χρησιμοποιούνται νευρωνικά δίκτυα πρόσθιας τροφοδότησης (feed forward Neural Networks-NN), δεδομένου ότι μπορούν να προσαρμόζουν καλά τα διαθέσιμα στοιχεία. Ονομάζονται πρόσθιας τροφοδότησης αφού όλοι οι υπολογισμοί που γίνονται ακολουθούν την ίδια κατεύθυνση, από την είσοδο προς την έξοδο. Στο Σχήμα 1.7 απεικονίζεται η συνδεσμολογία ενός νευρωνικού δικτύου πρόσθιας τροφοδότησης και οι υπολογισμοί που ακολουθούνται για την εκτίμηση της πρόθεσης ψήφου του χρήστη.

Τα δίκτυα αυτά αποτελούνται από ένα σύνολο εισόδων (στην προκειμένη περίπτωση ως είσοδο λαμβάνονται οι U απαντήσεις που έδωσαν οι χρήστες στο ερωτηματολόγιο του VAA), ένα κρυφό στρώμα h



Σχήμα 1.7: Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης για την πρόβλεψη της πρόθεσης ψήφου

που αποτελείται από HN κρυφούς νευρώνες και μια έξοδο. Για την εκπόνηση της διατριβής η έξοδος αντιπροσωπεύει το κόμμα i που δήλωσε ο κάθε χρήστης πως προτίθεται να ψηφίσει. Κάθε είσοδος x_k σταθμίζεται με ένα w_{ij} βάρος, κατευθύνεται προς τον $j - th$ κρυμμένο νευρώνα και αθροίζεται (μαζί με ένα shifting bias b_j) που παράγει μια ποιότητα y_j . Στη συνέχεια, μια συνάρτηση $g(y_j)$ σιγμοειδής ενεργοποιείται οδηγώντας από τη y_j στη h_j , η οποία είναι η έξοδος του κρυφού νευρώνα j . Οι έξοδοι όλων των κρυμμένων νευρώνων σταθμίζονται από τις παραμέτρους w_j και αθροίζονται (μαζί με ένα shifting bias z) για να παραχθεί η τελική p_i έξοδος. Τα μοντέλα των νευρωνικών δικτύων εκπαιδεύτηκαν δέκα φορές για κάθε χρήστη και διατηρήθηκε η μέση απόδοση καθενός από αυτά.

Βαθιά Μάθηση

Οι συμβατικές τεχνικές μηχανικής μάθησης έχουν περιορισμένη ικανότητά στο να επεξεργάζονται φυσικά δεδομένα στην ακατέργαστη μορφή τους. Οι μέθοδοι βαθιάς μάθησης (deep learning) επιτρέπουν την εισαγωγή ακατέργαστων δεδομένων και την αυτόματη επεξεργασία τους για την μετέπειτα ταξινόμηση τους στις διάφορες κατηγορίες. Η βαθιά μάθηση έχει αποδειχθεί πολύ καλή στον εντοπισμό περίπλοκων δομών σε δεδομένα μεγάλων διαστάσεων [79].

Πολλές εφαρμογές βαθιάς μάθησης χρησιμοποιούν αρχιτεκτονικές νευρωνικών δικτύων πρόσθιας τροφοδότησης. Ως εκ τούτου τα μοντέλα που δημιουργούνται συνήθως καλούνται βαθιά νευρωνικά δίκτυα (deep neural networks), με τον όρο «βαθύ» να αναφέρεται στον αριθμό των κρυμμένων στρωμάτων που βρίσκονται στο νευρωνικό δίκτυο. Οι υπολογισμοί εκτελούνται από πολλά στρώματα μαθαίνοντας κατευθείαν τα χαρακτηριστικά του δείγματος από τα δεδομένα, χωρίς να χρειάζεται να καθοριστούν οι τιμές από πριν ([24] pp.417-466).

Μηχανές Διανυσμάτων Υποστήριξης

Οι Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines- SVMs) είναι μια δημοφιλής τεχνική μηχανικής μάθησης που χρησιμοποιείται για να ταξινομήσει τα δεδομένα σε μια από δύο κατηγορίες. Για να συμβεί αυτό οι SVMs υπολογίζουν ένα υπερεπίπεδο (hyperplane), το οποίο διαχωρίζει τα δεδομένα των δύο κατηγοριών αφήνοντας τη μεγαλύτερη δυνατή απόσταση μεταξύ των σημείων της μιας κατηγορίας που βρίσκονται πιο κοντά στην άλλη κατηγορία. Τα σημεία αυτά ονομάζονται διανύσματα υποστήριξης (support vectors) [76, 77]. Ο στόχος του SVM είναι να ταξινομήσει τα νέα δεδομένα σε μια από τις δύο κατηγορίες, ανάλογα με την πλευρά του υπερεπιπέδου που εκτιμήθηκε ότι ανήκουν τα

δεδομένα αυτά.

Στις περιπτώσεις που δεν υπάρχει ένα απλό υπερεπίπεδο για να χρησιμοποιηθεί ως κριτήριο διαχωρισμού, οι SVMs εφαρμόζουν μη γραμμική ταξινόμηση προσεγγίζοντας το πρόβλημα ταξινόμησης με συναρτήσεις πυρήνα (kernel functions) [78, 77]. Υπό αυτές τις συνθήκες τα διανύσματα εισόδου μετασχηματίζονται για να μεταφερθούν σε ένα χώρο χαρακτηριστικών υψηλότερης διάστασης όπου ο διαχωρισμός γίνεται πιο εύκολα.

Το πρόβλημα ταξινόμησης που καλείται να λύσει το SVAA είναι η κατανομή των χρηστών σε ένα από τα κόμματα που εμπεριέχονται στο σχεδιασμό του VAA. Το πρόβλημα αυτό περιέχει περισσότερες από μια κατηγορίες, μιας και ο αριθμός των κόμματων σε όλες τις χώρες είναι μεγαλύτερος του δύο. Έτσι στην παρούσα εργασία γίνεται χρήση της βιβλιοθήκης LIBSVM [76], η οποία υποστηρίζει την ταξινόμηση SVM πολλαπλών κλάσεων και εφαρμόζει την προσέγγιση “one-against-one”. Αν D είναι ο αριθμός των πολιτικών κομμάτων τότε δημιουργούνται $\frac{D(D-1)}{2}$ ταξινομητές και κάθε ένας εκπαιδεύει τα δεδομένα από δύο ομάδες. Κάθε δυαδική ταξινόμηση δίνει μία ψήφο στη νικήτρια ομάδα και ο χρήστης ταξινομείται στην ομάδα κόμματος με τον μέγιστο αριθμό ψήφων.

1.7.2 Εκπαίδευση, δοκιμή και αξιολόγηση

Κάθε ερευνητικό ερώτημα που καταπιάνεται με τεχνικές μηχανικές μάθησης αντιμετωπίζεται ως ένα πρόβλημα μάθησης. Για την εκπαίδευση του ταξινομητή του προβλήματος αυτού γίνεται χρήση του συνόλου εκπαίδευσης (training set) $\mathcal{D}_{train} = \{(\vec{x}_j, v_j) | j = 1 \dots N_l, v_j \neq \emptyset\}$, το οποίο είναι ένα σύνολο διανυσμάτων \vec{x}_j , που αναφέρεται στις απαντήσεις των χρηστών στο EUvox, μαζί με την αντίστοιχη πρόθεση ψήφου v_j . Εδώ λαμβάνονται υπόψη μόνο οι χρήστες που εξέφρασαν πρόθεση ψήφου για ένα συγκεκριμένο κόμμα, έτσι $v_j \neq \emptyset$.

Αφού το σύστημα εκπαιδευτεί, είναι απαραίτητο να διεξαχθούν κάποιες δοκιμές έτσι ώστε να αξιολογηθεί η απόδοση του. Οι δοκιμές αυτές γίνονται με τη χρήση ενός συνόλου δοκιμής (testing set) $\mathcal{D}_{test} = \{(\vec{x}_t, v_t) | (\vec{x}_t, v_t) \notin \mathcal{D}_{train}, t = 1 \dots N_t, v_t \neq \emptyset\}$ που είναι ένα σύνολο διανυσμάτων \vec{x}_t , τα οποία αναφέρονται στις απαντήσεις των χρηστών που δεν χρησιμοποιούνται στο σύνολο εκπαίδευσης, μαζί με την αντίστοιχη πρόθεση ψήφου v_t .

Η αξιολόγηση της επίδοσης του EUvox όσον αφορά την πρόβλεψη ψήφου, κατά πόσο δηλαδή ταξινομήθηκε σωστά ο κάθε χρήστης σε ένα συγκεκριμένο κόμμα, το οποίο σημαίνει πως προβλέφθηκε σωστά η πρόθεση ψήφου που δήλωσε ο χρήστης, γίνεται με τη χρήση γνωστών μετρικών που ορίζονται στο πεδίο της ανάκτησης πληροφορίας [80]:

$$Precision = P_r^q = \frac{|\mathcal{G}_q|}{|\mathcal{G}_q| + |\mathcal{F}_q|} \quad (1.11)$$

$$Recall = R_e^q = \frac{|\mathcal{G}_q|}{|\mathcal{X}_q|} \quad (1.12)$$

$$Fmeasure = F_1^q = \frac{2 \cdot R_e^q \cdot P_r^q}{R_e^q + P_r^q} \quad (1.13)$$

όπου \mathcal{X}_q είναι το σύνολο των χρηστών που εξέφρασαν πρόθεση ψήφου για το κόμμα p_q , \mathcal{G}_q είναι το σύνολο των χρηστών που εξέφρασαν πρόθεση ψήφου για το κόμμα p_q και τους έγινε (σωστά) ταξινόμηση στο κόμμα p_q και \mathcal{F}_q είναι το σύνολο των χρηστών οι οποίοι εξέφρασαν πρόθεση ψήφου για ένα κόμμα διαφορετικό από το p_q , αλλά ταξινομήθηκαν (εσφαλμένα) στο κόμμα p_q . Η αξιολόγηση γίνεται

ανά κόμμα και συνολικά, με τις συνολικές τιμές των Precision (P_r), Recall (R_e) και F-measure (F_1) να υπολογίζονται ως τα σταθμισμένα αθροίσματα των ατόμων που βρίσκονται σε κάθε κόμμα. Με $| \cdot |$ συμβολίζεται ο πληθάρσιμος του αντίστοιχου συνόλου.

Επίσης υπολογίζεται ο αριθμητικός μέσος ακριβείας (Mean Average Precision - MAP) για κάθε σύνολο δεδομένων. Ο αριθμητικός μέσος ακριβείας υπολογίζει τη μέση σειρά κατάταξης που πήρε η πραγματική ψήφος στη λίστα κατάταξης που δημιουργήθηκε από το SVAA, δείχνοντας στην πρώτη θέση το κόμμα που το SVAA προτείνει περισσότερο στο χρήστη και στην τελευταία το κόμμα που προτείνει λιγότερο, για όλους τους χρήστες που βρίσκονται στα δεδομένα δοκιμών. Αυτό είναι χρήσιμο γιατί στην περίπτωση που η πραγματική πρόθεση ψήφου του χρήστη δεν ταξινομείται στην πρώτη θέση της λίστας κατάταξης του SVAA αλλά στη δεύτερη ή στην τρίτη θέση, ο χρήστης πάλι μένει ικανοποιημένος [81].

1.7.3 Δεδομένα

Κύριο μέλημα αυτής της διατριβής είναι να επεξεργαστεί τις απαντήσεις που έδωσαν Ευρωπαίοι πολίτες στο EUnox, το πανευρωπαϊκό VAA που τέθηκε σε εφαρμογή για τις εκλογές του 2014 για το Ευρωπαϊκό Κοινοβούλιο, για τη διερεύνηση του προφίλ του Ευρωπαίου Ψηφοφόρου και τη βελτιστοποίηση της σύστασης ψήφου. Οι χρήστες του EUnox έπρεπε να υποβάλουν τη γνώμη τους σε 30 ερωτήσεις που αφορούσαν διάφορα πολιτικά ζητήματα σχετικά με την Ευρωπαϊκή Ένωση (ΕΕ). Οι ερωτήσεις αυτές ήταν διαφορετικές για κάθε χώρα, παρόλο που υπήρχε ένα κοινό υποσύνολο, και διαμορφώνονταν ανάλογα με τα τρέχοντα θέματα που απασχολούσαν το έθνος σε σχέση με την ΕΕ.

Επιπλέον των 30 ερωτήσεων, οι χρήστες είχαν την επιλογή να ανταποκριθούν σε ορισμένες συμπληρωματικές ερωτήσεις, κάποιες από τις οποίες ζητούσαν τις δημογραφικές τους πληροφορίες, την πρόθεση ψήφου τους, το ενδιαφέρον τους για την πολιτική, τα κριτήρια με τα οποία ψηφίζουν και την αυτοτοποθέτηση τους στον πολιτικό χάρτη (αριστερά/ δεξιά, παραδοσιακή/ φιλελεύθερη). Τα δεδομένα προεπεξεργάστηκαν, προκειμένου να εξαιρεθούν από το δείγμα όσοι χρήστες δε δήλωσαν την πρόθεση ψήφου τους και όσοι δεν κατάφεραν να απαντήσουν σε όλες τις 30 ερωτήσεις. Η πρόθεση ψήφου είναι αναγκαία για τα πειράματα που διεξάγονται κατά την εκπόνηση της συγκεκριμένης διατριβής, έτσι ώστε να μετρηθεί η απόδοση των τεχνικών που προτείνονται για την επίλυση των ερευνητικών ερωτημάτων.

Σε όλες τις χώρες, το 40% περίπου των χρηστών δήλωσε πρόθεση ψήφου. Από εκείνους το 70% περίπου απάντησε σε όλες τις συμπληρωματικές ερωτήσεις. Ο Πίνακας 1.5 παρουσιάζει τα δεδομένα που χρησιμοποιήθηκαν, όπως αυτά προέκυψαν μετά την προ-επεξεργασία. Στο Παράρτημα I βρίσκονται τα ονόματα που αντιστοιχούν σε κάθε κόμμα.

Πίνακας 1.5: Δεδομένα

Σύνολο Δεδομένων	# Δειγμάτων	# Δειγμάτων στο Σύνολο Εκπ/σης	# Δειγμάτων στο Σύνολο Δοκιμής	# κομμάτων που μοντελοποιήθηκαν
Αυστρία	5030	3018	2012	8
Βουλγαρία	2570	1557	1013	5
Κύπρος	1839	1103	736	6
Τσεχία	15278	9167	6111	10
Γερμανία	5180	3108	2072	7
Δανία	53284	31970	21314	7
Εσθονία	8325	4995	3330	6
ΗΒ-Αγγλία	45523	27281	18242	5
ΗΒ-Βόρεια Ιρλανδία	1455	875	580	6
ΗΒ-Σκωτία	4749	2846	1903	6
ΗΒ-Ουαλία	2736	1640	1096	6
Ισπανία	107025	64215	42810	12
Φινλανδία	3611	2167	1444	8
Γαλλία	3818	2291	1527	7
Ελλάδα	26243	15746	10497	9
Κροατία	2512	1507	1005	5
Ουγγαρία	3138	1884	1254	5
Ιρλανδία	4688	2818	1870	7
Ιταλία	15906	9544	6362	8
Λιθουανία	3525	2115	1410	4
Λετονία	1319	791	528	3
Ολλανδία	815	489	326	6
Πολωνία	36590	21981	14609	8
Πορτογαλία	21682	13111	8571	6
Ρουμανία	3896	2343	1553	5
Σουηδία	1978	1187	791	10
Σλοβακία	2776	1657	1119	8

Κεφάλαιο 2

Χρήση Μηχανικής Μάθησης στα VAAs

2.1 Δημιουργία της σύστασης ψήφου με βάση τις ερωτήσεις του VAA

Η ενότητα αυτή εξετάζει την απόδοση του VAA όταν χρησιμοποιούνται μόνο οι δηλώσεις πολιτικής του online ερωτηματολογίου. Γίνεται αναφορά στις τεχνικές μηχανικής μάθησης που χρησιμοποιούνται και στην αποτελεσματικότητά τους. Στη συνέχεια γίνεται σύγκριση μεταξύ του παραδοσιακού τρόπου σύστασης και του κοινωνικού.

2.1.1 Οι παράγοντες που επηρεάζουν τη δημιουργία της σύστασης ψήφου

Ένα μεγάλο μέρος των ερευνητών που ασχολείται τόσο με τα ΣΣ (Bobadilla et al. [82], Breese et al. [20], Ricci et al. [30]) όσο και με τα VAAs (Katakis et al. [63], Louwerse & Rosema [43], Teran et al. [83], Tsapatsoulis et al. [81]), επικεντρώνεται στις μεθόδους και στους αλγορίθμους που χρησιμοποιούνται για την παραγωγή των συστάσεων τους. Ο τρόπος με τον οποίο δημιουργείται μια σύσταση για ένα συγκεκριμένο χρήστη, εξαρτάται από πολλούς παράγοντες. Ένας παράγοντας είναι η μορφή που έχουν τα δεδομένα στη βάση δεδομένων. Τα δεδομένα μπορούν να εμφανίζονται με τη μορφή αξιολογήσεων, οι οποίες γίνονται από τους χρήστες στα προϊόντα ή στις υπηρεσίες που χρησιμοποιούν, μπορούν να είναι πληροφορίες που έδωσαν οι χρήστες κατά την εγγραφή τους στο σύστημα, όπως επίσης πιθανόν να πηγάζουν από τις κοινωνικές σχέσεις μεταξύ των χρηστών. Ανάλογα με τον τρόπο που εμφανίζονται τα δεδομένα στο σύστημα, γίνεται και η κατάλληλη επεξεργασία χρησιμοποιώντας τις σωστές τεχνικές για κάθε περίπτωση (Bambini et al. [84]).

Στην περίπτωση του EUvox, τα δεδομένα που αφορούν τις δηλώσεις πολιτικής παίρνουν τις τιμές $\mathcal{L} = \{1, 2, 3, 4, 5, 6\}$, όπου με 1 και 2 εκφράζεται ο βαθμός με τον οποίο συμφωνεί ο χρήστης με το περιεχόμενο της δήλωσης, με 4 και 5 δηλώνεται το κατά πόσο διαφωνεί με αυτό και με την τιμή 3 δίνεται η ουδέτερη στάση του. Η έκφραση “δεν έχω άποψη” συμβολίζεται με την τιμή 6. Στο Σχήμα 2.1 δίνεται ένα παράδειγμα ερώτησης όπως εμφανίστηκε στο EUvox 2014 για του Κύπριους πολίτες με τις επιλογές που είχε ο χρήστης για να εκφράσει την άποψη του ως προς το περιεχόμενο της δήλωσης πολιτικής. Η επιλογή “διαφωνώ απόλυτα” καταγράφεται με την τιμή 5 στα δεδομένα του χρήστη, ενώ η επιλογή “συμφωνώ απόλυτα” παίρνει την τιμή 1. Επίσης κάθε ερώτηση εμπίπτει σε μια από τρεις ενότητες: Ευρωπαϊκή Ένωση, Οικονομίας και Πολιτική-Κουλτούρα. Η ερώτηση στο Σχήμα 2.1 βρίσκεται στην κατηγορία Ευρωπαϊκή Ένωση.

EUvox 2014

Σχετικά με εμάς Συχνές ερωτήσεις

← 1 →

ΕΥΡΩΠΑΪΚΗ ΈΝΩΣΗ

Η Κύπρος θα πρέπει να εγκαταλείψει το Ευρώ

Συμφωνώ απόλυτα Συμφωνώ Ούτε συμφωνώ, ούτε διαφωνώ Διαφωνώ Διαφωνώ απόλυτα

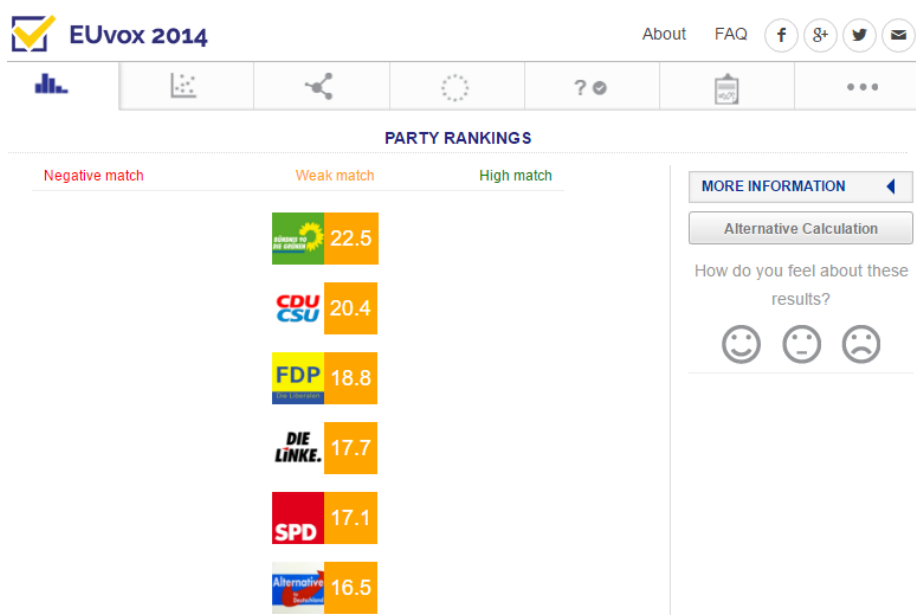
Δεν έχω άποψη

Σχήμα 2.1: Μια δήλωση πολιτικής από το EUvox 2014 μαζί με τις επιλογές που δίνονται στο χρήστη για να απαντήσει

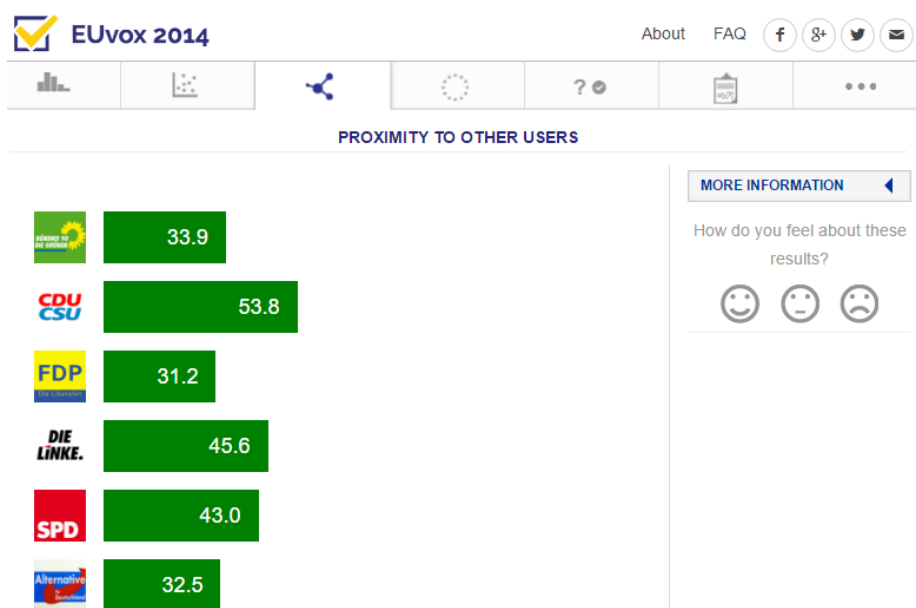
Ένας άλλος παράγοντας που μετρά στο πώς πρόκειται να καταλήξει το σύστημα στη σύσταση του προς το χρήστη, είναι ο τρόπος με τον οποίο προσεγγίζεται το πρόβλημα σύστασης. Για παράδειγμα εάν βασίζεται στο περιεχόμενο του προϊόντος που προτείνει, εάν χρησιμοποιεί τα δημογραφικά χαρακτηριστικά του χρήστη, εάν εκμεταλλεύεται τις κριτικές που έκαναν σε προϊόντα ή υπηρεσίες παρόμοιοι χρήστες με τον εξεταζόμενο χρήστη. Συγκεκριμένα στην τελευταία περίπτωση, που γίνεται σύσταση ανάμεσα σε παρόμοιους χρήστες, τεχνικές συνεργατικού φιλτραρίσματος θα μπορούσαν να χρησιμοποιηθούν για να υπολογίσουν την ομοιότητα/ απόσταση που έχουν οι χρήστες μεταξύ τους, σχετικά με τον τρόπο που αξιολογούν ίδια αντικείμενα. Γνωστά μέτρα για τον υπολογισμό της ομοιότητας/απόστασης ανάμεσα σε δύο χρήστες είναι η Ευκλείδεια απόσταση, η απόσταση Mahalanobis, η απόσταση Minkowski, η απόσταση Manhattan, η συσχέτιση συνημιτόνου και η συσχέτιση Pearson (Amatriain et al. [23], Segaran [77] σελ. 7-28, Witten et al. [24]).

Η παραδοσιακή προσέγγιση που ακολουθεί ένα VAA για τη δημιουργία σύστασης είναι να αξιολογεί την ομοιότητα μεταξύ χρήστη και κόμματος σύμφωνα με τις απαντήσεις που έδωσαν στις δηλώσεις του ερωτηματολογίου και να ταξινομεί σε φθίνουσα σειρά τα κόμματα ανάλογα με την ομοιότητα που υπολογίστηκε, κατατάσσοντας στην πρώτη θέση το κόμμα με τη μεγαλύτερη συσχέτιση με το χρήστη και στην τελευταία αυτό με τη λιγότερη συσχέτιση. Στο Σχήμα 2.2 φαίνεται ένα παράδειγμα από το EUvox 2014, όσον αφορά τα Γερμανικά κόμματα, όπου εμφανίζονται τα αποτελέσματα όπως υπολογίστηκαν με την παραδοσιακή προσέγγιση των VAAs.

Οι Katakis et al. [57] παρατήρησαν ότι οι ψηφοφόροι συχνά δεν συμφωνούν με τις πολιτικές θέσεις του κόμματος / υποψηφίου που προτίθενται να υποστηρίξουν στις επερχόμενες εκλογές, αλλά τείνουν να δίνουν την ψήφο τους επηρεαζόμενοι από την οικογένεια, τους φίλους και την κοινότητα. Ως εκ τούτου, εμπνευσμένοι από την επιρροή που έχει η κοινότητα σε έναν ψηφοφόρο, πρότειναν μια εναλλακτική προσέγγιση για τα VAAs, το λεγόμενο κοινωνικό VAA (social VAA – SVAA), που συγκρίνει τις απαντήσεις ενός χρήστη (δηλαδή, το προφίλ του \vec{x}_j – βλ. εξίσωση 1.1) με εκείνες των άλλων χρηστών του VAA για να εντοπίσει του παρόμοιους χρήστες. Στη συνέχεια, η σύσταση δίνεται με βάση την κατανομή της πρόθεσης ψήφου των παρόμοιων χρηστών. Η λογική πίσω από αυτήν την προσέγγιση είναι πως στην ουσία τα iVAAs είναι ΣΣ που εφαρμόζονται στην πολιτική. Έτσι, σε αυτήν την περίπτωση θα μπορούσαν



Σχήμα 2.2: Ταξινόμηση των κομμάτων βάσει της ομοιότητας μεταξύ κόμματος και χρήστη, όπως υπολογίστηκε με την παραδοσιακή προσέγγιση των VAAs, EUvox 2014



Σχήμα 2.3: Ταξινόμηση των κομμάτων βάσει της ομοιότητας μεταξύ του μοντέλου κόμματος και των απαντήσεων του χρήστη, όπως υπολογίστηκε με την κοινωνική προσέγγιση των VAAs, EUvox 2014

να χρησιμοποιηθούν προσεγγίσεις του συνεργατικού φιλτραρίσματος για τη δημιουργία της σύστασης. Το Σχήμα 2.3 αναφέρεται στο ίδιο παράδειγμα με το Σχήμα 2.2, μόνο που εδώ τα αποτελέσματα εμφανίζονται όπως υπολογίστηκαν με την κοινωνική προσέγγιση των VAAs.

2.1.2 Χρήση τεχνικών μηχανικής μάθησης στο VAA

Όταν το πρόβλημα σύστασης στα VAAs καλείται να επιλυθεί με την κοινωνική προσέγγιση, τεχνικές μηχανικής μάθησης υιοθετούνται για να δημιουργήσουν μοντέλα κόμματος, τα οποία αντιστοιχούν στα κόμματα που παίρνουν μέρος στις εκλογές και περιλαμβάνονται στο σχεδιασμό του VAA. Αυτό που γίνεται ουσιαστικά είναι να ομαδοποιούνται μαζί οι χρήστες που δηλώνουν πρόθεση ψήφου για το ίδιο κόμμα και στη συνέχεια να δημιουργούνται τα μοντέλα κόμματος, που δείχνουν τον τρόπο με τον οποίο οι χρήστες που βρίσκονται στην ίδια ομάδα, συμπληρώνουν το online ερωτηματολόγιο. Στο τέλος, το SVAA, δημιουργεί μια λίστα κατάταξης για κάθε καινούργιο χρήστη, ο οποίος απαντά στο online ερωτηματολόγιο, όπου στην πρώτη θέση βρίσκεται το κόμμα που το SVAA προτείνει περισσότερο στο χρήστη, μιας και το μοντέλο του κόμματος αυτού ταιριάζει περισσότερο με τις απαντήσεις που έδωσε, και στην τελευταία το κόμμα που προτείνει λιγότερο σε αυτόν. Υπό αυτές τις συνθήκες προσεγγίσεις του συνεργατικού φιλτραρίσματος εφαρμόζονται για τη δημιουργία της σύστασης.

Για να γίνουν προτάσεις σε χρήστες από τα ΣΣ που προσεγγίζουν το πρόβλημα σύστασης με το συνεργατικό φιλτράρισμα, χρειάζεται να υπολογιστεί η ομοιότητα $sim(c_a, c_b)$ μεταξύ ενός χρήστη c_a και ενός c_b , βασισμένη στα αντικείμενα που και οι δύο χρήστες βαθμολόγησαν σε σχέση με το πόσο ικανοποιήθηκαν από αυτά. Οι πιο συνηθισμένοι τρόποι υπολογισμού της ομοιότητας μεταξύ των χρηστών είναι με τις συσχετίσεις Pearson και συνημίτονου μέτρου [29]. Ο υπολογισμός της ομοιότητας με βάση το συνημίτονο χρησιμοποιείται συχνά στην ανάκτηση πληροφορίας και ορίζεται ως ακολούθως:

$$sim(ca, cb) = \cos(x_{ca}, x_{cb}) = \frac{\sum_{i \in I_{cacb}} (r_{cai} r_{cbi})}{\sqrt{\sum_{i \in I_{ca}} r_{cai}^2 \sum_{j \in I_{cb}} r_{cbj}^2}} \quad (2.1)$$

Όπου x_{ca} και x_{cb} είναι τα διανύσματα του c_a και c_b χρήστη αντίστοιχα, και $x_{ca} = r_{cai}$ και $x_{cb} = r_{cbi}$ αν οι χρήστες αυτοί βαθμολόγησαν το αντικείμενο i , διαφορετικά ισούνται με μηδέν. I_{cacb} είναι τα αντικείμενα που βαθμολόγησαν και οι δύο χρήστες μαζί. Η ομοιότητα δυο χρηστών με τη συσχέτιση Pearson υπολογίζεται ως εξής:

$$sim(ca, cb) = pearson(ca, cb) = \frac{\sum_{i \in I_{cacb}} (r_{cai} - \bar{r}_{ca})(r_{cbi} - \bar{r}_{cb})}{\sqrt{\sum_{i \in I_{ca}} (r_{cai} - \bar{r}_{ca})^2 \sum_{i \in I_{cb}} (r_{cbi} - \bar{r}_{cb})^2}} \quad (2.2)$$

Όπου r_{cai} και r_{cbi} είναι οι βαθμολογίες που έδωσαν οι συγκεκριμένοι χρήστες στο αντικείμενο i , και \bar{r}_{ca} , \bar{r}_{cb} είναι η μέση βαθμολογία του κάθε χρήστη. Και οι δύο μέθοδοι παράγουν τιμές $sim(ca, cb) \in [-1, 1]$.

Όσο πιο κοντά στο 1 είναι η συσχέτιση τόσο μεγαλύτερη η ομοιότητα μεταξύ των χρηστών. Εάν η συσχέτιση είναι αρνητική σημαίνει ότι όταν ο ένας χρήστης βαθμολογεί ένα αντικείμενο με υψηλή βαθμολογία, τότε ο άλλος το αξιολογεί με χαμηλή και αντίστροφα, ενώ θετική συσχέτιση δείχνει ότι οι μεταβολές των αξιολογήσεων έχουν την ίδια κατεύθυνση [20, 85].

2.1.3 Η απόδοση της πρόβλεψης ψήφου με χρήση μόνο των απαντήσεων στις δηλώσεις πολιτικής

Σε αυτήν την ενότητα εξετάζεται κατά πόσο μπορεί να προβλεφθεί με ακρίβεια η πρόβλεψη ψήφου ενός χρήστη, χρησιμοποιώντας μόνο τις απαντήσεις στις δηλώσεις πολιτικής που έδωσε στο ερωτηματολόγιο του VAA (δηλαδή χωρίς να χρησιμοποιείται οποιαδήποτε άλλη πληροφορία που μπορεί να δώσει ο χρήστης, όπως είναι το ενδιαφέρον για την πολιτική ή η αυτοτοποθέτηση του στον πολιτικό χάρτη). Οι χρήστες προέρχονται από διαφορετικές χώρες της ΕΕ, οπότε γίνεται η εκτίμηση της απόδοσης της πρόβλεψης ψήφου για κάθε χώρα ξεχωριστά και στη συνέχεια πραγματοποιείται σύγκριση σε πανευρωπαϊκό επίπεδο. Με αυτόν τον τρόπο μπορεί να εξεταστεί εάν η απόδοση του VAA, όταν χρησιμοποιούνται μόνο οι απαντήσεις στις δηλώσεις πολιτικής, διαφέρει από χώρα σε χώρα. Η διερεύνηση αυτού του προβλήματος μπορεί να γίνει είτε με την παραδοσιακή είτε με την κοινωνική προσέγγιση των VAA.

Για την παρούσα εργασία λήφθηκαν υπόψη μόνο οι χρήστες που δήλωσαν την πρόθεση ψήφου τους στη συμπληρωματική ερώτηση του VAA ερωτηματολογίου (βλέπε Σχήμα 1.6). Για τη μέτρηση της αποδοτικότητας των τεχνικών που εφαρμόστηκαν, η πρόθεση ψήφου των χρηστών που βρίσκονται στο σύνολο δοκιμής (βλέπε Ενότητα 1.7.2) αρχικά αποκρύπτεται από το σύστημα και αποκαλύπτεται αργότερα, όταν παράγεται η σύσταση από το σύστημα. Με αυτόν τον τρόπο φανερώνεται εάν το πολιτικό κόμμα που προτείνεται από το σύστημα είναι το ίδιο ή διαφέρει από αυτό που δήλωσε ο χρήστης με την πρόθεση ψήφου του.

Η παραδοσιακή προσέγγιση

Στην περίπτωση που χρησιμοποιείται ο παραδοσιακός τρόπος σύστασης, οι απαντήσεις του κάθε χρήστη στις δηλώσεις πολιτικής συγκρίνονται με τις απαντήσεις των κομμάτων για να εκτιμηθεί η άγνωστη συνάφεια $h(j, i)$ του χρήστη j για το κόμμα i . Στη συνέχεια γίνεται κατάταξη των πολιτικών κομμάτων με βάση τη συνάφεια προς το χρήστη και το κόμμα. Ο πιο απλός τρόπος να εκτιμηθεί η $\hat{h}(\vec{x}, \vec{p})$ είναι με τη βοήθεια ενός μέτρου απόστασης $d(\vec{x}, \vec{p})$, όπως περιγράφεται στην εξίσωση 1.4. Στα αποτελέσματα που παρουσιάζονται στη συνέχεια, στην παρούσα ενότητα, ο υπολογισμός της ομοιότητας μεταξύ των απαντήσεων των χρηστών και των απαντήσεων των κομμάτων, γίνεται με τη συσχέτιση Pearson¹. Τα συνολικά αποτελέσματα για κάθε χώρα ξεχωριστά παρουσιάζονται στους Πίνακες 2.1-2.7. Τα αποτελέσματα για κάθε κόμμα σε κάθε χώρα παρατίθενται στο Παράρτημα II. Τα μηδενικά που υπάρχουν στα κόμματα των πινάκων του Παραρτήματος II, δείχνουν ότι δεν κατάφερε να ταξινομηθεί κανένας χρήστης στα συγκεκριμένα κόμματα με την αντίστοιχη μέθοδο. Το “Party Coding” αναφέρεται στην παραδοσιακή προσέγγιση, ενώ όλες οι άλλες τεχνικές μηχανικής μάθησης αφορούν την κοινωνική προσέγγιση.

Η απόδοση του “Party Coding”, ουσιαστικά, μπορεί να δείξει κατά πόσο ο χρήστης γνωρίζει τις θέσεις των κομμάτων στα πολιτικά ζητήματα που θέτει το ερωτηματολόγιο του VAA. Η παραδοσιακή προσέγγιση είναι πολύ απλή και αφορά τη σύγκριση μεταξύ των απαντήσεων που έδωσαν οι χρήστες και τα κόμματα (ή οι ειδικοί ερευνητές που απάντησαν εκ μέρους των κομμάτων/υποψηφίων στα θέματα αυτά). Η υψηλή απόδοση αυτής της τεχνικής αντιστοιχεί σε υψηλή συσχέτιση μεταξύ των απαντήσεων των κομμάτων και των χρηστών, πράγμα που δείχνει πως χρήστες και κόμματα απαντούν με τον ίδιο τρόπο και ασπάζονται τις ίδιες πολιτικές θέσεις.

Η απλότητα της παραδοσιακής μεθόδου είναι και ο κύριος λόγος που η μέθοδος αυτή υστερεί μπροστά στην κοινωνική. Είναι ξεκάθαρο, από τα αποτελέσματα, πως η κοινωνική προσέγγιση υπερτερεί της παραδοσιακής, αφού σε όλες τις περιπτώσεις τεχνικές που εφαρμόστηκαν για την κοινωνική προσέγγιση είχαν καλύτερη απόδοση, με τις Μηχανές Διανυσμάτων Υποστήριξης (SVMs) να εμφανίζουν τα

¹https://en.wikipedia.org/wiki/Pearson_correlation_coefficient

καλύτερα αποτελέσματα. Αυτό που διαφέρει από χώρα σε χώρα είναι η διαφορά μεταξύ της απόδοσης των τεχνικών μηχανικής μάθησης και του “Party Coding”. Δηλαδή, υπήρχαν περιπτώσεις που το Party Coding ήταν καλύτερο ή ισοδύναμο με κάποιες από τις τεχνικές μηχανικής μάθησης που χρησιμοποιήθηκαν.

Πίνακας 2.1: Τα συνολικά αποτελέσματα των ταξινομητών για Αυστρία, Βουλγαρία, Κύπρος και Τσεχία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
Αυστρία	HMM	0.4563	0.4618	0.4590	0.6563
	Naive Bayes	0.4448	0.4610	0.4527	0.6470
	KNN	0.4056	0.4180	0.4117	0.5888
	Mahalanobis	0.4414	0.5177	0.4765	0.6351
	NN3	0.4737	0.3509	0.4032	0.6693
	NN5	0.4975	0.4770	0.4870	0.6852
	NN7	0.4980	0.4812	0.4895	0.6838
	SVM	0.5000	0.5024	0.5012	0.6899
	Party Coding	0.3752	0.4083	0.3911	0.5994
Βουλγαρία	HMM	0.4174	0.4933	0.4522	0.6246
	Naive Bayes	0.4501	0.5172	0.4813	0.6575
	KNN	0.4474	0.4185	0.4324	0.6343
	Mahalanobis	0.3648	0.4495	0.4027	0.6016
	NN1	0.4510	0.3319	0.3824	0.6452
	NN2	0.4909	0.4092	0.4464	0.6792
	NN3	0.5064	0.4446	0.4735	0.6917
	NN4	0.4991	0.4630	0.4804	0.6870
	SVM	0.4991	0.4784	0.4886	0.6916
Party Coding	0.3430	0.3881	0.3642	0.5599	
Κύπρος	HMM	0.5968	0.6525	0.6234	0.7306
	Naive Bayes	0.6034	0.5951	0.5993	0.7355
	KNN	0.6166	0.5543	0.5838	0.6992
	Mahalanobis	0.6100	0.5936	0.6017	0.7384
	NN4	0.6219	0.5968	0.6091	0.7504
	NN6	0.6324	0.5943	0.6128	0.7545
	NN8	0.6337	0.5940	0.6132	0.7548
	NN9	0.6219	0.5900	0.6055	0.7475
	SVM	0.6667	0.6153	0.6399	0.7766
Party Coding	0.4875	0.6247	0.5746	0.6554	
Τσεχία	HMM	0.4934	0.5062	0.4997	0.6683
	Naive Bayes	0.4901	0.4743	0.4821	0.6674
	KNN	0.4101	0.4283	0.4190	0.5735
	Mahalanobis	0.4428	0.4856	0.4632	0.6277
	NN10	0.5022	0.4738	0.4876	0.6742
	NN11	0.5029	0.4768	0.4895	0.6757
	NN20	0.5086	0.4859	0.4970	0.6793
	NN26	0.5047	0.5011	0.5029	0.6721
	SVM	0.5163	0.5042	0.5102	0.6881
Party Coding	0.3980	0.4280	0.4124	0.5887	

Πίνακας 2.2: Τα συνολικά αποτελέσματα των ταξινομητών για Γερμανία, Δανία, Εσθονία και Ηνωμένο Βασίλειο-Αγγλία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
Γερμανία	HMM	0.5458	0.5811	0.5629	0.7203
	Naive Bayes	0.5333	0.5423	0.5378	0.7098
	KNN	0.4913	0.4995	0.4954	0.6625
	Mahalanobis	0.5106	0.5341	0.5221	0.6856
	NN3	0.5319	0.4200	0.4694	0.7093
	NN5	0.5792	0.5518	0.5651	0.7372
	NN10	0.5883	0.5817	0.5850	0.7435
	NN15	0.5719	0.6038	0.5874	0.7300
	SVM	0.5999	0.5837	0.5917	0.7527
	Party Coding	0.4913	0.4731	0.4820	0.6788
Δανία	HMM	0.4747	0.5647	0.5158	0.6772
	Naive Bayes	0.5189	0.5370	0.5278	0.7051
	KNN	0.4698	0.4772	0.4735	0.6559
	Mahalanobis	0.4671	0.5457	0.5033	0.6641
	NN3	0.4735	0.4591	0.4662	0.6795
	NN5	0.5373	0.4998	0.5179	0.7154
	NN10	0.5581	0.5484	0.5532	0.7294
	NN15	0.5593	0.5517	0.5554	0.7314
	SVM	0.5586	0.5493	0.5540	0.7323
	Party Coding	0.4616	0.4517	0.4566	0.6635
Εσθονία	HMM	0.3571	0.3945	0.3748	0.5887
	Naive Bayes	0.3664	0.4089	0.3865	0.6015
	KNN	0.3571	0.3819	0.3691	0.5663
	Mahalanobis	0.3159	0.3830	0.3462	0.5500
	NN3	0.4165	0.3425	0.3759	0.6381
	NN5	0.4408	0.4454	0.4431	0.6547
	NN10	0.4243	0.4230	0.4237	0.6421
	NN15	0.4153	0.4200	0.4177	0.6341
	SVM	0.4411	0.4390	0.4401	0.6532
	Party Coding	0.2928	0.2789	0.2857	0.5209
HB-Αγγλία	HMM	0.5346	0.5595	0.5468	0.7359
	Naive Bayes	0.4989	0.6091	0.5485	0.7157
	KNN	0.5080	0.5127	0.5103	0.6969
	Mahalanobis	0.5408	0.5566	0.5486	0.7344
	NN3	0.5509	0.5091	0.5292	0.7407
	NN5	0.5719	0.5847	0.5782	0.7553
	NN10	0.5752	0.5827	0.5789	0.7577
	SVM	0.5748	0.5896	0.5821	0.7584
	Party Coding	0.4846	0.5148	0.4992	0.6915

Στις χώρες Τσεχία, Γερμανία, Δανία, Σουηδία και Ολλανδία, το Party Coding είχε παρόμοια απόδοση με τη μέθοδο των k -κοντινότερων γειτόνων (KNN). Αυτό φανερώνει πως στις περιπτώσεις αυτές, οι k -κοντινότεροι γείτονες που επιλέχθηκαν για κάθε χρήση σε κάθε κόμμα, αντιπροσωπεύουν την πλειο-

ψηφία των χρηστών του κάθε κόμματος. Έτσι, αν και η παραδοσιακή προσέγγιση στηρίζει την απόδοση της στις απαντήσεις όλων των χρηστών ενώ η KNN τεχνική ασχολείται μόνο με τις απαντήσεις των k χρηστών (που είναι ένας πολύ μικρός αριθμός σε σχέση με όλο το δείγμα), η απόδοση των δύο μεθόδων είναι παρεμφερής.

Όσον αφορά τη Λιθουανία, το Party Coding είχε καλύτερα αποτελέσματα από τις τεχνικές Hidden Markov Models (HMMs) και Mahalanobis classifier. Αυτό συνέβη γιατί τα κόμματα που λήφθηκαν υπόψη στην πειραματική διαδικασία της παρούσας εργασίας ήταν πολύ λίγα σε σχέση με τις άλλες χώρες (μόλις 4 κόμματα) και ο διαχωρισμός μεταξύ των απαντήσεων που έδωσαν οι χρήστες που τα υποστήριζαν (που δήλωσαν δηλαδή πρόθεση ψήφου για τα συγκεκριμένα κόμματα) δεν ήταν ξεκάθαρος. Έτσι, αν και η πλειοψηφία των χρηστών συγκεντρώνεται στο δεύτερο και τρίτο κόμμα, οι συγκεκριμένες προσεγγίσεις είχαν πολύ χαμηλό Recall σε αυτά, πράγμα που δείχνει πως ταξινόμησαν ορθά λιγότερους χρήστες από αυτούς που ανήκαν εκεί πραγματικά. Τοιούτοτρόπως ταξινόμησαν περισσότερους χρήστες από ότι έπρεπε στο πρώτο και τελευταίο κόμμα, καταλήγοντας αυτή τη φορά, πολύ χαμηλό Precision.

Στο δείγμα της Πολωνίας, το Party Coding παρουσίασε παρόμοια συμπεριφορά με τη Λιθουανία για τις τεχνικές HMMs και Mahalanobis, αν και το δείγμα αυτή τη φορά είχε διαφορετικά χαρακτηριστικά: ήταν ένα από τα πιο μεγάλα σε μέγεθος και είχε τα διπλάσια κόμματα από τη Λιθουανία. Εδώ φαίνεται ξεκάθαρα το πρόβλημα που δημιουργεί σε αυτές τις μεθόδους το θέμα της διαφορετικότητας μεταξύ των χρηστών σε κάθε κόμμα, αφού στα κόμματα που υπήρχε συσπείρωση (ιδεολογική ταύτιση) μεταξύ των χρηστών η απόδοση ήταν αρκετά καλή (βλέπε Παράρτημα II, Πίνακας II.23 κόμμα 5). Τα προβλήματα στην απόδοση παρουσιάστηκαν στα κόμματα που χαρακτηρίζονται από πολυμορφία, δηλαδή στις περιπτώσεις που οι χρήστες ενός συγκεκριμένου κόμματος διαφέρουν κατά πολύ μεταξύ τους και οι απαντήσεις τους ποικίλουν ανάλογα με το κοινωνικό και πολιτικό τους υπόβαθρο.

Το Party Coding είχε καλύτερη απόδοση από το Mahalanobis classifier και στο δείγμα της Ολλανδίας. Το μεγάλο μειονέκτημα αυτής της μεθόδου στο συγκεκριμένο σύνολο δεδομένων, είναι πως δεν μπορεί να λειτουργήσει αποτελεσματικά με λίγα δεδομένα και στην περίπτωση αυτή η Ολλανδία έχει το μικρότερο μέγεθος από όλες τις άλλες χώρες. Επίσης η παραδοσιακή σύσταση είχε καλύτερα αποτελέσματα από τη πιθανοτική μέθοδο Naïve Bayes στα δεδομένα της Σκωτίας, της οποίας η απόδοση για τα κόμματα 4 και 5 ήταν εξαιρετικά χαμηλή (βλέπε Παράρτημα II, Πίνακας II.10). Αυτό συνέβη επειδή πολύ πιθανόν να υπάρχει υψηλή συσχέτιση μεταξύ των χρηστών στα κόμματα αυτά, την οποία ο αλγόριθμος δεν κατάφερε να ερμηνεύσει μιας και υποστηρίζει την ανεξαρτησία των χαρακτηριστικών σε κάθε κατηγορία (κόμμα).

Στην Κύπρο αν και η παραδοσιακή προσέγγιση είχε χαμηλό ποσοστό Recall σε σχέση με τις άλλες μεθόδους, το Precision ήταν αρκετά υψηλό. Αυτό συνέβη γιατί, όπως φαίνεται στο Πίνακα II.3, το Precision ήταν εξαιρετικά υψηλό για τα πρώτα δύο κόμματα (80% για το πρώτο κόμμα και 88% για το δεύτερο), πράγμα που δείχνει πως το σύστημα ταξινόμησε σε αυτά τα δύο κόμματα λιγότερους χρήστες από αυτούς που ανήκουν στην πραγματικότητα, αλλά η μεγάλη πλειοψηφία ταξινομήθηκε σωστά. Η συγκεκριμένη παρατήρηση είναι δικαιολογημένη, αφού η Κύπρος έχει ένα μικρό δείγμα χρηστών, οι περισσότεροι εκ των οποίων δήλωσαν πρόθεση ψήφου για τα δύο πρώτα κόμματα.

Ο παραδοσιακός τρόπος σύστασης είχε τη χειρότερη απόδοση στα Ισπανικά δεδομένα, γεγονός που πολύ πιθανόν να δείχνει πως οι Ισπανοί χρήστες δε γνώριζαν σε σημαντικό βαθμό τις θέσεις των πολιτικών κομμάτων που δήλωσαν πως θα υποστηρίξουν. Το δείγμα όμως των χρηστών και ο αριθμός των κομμάτων υπερτερούσαν σε μέγεθος των άλλων χωρών. Αυτό στοιχίζει στη συνολική απόδοση του Party Coding, ειδικά στην περίπτωση που οι χρήστες της κάθε κατηγορίας (κάθε κατηγορία αντιστοιχεί σε ένα από τα κόμματα που έλαβαν μέρος στο σχεδιασμό του VAA) χαρακτηρίζονται από πολυμορφία και δε δίνουν όλοι παρόμοιες απαντήσεις με τα πολιτικά κόμματα. Επιπρόσθετα, ο μεγάλος αριθμός κομμάτων

περιορίζει τις απαντήσεις που μπορούν να δοθούν εκ μέρους τους, με αποτέλεσμα να μην υπάρχει ξεκάθαρος διαχωρισμός μεταξύ των κομμάτων αυτών, όσον αφορά την ανταπόκριση τους στο συγκεκριμένο ερωτηματολόγιο. Κάτι τέτοιο μπορεί να συγχέει τους πολίτες, αφού τις πλείστες φορές οι απαντήσεις που δίνουν είναι παρόμοιες με περισσότερα από ένα κόμματα. Αυτό είναι και το αδύναμο σημείο της παραδοσιακής μεθόδου, αφού στηρίζεται αποκλειστικά στις απαντήσεις χρηστών - κομμάτων.

Η κοινωνική προσέγγιση

Με την κοινωνική προσέγγιση γίνεται αξιοποίηση της κοινότητας του Ευνοχ και αγνοούνται οι πληροφορίες από τα προφίλ των πολιτικών κομμάτων. Η ιδέα πίσω από αυτήν την προσέγγιση είναι πως οι χρήστες που δίνουν παρόμοιες απαντήσεις στο ερωτηματολόγιο του VAA πιθανόν να υποστηρίζουν το ίδιο πολιτικό κόμμα στις επερχόμενες εκλογές. Υπό αυτές τις συνθήκες, για τη σύσταση λαμβάνεται υπόψη μόνο η πρόθεση ψήφου των χρηστών, μετατρέποντας το πρόβλημα σύστασης σε πρόβλημα ταξινόμησης, όπου ο χρήστης συνιστάται να ψηφίσει το κόμμα, στο οποίο έχει τη μεγαλύτερη πιθανότητα να ανήκει. Το κόμμα, δηλαδή, που δήλωσαν πως πρόκειται να υποστηρίξουν οι χρήστες που απάντησαν παρόμοια το ερωτηματολόγιο με αυτόν. Η διαδικασία που ακολουθείται σε αυτήν την περίπτωση είναι η εξής:

- Αρχικά καθορίζονται πόσες ομάδες μπορούν να δημιουργηθούν ανά χώρα. Κάθε ομάδα αντιστοιχεί σε ένα συγκεκριμένο κόμμα που λαμβάνει μέρος στις επερχόμενες εκλογές της χώρας και βρίσκεται στο σχεδιασμό του VAA.
- Στη συνέχεια οι χρήστες που δήλωσαν πρόθεση ψήφου χωρίζονται στο σύνολο δοκιμής και εκπαίδευσης. Ο καθένας από τους χρήστες στο σύνολο εκπαίδευσης τοποθετείται στην ομάδα που αντιπροσωπεύει το κόμμα, το οποίο εξέφρασε πως θα υποστηρίξει με την πρόθεση ψήφου του.
- Από τις απαντήσεις των χρηστών που βρίσκονται στις ομάδες, δημιουργούνται τα μοντέλα κόμματος. Το μοντέλο μιας ομάδας ή αλλιώς το μοντέλο ενός κόμματος αντιπροσωπεύει τον τρόπο με τον οποίο οι υποστηρικτές αυτού του κόμματος συνηθίζουν να απαντούν στο ερωτηματολόγιο του VAA.
- Για το επόμενο βήμα χρησιμοποιούνται οι χρήστες από το σύνολο δοκιμής, έτσι ώστε να συγκριθούν οι απαντήσεις που έδωσαν με κάθε μοντέλο κόμματος ξεχωριστά. Στο τέλος προτείνεται σε κάθε χρήστη από το σύνολο δοκιμής το κόμμα, του οποίου το μοντέλο ταιριάζει καλύτερα με τις απαντήσεις που έδωσε. Τα μοντέλα αυτά μπορούν να είναι είτε γραμμικά είτε μη γραμμικά, συγκρίνοντας την απόδοση σε κάθε περίπτωση ξεχωριστά.
- Για τη δημιουργία των μοντέλων μπορούν να χρησιμοποιηθούν προσεγγίσεις από τη μηχανική μάθηση και όπως ήδη αναφέρθηκε στη Μεθοδολογία της Έρευνας στην Ενότητα 1.7.1 για την παρούσα εργασία υιοθετήθηκαν οι πιο κάτω μέθοδοι:
 1. Κρυμμένα Μοντέλα Markov (HMM).
 2. Η πιθανοτική μέθοδος Naïve Bayes.
 3. Ο αλγόριθμος των k-κοντινότερων γειτόνων (KNN).
 4. Ο ταξινομητής ελάχιστης Mahalanobis απόστασης.
 5. Τεχνητά Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης (NN).
 6. Μηχανές Διανυσμάτων Υποστήριξης (SVM).

Τα αποτελέσματα για κάθε κόμμα σε κάθε χώρα βρίσκονται στο Παράρτημα ΙΙ ενώ τα συγκεντρωτικά αποτελέσματα για κάθε χώρα ξεχωριστά παρουσιάζονται στους Πίνακες 2.1-2.7. Στους πίνακες αυτούς με κίτρινο χρώμα σημειώνεται η τεχνική με την καλύτερη απόδοση.

Πίνακας 2.3: Τα συνολικά αποτελέσματα των ταξινομητών για Ηνωμένο Βασίλειο-Βόρεια Ιρλανδία, Ηνωμένο Βασίλειο-Σκωτία, Ηνωμένο Βασίλειο-Ουαλία και Ισπανία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
HB-Βόρεια Ιρλανδία	HMM	0.4653	0.4794	0.4723	0.6660
	Naive Bayes	0.4455	0.4591	0.4522	0.6637
	KNN	0.4191	0.4096	0.4143	0.6136
	Mahalanobis	0.4604	0.4220	0.4404	0.6527
	NN3	0.4769	0.3428	0.3989	0.6728
	NN5	0.4703	0.4594	0.4648	0.6763
	NN10	0.4934	0.4834	0.4883	0.6801
	NN15	0.4488	0.4346	0.4416	0.6284
	SVM	0.5099	0.5159	0.5129	0.7065
	Party Coding	0.3759	0.3855	0.3806	0.5994
HB-Σκωτία	HMM	0.4115	0.5841	0.4828	0.6287
	Naive Bayes	0.3988	0.4792	0.4353	0.6419
	KNN	0.5123	0.5152	0.5138	0.6959
	Mahalanobis	0.4430	0.5291	0.4822	0.6648
	NN3	0.5281	0.5246	0.5264	0.7278
	NN5	0.5681	0.5553	0.5616	0.7504
	NN10	0.5733	0.5733	0.5733	0.7489
	SVM	0.5686	0.5802	0.5743	0.7549
	Party Coding	0.4808	0.5202	0.4997	0.6847
	HB-Ουαλία	HMM	0.4546	0.4913	0.4723
Naive Bayes		0.4120	0.4411	0.4260	0.6238
KNN		0.4220	0.4272	0.4246	0.6142
Mahalanobis		0.4274	0.4333	0.4303	0.6363
NN3		0.4474	0.3395	0.3860	0.6604
NN5		0.4846	0.4632	0.4736	0.6748
NN10		0.4619	0.4669	0.4644	0.6559
SVM		0.4973	0.5043	0.5008	0.6977
Party Coding		0.3823	0.4038	0.3928	0.6050
Ισπανία		HMM	0.2220	0.3336	0.2666
	Naive Bayes	0.2968	0.2973	0.2971	0.5112
	KNN	0.2696	0.2785	0.274	0.4599
	Mahalanobis	0.2334	0.3699	0.2862	0.4345
	NN3	0.2907	0.2179	0.2491	0.5179
	NN5	0.3258	0.2599	0.2891	0.5421
	NN10	0.3337	0.3271	0.3303	0.5504
	NN15	0.3381	0.3426	0.3403	0.5523
	SVM	0.3718	0.3761	0.3739	0.5923
	Party Coding	0.1480	0.2281	0.1795	0.3743

Παρατηρούμε πως σε όλες τις χώρες, εκτός της Ολλανδίας, η μέγιστη απόδοση επιτεύχθηκε από τις SVM. Επιπρόσθετα στις πλείστες χώρες τα νευρωνικά δίκτυα κατάφεραν τη δεύτερη πιο υψηλή επίδοση, ενώ σε κάποιες περιπτώσεις η απόδοσή τους ήταν σχεδόν ίδια με αυτήν των SVM, όπως συνέβη με τη Δανία, την Εσθονία, την Αγγλία, την Ιταλία, τη Λιθουανία και την Πολωνία.

Πίνακας 2.4: Τα συνολικά αποτελέσματα των ταξινομητών για Φινλανδία, Γαλλία, Ελλάδα και Κροατία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
Φινλανδία	HMM	0.5499	0.6219	0.5837	0.7118
	Naive Bayes	0.5637	0.5361	0.5495	0.7241
	KNN	0.5229	0.5448	0.5336	0.6793
	Mahalanobis	0.5416	0.5124	0.5265	0.7040
	NN3	0.5533	0.4302	0.4841	0.7213
	NN5	0.606	0.5721	0.5885	0.7512
	NN10	0.5914	0.5859	0.5887	0.7374
	SVM	0.6260	0.6221	0.6241	0.7684
	Party Coding	0.4204	0.5384	0.4721	0.6270
Γαλλία	HMM	0.5056	0.5613	0.5320	0.6997
	Naive Bayes	0.5180	0.5339	0.5258	0.7070
	KNN	0.4748	0.4731	0.4740	0.6521
	Mahalanobis	0.4532	0.5160	0.4820	0.6492
	NN3	0.4846	0.4682	0.4763	0.6909
	NN5	0.5671	0.5544	0.5607	0.7355
	NN10	0.5521	0.5352	0.5435	0.7250
	NN15	0.5409	0.5556	0.5481	0.7105
	SVM	0.5704	0.5641	0.5673	0.7433
Party Coding	0.4257	0.5058	0.4623	0.6500	
Ελλάδα	HMM	0.3966	0.5237	0.4514	0.6099
	Naive Bayes	0.5300	0.5041	0.5167	0.6951
	KNN	0.4794	0.4545	0.4666	0.6336
	Mahalanobis	0.3732	0.5178	0.4338	0.5798
	NN3	0.5337	0.4063	0.4614	0.695
	NN5	0.5502	0.4887	0.5176	0.7094
	NN10	0.5563	0.5131	0.5338	0.7153
	NN15	0.5559	0.5156	0.5350	0.7141
	SVM	0.5581	0.5437	0.5508	0.7202
Party Coding	0.3574	0.4783	0.4091	0.5627	
Κροατία	HMM	0.4328	0.5565	0.4870	0.6694
	Naive Bayes	0.5045	0.5161	0.5102	0.7117
	KNN	0.4756	0.4752	0.4754	0.6743
	Mahalanobis	0.5403	0.5109	0.5252	0.7241
	NN3	0.5652	0.4959	0.5283	0.7445
	NN5	0.5522	0.5343	0.5431	0.7356
	NN10	0.5453	0.5206	0.5326	0.7224
	NN15	0.5353	0.5124	0.5236	0.7014
	SVM	0.5920	0.5522	0.5714	0.7609
Party Coding	0.4090	0.5303	0.4618	0.6496	

Πίνακας 2.5: Τα συνολικά αποτελέσματα των ταξινομητών για Ουγγαρία, Ιρλανδία, Ιταλία και Λιθουανία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
Ουγγαρία	HMM	0.5587	0.5986	0.5779	0.7342
	Naive Bayes	0.6003	0.5853	0.5927	0.7647
	KNN	0.5314	0.5589	0.5448	0.6956
	Mahalanobis	0.5473	0.5408	0.5440	0.7251
	NN3	0.5935	0.5278	0.5587	0.7568
	NN5	0.5897	0.5243	0.5551	0.7534
	NN10	0.5549	0.5408	0.5478	0.7260
	NN15	0.5450	0.5429	0.5440	0.7206
	SVM	0.6238	0.5892	0.6060	0.7783
	Party Coding	0.4307	0.5021	0.4637	0.6378
Ιρλανδία	HMM	0.3397	0.3733	0.3557	0.5619
	Naive Bayes	0.3454	0.3425	0.3439	0.5708
	KNN	0.2947	0.2928	0.2937	0.4963
	Mahalanobis	0.3128	0.3764	0.3417	0.5370
	NN3	0.3557	0.2411	0.2874	0.5754
	NN5	0.3816	0.3394	0.3593	0.5961
	NN10	0.3795	0.3636	0.3714	0.5917
	NN15	0.3769	0.3624	0.3695	0.5846
	SVM	0.3919	0.3749	0.3832	0.6034
	Party Coding	0.2674	0.3282	0.2947	0.4865
Ιταλία	HMM	0.5329	0.5807	0.5558	0.7087
	Naive Bayes	0.5247	0.5974	0.5587	0.7137
	KNN	0.5192	0.5171	0.5181	0.6722
	Mahalanobis	0.5341	0.5902	0.5607	0.7092
	NN3	0.5937	0.5231	0.5562	0.7544
	NN5	0.6036	0.5643	0.5833	0.7597
	NN10	0.6091	0.5926	0.6007	0.7630
	NN15	0.6052	0.5830	0.5939	0.7606
	SVM	0.6022	0.5974	0.5998	0.7631
	Party Coding	0.4890	0.5259	0.5068	0.6901
Λιθουανία	HMM	0.2993	0.5882	0.3967	0.5836
	Naive Bayes	0.5426	0.5461	0.5443	0.7462
	KNN	0.5447	0.5478	0.5462	0.7339
	Mahalanobis	0.3447	0.5893	0.4350	0.6077
	NN3	0.6184	0.5974	0.6077	0.7855
	NN5	0.5872	0.5723	0.5797	0.7641
	NN10	0.5667	0.5517	0.5591	0.7500
	NN15	0.5227	0.5127	0.5176	0.7180
	SVM	0.6142	0.5953	0.6046	0.7843
	Party Coding	0.5418	0.5763	0.5586	0.7350

Πίνακας 2.6: Τα συνολικά αποτελέσματα των ταξινομητών για Λετονία, Ολλανδία, Πολωνία και Πορτογαλία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
Λετονία	HMM	0.6799	0.7075	0.6934	0.8327
	Naive Bayes	0.7292	0.7115	0.7202	0.8605
	KNN	0.6818	0.6584	0.6699	0.8176
	Mahalanobis	0.6629	0.6405	0.6515	0.8251
	NN3	0.7254	0.7061	0.7156	0.8580
	NN5	0.6383	0.6535	0.6458	0.8040
	NN10	0.5909	0.6067	0.5987	0.7541
	NN15	0.5455	0.6057	0.5740	0.7383
	SVM	0.7367	0.7084	0.7223	0.8643
	Party Coding	0.6307	0.6709	0.6502	0.8074
Ολλανδία	HMM	0.5859	0.6146	0.5999	0.7472
	Naive Bayes	0.7117	0.7046	0.7081	0.8385
	KNN	0.5828	0.5835	0.5831	0.7443
	Mahalanobis	0.4356	0.5254	0.4763	0.6543
	NN3	0.6043	0.5176	0.5576	0.7626
	NN5	0.5920	0.5722	0.5819	0.7359
	NN10	0.4693	0.4504	0.4597	0.6386
	NN15	0.4540	0.4775	0.4654	0.6465
	SVM	0.6810	0.6855	0.6832	0.8233
	Party Coding	0.5613	0.5950	0.5777	0.7480
Πολωνία	HMM	0.2772	0.6967	0.3966	0.5405
	Naive Bayes	0.6147	0.5746	0.5939	0.7560
	KNN	0.5759	0.5483	0.5618	0.7073
	Mahalanobis	0.4587	0.6393	0.5342	0.6492
	NN3	0.6167	0.4742	0.5361	0.7516
	NN5	0.6299	0.5225	0.5712	0.7583
	NN10	0.6423	0.596	0.6183	0.7734
	NN15	0.6486	0.6088	0.6281	0.7774
	SVM	0.6442	0.6088	0.6260	0.7769
	Party Coding	0.5119	0.6143	0.5584	0.6853
Πορτογαλία	HMM	0.3709	0.4646	0.4125	0.5693
	Naive Bayes	0.4154	0.3480	0.3787	0.6039
	KNN	0.3985	0.3710	0.3842	0.5751
	Mahalanobis	0.3590	0.4789	0.4104	0.5566
	NN3	0.4504	0.4759	0.4628	0.6293
	NN5	0.4522	0.3931	0.4206	0.6323
	NN10	0.4564	0.3875	0.4191	0.6358
	NN15	0.4582	0.3916	0.4223	0.6351
	SVM	0.4609	0.4075	0.4325	0.6407
	Party Coding	0.3159	0.3957	0.3514	0.5230

Πίνακας 2.7: Τα συνολικά αποτελέσματα των ταξινομητών για Ρουμανία, Σουηδία και Σλοβακία

Σύνολο Δεδομένων	Μέθοδος	Recall	Precision	F1	MAP
Ρουμανία	HMM	0.3608	0.5156	0.4245	0.5864
	Naive Bayes	0.4236	0.4669	0.4442	0.6497
	KNN	0.4456	0.4236	0.4343	0.6510
	Mahalanobis	0.3891	0.5005	0.4378	0.6160
	NN3	0.5148	0.4179	0.4613	0.7052
	NN5	0.5160	0.4203	0.4633	0.7020
	NN10	0.4997	0.4799	0.4896	0.6933
	NN15	0.5003	0.4577	0.4781	0.6866
	SVM	0.5185	0.4754	0.4960	0.7103
	Party Coding	0.3237	0.4034	0.3592	0.5630
Σουηδία	HMM	0.4147	0.4040	0.4093	0.6111
	Naive Bayes	0.4286	0.4028	0.4153	0.6209
	KNN	0.3515	0.3656	0.3584	0.5353
	Mahalanobis	0.4083	0.4185	0.4134	0.6025
	NN3	0.3793	0.2367	0.2915	0.5832
	NN5	0.3970	0.3106	0.3485	0.6042
	NN10	0.4083	0.4135	0.4109	0.6080
	NN15	0.4033	0.4005	0.4019	0.5980
	SVM	0.4589	0.4499	0.4544	0.6511
	Party Coding	0.3148	0.3575	0.3348	0.5349
Σλοβακία	HMM	0.2681	0.3798	0.3143	0.4910
	Naive Bayes	0.3795	0.3357	0.3562	0.5741
	KNN	0.3507	0.3435	0.3471	0.5263
	Mahalanobis	0.2985	0.4026	0.3428	0.5319
	NN3	0.4091	0.2792	0.3319	0.5960
	NN5	0.4117	0.3531	0.3801	0.5977
	NN10	0.3986	0.3506	0.3731	0.5827
	NN15	0.3925	0.3492	0.3696	0.5754
	SVM	0.4204	0.3803	0.3993	0.6113
	Party Coding	0.2977	0.3669	0.3287	0.4962

Η καλή απόδοση των νευρωνικών δικτύων και των μηχανών διανυσμάτων υποστήριξης δείχνει ότι τα μοντέλα των χρηστών έχουν ένα βαθμό πολυπλοκότητας, καθώς οι τεχνικές αυτές μπορούν να καταγράψουν πιο σύνθετες σχέσεις μεταξύ των χαρακτηριστικών εισόδου. Επιπρόσθετα με την αύξηση του αριθμού των κρυφών νευρώνων βελτιώνεται η απόδοση της μοντελοποίησης που βασίζεται στο νευρωνικό δίκτυο, κάτι που επιβεβαιώνει την πολυπλοκότητα στα μοντέλα των χρηστών αφού όσο περισσότεροι νευρώνες χρησιμοποιούνται τόσο πιο σύνθετο είναι το πρόβλημα που παρουσιάζεται. Το τι διαφέρει από χώρα σε χώρα είναι ο βαθμός πολυπλοκότητας, αφού κάποιες χρειάζονται περισσότερους και κάποιες λιγότερους νευρώνες για να επιτευχθεί η βέλτιστη απόδοση.

Η μέθοδος Naïve Bayes παρουσίασε επίσης πολύ καλά αποτελέσματα, φτάνοντας στην πρώτη θέση στην κατάταξη των αποτελεσμάτων όσον αφορά το σύνολο δεδομένων της Ολλανδίας και επιτυγχάνοντας εξίσου καλή εμφάνιση με τη μέθοδο SVM στη Λετονία.

Όσον αφορά την Ολλανδία και Λετονία, τα πολύ καλά αποτελέσματα της μεθόδου Naïve Bayes για αυ-

τές τις χώρες αποδεικνύουν πως ο τρόπος που απάντησαν οι συγκεκριμένοι χρήστες το ερωτηματολόγιο του VAA δεν είναι τόσο πολύπλοκος όπως στις άλλες χώρες. Αυτό υποστηρίζεται και από τη συμπεριφορά των νευρωνικών δικτύων στην Ολλανδία που παρουσιάζουν τη μέγιστη τους απόδοση με μόλις 3 κρυφούς νευρώνες τη στιγμή που η Ολλανδία περιλαμβάνει 6 πολιτικά κόμματα. Βέβαια πρόκειται για τα δύο πιο μικρά σύνολα δεδομένων, με τη Δετονία να κατέχει τον πιο μικρό αριθμό πολιτικών κομμάτων, πράγμα που δείχνει πως αυτή η μέθοδος λειτουργεί πολύ καλά σε πολύ μικρά δείγματα.

2.2 Ενίσχυση της σύστασης ψήφου με χρήση των συμπληρωματικών ερωτήσεων

Η πρώτη έκδοση του VAA, που δημιουργήθηκε στην Ολλανδία με το όνομα ‘Stemwíjzer’ το 1989, είχε τη μορφή ενός απλού γραπτού τεστ από 60 δηλώσεις των πολιτικών κομμάτων (κυρίως ιδεολογικές) και μία δισκέτα. Η έκδοση αυτή τοποθετούσε όλα τα πολιτικά κόμματα σε μια αριστερή-δεξιά κλίμακα και ο αναγνώστης έπρεπε να απαντά με ποιες από τις 60 δηλώσεις που εμπεριέχονταν στο τεστ συμφωνούσε. Κάθε δήλωση έπαιρνε από 0 έως 50 μονάδες, ανάλογα με το κόμμα που κρυβόταν πίσω από αυτήν. Διαιρώντας το συνολικό ποσό των πόντων με τον αριθμό των δηλώσεων, με τις οποίες είχε συμφωνήσει ο χρήστης, υπολογιζόταν με ποιο κόμμα ταίριαζε περισσότερο. Σκοπός του συγκεκριμένου εγχειρήματος ήταν να ενημερώσει για τις διαφορές και ομοιότητες μεταξύ των πολιτικών κομμάτων και να βοηθήσει τον κόσμο στο να επιλέξει το πολιτικό κόμμα που του ταιριάζει καλύτερα [48].

Σήμερα τα VAAs είναι διαθέσιμα στο διαδίκτυο και έχουν τύχει μεγάλης ανταπόκρισης από το κοινό [42, 39]. Οι Garzia και Marshall ([51]) υποστηρίζουν ότι υπάρχουν πολλές διαφορές μεταξύ των ποικίλων VAAs σε διάφορες πτυχές. Ένα VAA διαφοροποιείται από κάποιο άλλο βάσει της οργάνωσης που είναι υπεύθυνη για το σχεδιασμό και τους οικονομικούς πόρους, οι οποίοι μπορούν να αντληθούν για τη δημιουργία, την προώθηση και τη συντήρηση του. Επίσης, κάθε VAA χειρίζεται με διαφορετικό τρόπο τον αριθμό των κομμάτων που λαμβάνουν μέρος στις επερχόμενες εκλογές, κάποια χρησιμοποιούν όλα τα κόμματα ενώ άλλα προτιμούν να αναφέρονται μόνο στα πιο δημοφιλή. Επιπρόσθετα ο αριθμός των ερωτήσεων ποικίλει από VAA σε VAA, όπως επίσης δε διαμορφώνονται όλες οι ερωτήσεις με τον ίδιο τρόπο. Επιπλέον η καταγραφή των θέσεων των κομμάτων ή υποψηφίων και οι επιλογές που δίνονται στο χρήστη για να εκφράσει κατά πόσο συμφωνεί ή διαφωνεί με τα ερωτήματα που του τίθενται μέσω του VAA, είναι μια ακόμα παραλλαγή μεταξύ των VAAs. Ακόμη, ο τρόπος με τον οποίο εκτιμάται ο κατάλληλος υποψήφιος για κάθε χρήστη και το πώς εμφανίζεται το τελικό αποτέλεσμα στο χρήστη, μπορεί να διαφέρει από VAA σε VAA.

Το EUvox 2014 είναι μια πανευρωπαϊκή εφαρμογή ηλεκτρονικού συμβούλου ψήφου που δεν αναφέρεται σε μια συγκεκριμένη χώρα αλλά σε ολόκληρη την Ευρωπαϊκή Ένωση. Εκτός από τις 30 ερωτήσεις που θέτει το χρήστη να απαντήσει, υπάρχουν και κάποιες συμπληρωματικές ερωτήσεις, στις οποίες μπορεί να ανταποκριθεί. Το Σχήμα 2.4 απεικονίζει τον τρόπο που εμφανίζονταν κάποιες επιπρόσθετες ερωτήσεις πριν από την έναρξη του EUvox 2014. Οι επιπλέον αυτές ερωτήσεις, βοηθούν το σύστημα να «γνωρίσει» λίγο καλύτερα το χρήστη για να του παρέχει πιο έγκυρες εξατομικευμένες συστάσεις ψήφου.

Στην ενότητα αυτή διερευνάται κατά πόσο βελτιώνεται η απόδοση της πρόβλεψης ψήφου του χρήστη. Δηλαδή πόσο καλύτερα μπορεί να εκτιμήσει το σύστημα τι απάντησαν οι χρήστες στη συμπληρωματική ερώτηση “Στις Ευρωεκλογές, ποιο κόμμα προτίθεστε να ψηφίσετε;” (βλέπε Σχήμα 1.6), όταν μαζί με τις απαντήσεις στις δηλώσεις πολιτικής λαμβάνονται υπόψη και κάποιες από τις απαντήσεις που έδωσαν οι χρήστες στις συμπληρωματικές ερωτήσεις. Η αποτελεσματικότητα της πρόβλεψης ψήφου αναμένεται να διαφέρει από χώρα σε χώρα, δεδομένου ότι τα χαρακτηριστικά των χρηστών, ο αριθμός των

κομμάτων, οι ερωτήσεις που καλούνται χρήστες και κόμματα να απαντήσουν καθώς και το μέγεθος του δείγματος ποικίλουν μεταξύ των χωρών.

Τα αποτελέσματα με τη χρήση των απαντήσεων στις συμπληρωματικές ερωτήσεις και των απαντήσεων στις δηλώσεις πολιτικής, πολύ πιθανόν να είναι καλύτερα από τα αποτελέσματα στην Ενότητα 2.1, όπου χρησιμοποιήθηκαν μόνο οι απαντήσεις των χρηστών στις δηλώσεις πολιτικής, μιας και ενισχύεται η γνώση του κάθε χρήστη από το σύστημα. Για παράδειγμα, οι γυναίκες που υποστηρίζουν ένα κόμμα μπορεί να δίνουν διαφορετικές απαντήσεις από τους άντρες που υποστηρίζουν το ίδιο κόμμα, έτσι σε αυτήν την περίπτωση το σύστημα μπορεί να ανταποκριθεί καλύτερα και να προβεί σε πιο ακριβείς προβλέψεις εάν γνωρίζει το φύλο του χρήστη. Ακόμα και έτσι όμως, δεν προβλέπεται πως όλες οι επιπρόσθετες ερωτήσεις θα συνεισφέρουν θετικά στην ενίσχυση της πρόβλεψης ψήφου και πιθανόν οι ερωτήσεις αυτές να έχουν διαφορετικό αντίκτυπο σε κάθε χώρα.

Τα ερωτήματα που δημιουργούνται για αυτήν την ενότητα είναι τα ακόλουθα:

- Οι συμπληρωματικές ερωτήσεις βοηθούν τις εξατομικευμένες προτάσεις που γίνονται από το VAA;
- Σε ποιο βαθμό βελτιώνονται τα αποτελέσματα της πρόβλεψης ψήφου, ενισχύοντας τις απαντήσεις στις δηλώσεις πολιτικής με τις απαντήσεις στις συμπληρωματικές ερωτήσεις;
- Ποιές συμπληρωματικές ερωτήσεις συμβάλουν ουσιαστικά στην ενίσχυση της γνώσης του χρήστη;
- Ποιές συμπληρωματικές ερωτήσεις λειτουργούν καλύτερα για κάθε χώρα;

2.2.1 Συμπληρωματικές ερωτήσεις

Στην παρούσα Ενότητα προσπαθούμε να μετρήσουμε τη σημαντικότητα των συμπληρωματικών ερωτήσεων για την αποτελεσματικότητα της πρόβλεψης ψήφου. Έτσι εξετάζεται κατά πόσο μπορούν να βοηθήσουν την απόδοση του συστήματος οι συμπληρωματικές ερωτήσεις που δείχνουν τον τρόπο με τον οποίο ψηφίζουν οι χρήστες και το ενδιαφέρον τους για την πολιτική καθώς και αυτές που αφορούν το φύλο, την ηλικία και το μορφωτικό επίπεδο του χρήστη. Επιπρόσθετα μελετά την αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη, μιας και ο τρόπος που τοποθετεί ο χρήστης τον εαυτό του στην Οικονομική κλίμακα, στη Φιλελεύθερη-Συντηρητική κλίμακα και σε αυτήν που αφορά την Ευρωπαϊκή Ένωση, δίνουν επιπλέον γνώση για τις πολιτικές του απόψεις.

The screenshot shows the EUvox 2014 survey interface. At the top left is the EUvox 2014 logo. At the top right are links for 'Σχετικά με εμάς' and 'Συχνές ερωτήσεις'. Below the header, a message reads: 'Πριν από την έναρξη του EUVOX παρακαλούμε να απαντήσετε στις ακόλουθες ερωτήσεις'. The survey consists of six questions, each with a dropdown menu for selection:

- Είμαι κάτοικος της παρακάτω χώρας :
- Σε ποιά εκλογική περιφέρεια ψηφίζετε;
- Φύλο :
- Έτος γέννησης :
- Εκπαίδευση :
- Σε τι βαθμό θα λέγατε πως ενδιαφέρεστε για την πολιτική;

At the bottom right of the form is a right-pointing arrow button.

Σχήμα 2.4: Επιπρόσθετες ερωτήσεις που υπάρχουν στο EUvox 2014

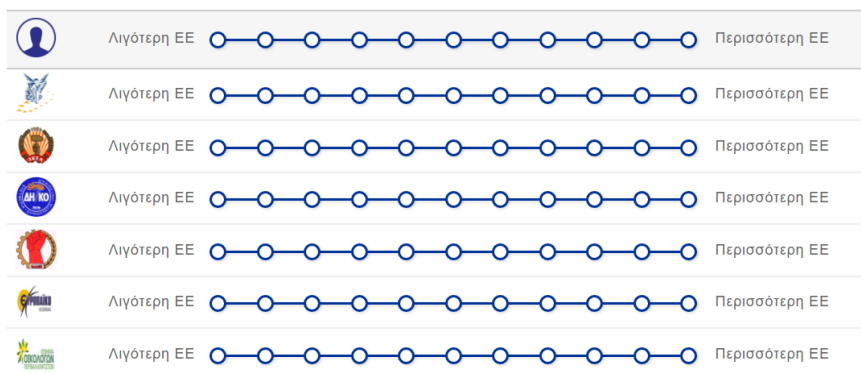
Πίνακας 2.8: Κατανομή με βάση το φύλο σε κάθε σύνολο δεδομένων

A/A	Σύνολο Δεδομένων	Ποσοστό Γυναικών	Ποσοστό Αντρών
1	Αυστρία	0.27	0.73
2	Βουλγαρία	0.29	0.71
3	Κύπρος	0.20	0.80
4	Τσεχία	0.21	0.79
5	Γερμανία	0.16	0.84
6	Δανία	0.35	0.65
7	Εσθονία	0.35	0.65
8	ΗΒ-Αγγλία	0.29	0.71
9	ΗΒ-Βόρεια Ιρλανδία	0.24	0.76
10	ΗΒ-Σκωτία	0.30	0.70
11	ΗΒ-Ουαλία	0.30	0.70
12	Ισπανία	0.26	0.74
13	Φινλανδία	0.28	0.72
14	Γαλλία	0.25	0.75
15	Ελλάδα	0.21	0.79
16	Κροατία	0.29	0.71
17	Ουγγαρία	0.31	0.69
18	Ιρλανδία	0.22	0.78
19	Ιταλία	0.24	0.76
20	Λιθουανία	0.40	0.60
21	Λετονία	0.39	0.61
22	Ολλανδία	0.21	0.79
23	Πολωνία	0.18	0.82
24	Πορτογαλία	0.33	0.67
25	Ρουμανία	0.28	0.72
26	Σουηδία	0.26	0.74
27	Σλοβακία	0.22	0.78

Πίνακας 2.9: Κατανομή των χρηστών των VAAs ανά ηλικιακή ομάδα για κάθε σύνολο δεδομένων

A/A	Σύνολο Δεδομένων	Ποσοστό ηλικίας 18-24	Ποσοστό ηλικίας 25-34	Ποσοστό ηλικίας 35-44	Ποσοστό ηλικίας 45-54	Ποσοστό ηλικίας 55-65	Ποσοστό ηλικίας 65 +
1	Αυστρία	0.30	0.29	0.14	0.13	0.08	0.06
2	Βουλγαρία	0.23	0.26	0.19	0.15	0.13	0.04
3	Κύπρος	0.21	0.32	0.20	0.16	0.08	0.03
4	Τσεχία	0.19	0.29	0.19	0.13	0.10	0.10
5	Γερμανία	0.36	0.27	0.14	0.12	0.07	0.04
6	Δανία	0.19	0.24	0.19	0.18	0.12	0.08
7	Εσθονία	0.16	0.30	0.20	0.16	0.11	0.07
8	ΗΒ-Αγγλία	0.23	0.25	0.15	0.14	0.12	0.11
9	ΗΒ-Βόρεια Ιρλανδία	0.32	0.28	0.15	0.13	0.08	0.04
10	ΗΒ-Σκωτία	0.22	0.30	0.16	0.15	0.10	0.07
11	ΗΒ-Ουαλία	0.24	0.23	0.13	0.14	0.14	0.12
12	Ισπανία	0.29	0.31	0.18	0.12	0.07	0.03
13	Φινλανδία	0.24	0.28	0.16	0.14	0.11	0.07
14	Γαλλία	0.36	0.22	0.14	0.11	0.10	0.07
15	Ελλάδα	0.19	0.27	0.23	0.18	0.10	0.03
16	Κροατία	0.28	0.28	0.18	0.12	0.10	0.04
17	Ουγγαρία	0.32	0.22	0.13	0.09	0.15	0.09
18	Ιρλανδία	0.21	0.30	0.22	0.14	0.08	0.05
19	Ιταλία	0.34	0.28	0.13	0.11	0.10	0.04
20	Λιθουανία	0.41	0.38	0.11	0.05	0.03	0.02
21	Λετονία	0.25	0.36	0.16	0.10	0.07	0.06
22	Ολλανδία	0.30	0.30	0.12	0.12	0.10	0.06
23	Πολωνία	0.43	0.27	0.09	0.07	0.09	0.05
24	Πορτογαλία	0.25	0.30	0.22	0.11	0.08	0.04
25	Ρουμανία	0.20	0.22	0.14	0.20	0.17	0.07
26	Σουηδία	0.38	0.28	0.14	0.11	0.07	0.02
27	Σλοβακία	0.27	0.32	0.18	0.10	0.08	0.05

Παρακαλούμε σημειώστε που θα τοποθετούσατε τον εαυτό σας (η πρώτη γραμμή) και τα πολιτικά κόμματα στην κλίμακα που ακολουθεί. Αν δεν είστε σίγουρος σχετικά με το πού να τοποθετήσετε κάποιο κόμμα, μπορείτε να αφήσετε τη γραμμή κενή. Λιγότερη ΕΕ σημαίνει πως θα θέλατε να αντιστραφεί η πορεία της Ευρωπαϊκής ολοκλήρωσης. Περισσότερη ΕΕ σημαίνει πως επιθυμείτε την εντατικοποίηση της πορείας της Ευρωπαϊκής ολοκλήρωσης.



Σχήμα 2.5: Ο τρόπος που εμφανίζεται η ερώτηση για την αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη στο EUvox 2014 σε σχέση με την κλίμακα “λιγότερη ΕΕ - περισσότερη ΕΕ”

Στο Σχήμα 2.5 φαίνεται ο τρόπος με τον οποίο μπορεί να τοποθετεί ο χρήστης τον εαυτό του και τα πολιτικά κόμματα στον πολιτικό χάρτη. Με το “λιγότερη ΕΕ” εκφράζεται η επιθυμία για να αναστραφεί η πορεία της Ευρωπαϊκής ολοκλήρωσης και με το “περισσότερη ΕΕ” δηλώνεται η θέληση για εντατικοποίηση της πορείας της Ευρωπαϊκής ολοκλήρωσης. Όσον αφορά την οικονομική κλίμακα του πολιτικού χάρτη, με την αυτοτοποθέτηση του χρήστη προς την «Αριστερά» δηλώνεται η προτίμηση του για ευρεία εμπλοκή του κράτους στον έλεγχο της οικονομίας, ενώ προς τη «Δεξιά» εκφράζεται η εμπιστοσύνη του στις αρχές της ελεύθερης αγοράς χωρίς κρατικό παρεμβατισμό. Σχετικά με τη Φιλελεύθερη - Συντηρητική κλίμακα, οι κοινωνικά φιλελεύθεροι χρήστες “αποδέχονται εναλλακτικούς τρόπους ζωής και δίνουν μικρότερη έμφαση στις παραδόσεις και τη διατήρηση της τάξης”, ενώ οι κοινωνικά συντηρητικοί “τονίζουν τη σημασία παραδοσιακών αξιών, όπως ο νόμος και η τάξη, η οικογένεια και ο πατριωτισμός”.

Στους Πίνακες 2.8 - 2.10 βρίσκονται τα δημογραφικά χαρακτηριστικά των χρηστών VAA της κάθε χώρας (φύλο, ηλικία και μορφωτικό επίπεδο αντίστοιχα). Όπως φαίνεται από αυτά, οι άντρες χρήστες είναι κατά πολύ περισσότεροι από τις γυναίκες (βλέπε Πίνακα 2.8) καθώς, κατά μέσο όρο το 70% των χρηστών των VAAs, είναι άνδρες.

Η κατανομή όσον αφορά τις ηλικιακές ομάδες ποικίλει από χώρα σε χώρα (Πίνακας 2.9), ενώ το μορφωτικό επίπεδο της μεγάλης πλειοψηφίας των χρηστών για όλες τις χώρες, με εξαίρεση τη Λετονία, συγκεντρώνεται στις δύο τελευταίες ομάδες που αφορούν τους απόφοιτους λυκείου / τεχνικού λυκείου / επαγγελματικής σχολής και τα άτομα με πτυχίο, μεταπτυχιακό ή και διδακτορικό (Πίνακας 2.10). Στη Λετονία, το μορφωτικό επίπεδο των χρηστών περιορίζεται στους απόφοιτους γυμνασίου και λυκείου / τεχνικού λυκείου / επαγγελματικής σχολής.

Η συντριπτική πλειοψηφία των χρηστών των VAAs (πάνω από το 60%), σχεδόν για όλες τις χώρες, δήλωσαν πως ψηφίζουν ένα πολιτικό κόμμα γιατί η ιδεολογία του πλησιάζει τη δική τους (Πίνακας 2.11). Εξαίρεση πάλι ήταν η Λετονία, όπου το 42% υποστήριξε πως ψηφίζει για ένα κόμμα βάσει της ιδεολογίας του και το 36% επειδή το κόμμα είναι ικανότερο από τα υπόλοιπα. Κάτι αντίστοιχο συμβαίνει και με τη Βουλγαρία με τα ποσοστά της να διαμορφώνονται 49% για την ιδεολογία, 29% για την ικανότητα του κόμματος και 12% για τη βοήθεια που προσφέρει το κόμμα σε ανθρώπους όπως το χρήστη.

Πίνακας 2.10: Κατανομή των χρηστών των VAAs ανά μορφωτικό επίπεδο για κάθε σύνολο δεδομένων

A/A	Σύνολο Δεδομένων	Καθόλου εκπαίδευση	Γυμνάσιο	Λύκειο/Τεχνικό Λύκειο /Επαγγελματική σχολή	Πτυχίο/μεταπτυχιακό /διδακτορικό
1	Αυστρία	0.012	0.030	0.48	0.48
2	Βουλγαρία	0.003	0.018	0.31	0.66
3	Κύπρος	0.0015	0.012	0.22	0.76
4	Τσεχία	0.00091	0.020	0.41	0.57
5	Γερμανία	0.004	0.041	0.55	0.40
6	Δανία	0.0025	0.078	0.53	0.39
7	Εσθονία	0.0023	0.024	0.36	0.61
8	ΗΒ-Αγγλία	0.012	0.093	0.23	0.66
9	ΗΒ-Βόρεια Ιρλανδία	0.0048	0.044	0.18	0.77
10	ΗΒ-Σκωτία	0.0084	0.071	0.20	0.72
11	ΗΒ-Ουαλία	0.015	0.091	0.24	0.66
12	Ισπανία	0.012	0.055	0.36	0.57
13	Φινλανδία	0.0019	0.035	0.42	0.55
14	Γαλλία	0.022	0.110	0.38	0.49
15	Ελλάδα	0.0023	0.018	0.23	0.75
16	Κροατία	0.003	0.350	0.54	0.11
17	Ουγγαρία	0.002	0.070	0.83	0.10
18	Ιρλανδία	0.0045	0.029	0.13	0.84
19	Ιταλία	0.004	0.061	0.46	0.48
20	Λιθουανία	0.0039	0.024	0.26	0.71
21	Λετονία	0.017	0.230	0.69	0.055
22	Ολλανδία	0.0017	0.007	0.23	0.76
23	Πολωνία	0.00063	0.021	0.45	0.53
24	Πορτογαλία	0.0086	0.022	0.20	0.76
25	Ρουμανία	0.00034	0.014	0.72	0.26
26	Σουηδία	0.0055	0.045	0.35	0.60
27	Σλοβακία	0.0023	0.010	0.38	0.61

Πίνακας 2.11: Τα κριτήρια με τα οποία ψηφίζουν οι χρήστες των VAAs σε κάθε χώρα

A/A	Σύνολο Δεδομένων	Ικανότερο κόμμα	Ιδεολογία κόμματος	Βοήθεια που προσφέρει το κόμμα	Οικογένεια και φίλοι	Ηγεσία κόμματος
1	Αυστρία	0.15	0.77	0.032	0.014	0.03
2	Βουλγαρία	0.29	0.49	0.12	0.025	0.081
3	Κύπρος	0.22	0.69	0.024	0.026	0.042
4	Τσεχία	0.18	0.66	0.037	0.024	0.1
5	Γερμανία	0.11	0.75	0.068	0.025	0.045
6	Δανία	0.097	0.83	0.02	0.011	0.044
7	Εσθονία	0.30	0.58	0.027	0.017	0.073
8	ΗΒ-Αγγλία	0.12	0.80	0.038	0.022	0.022
9	ΗΒ-Βόρεια Ιρλανδία	0.15	0.72	0.047	0.045	0.041
10	ΗΒ-Σκωτία	0.15	0.79	0.031	0.017	0.02
11	ΗΒ-Ουαλία	0.097	0.80	0.06	0.027	0.02
12	Ισπανία	0.11	0.77	0.062	0.0079	0.045
13	Φινλανδία	0.077	0.75	0.15	0.013	0.014
14	Γαλλία	0.12	0.78	0.036	0.013	0.043
15	Ελλάδα	0.22	0.70	0.028	0.011	0.04
16	Κροατία	0.12	0.71	0.026	0.028	0.12
17	Ουγγαρία	0.30	0.59	0.04	0.02	0.05
18	Ιρλανδία	0.16	0.69	0.048	0.043	0.052
19	Ιταλία	0.12	0.75	0.021	0.011	0.093
20	Λιθουανία	0.30	0.61	0.013	0.021	0.055
21	Λετονία	0.36	0.42	0.0098	0.038	0.17
22	Ολλανδία	0.11	0.83	0.024	0.01	0.03
23	Πολωνία	0.17	0.70	0.027	0.015	0.097
24	Πορτογαλία	0.19	0.66	0.023	0.02	0.11
25	Ρουμανία	0.20	0.62	0.083	0.026	0.07
26	Σουηδία	0.13	0.80	0.029	0.0069	0.032
27	Σλοβακία	0.23	0.64	0.031	0.03	0.066

Το ποσοστό των χρηστών που εξέφρασαν πως δεν ενδιαφέρονται καθόλου για την πολιτική ήταν κάτω από το 2% για όλες τις χώρες (Πίνακας 2.12). Λίγοι ήταν και οι χρήστες που δήλωσαν ότι ενδιαφέρονται ελάχιστα για την πολιτική (ποσοστό λιγότερο από 20% για όλες τις χώρες με εξαίρεση τη Λετονία με 24%). Όλοι οι υπόλοιποι χρήστες ενδιαφέρονται από πολύ μέχρι αρκετά για την πολιτική.

2.2.2 Η πρόβλεψη της πρόθεσης ψήφου συνδυάζοντας τις δηλώσεις πολιτικής μαζί με κάποιες από τις συμπληρωματικές ερωτήσεις

Στην παρούσα ενότητα εξετάζεται κατά πόσο τα αποτελέσματα της πρόβλεψης ψήφου που είδαμε στην προηγούμενη Ενότητα (Ενότητα 2.1), όπου το πρόβλημα σύστασης προσεγγίστηκε με την κοινωνική και παραδοσιακή μέθοδο χρησιμοποιώντας μόνο τις απαντήσεις των χρηστών στις δηλώσεις πολιτικής του VAA ερωτηματολογίου, μπορούν να βελτιωθούν όταν το σύστημα χρησιμοποιεί και τις απαντήσεις που έδωσε ο χρήστης σε κάποιες επιπρόσθετες ερωτήσεις οι οποίες δεν αφορούν πολιτικές θέσεις. Στους Πίνακες IV.1 - IV.25 βρίσκονται τα αποτελέσματα για κάθε χώρα ξεχωριστά, όταν γίνονται διά-

φοροι συνδυασμοί μεταξύ των απαντήσεων στις δηλώσεις πολιτικής και των απαντήσεων στις συμπληρωματικές ερωτήσεις που αναφέρθηκαν στην Ενότητα 2.4.

Τα παρακάτω σύμβολα αντιστοιχούν στις ακόλουθες συμπληρωματικές ερωτήσεις:

- Q_3 : Τα κριτήρια με τα οποία ψηφίζει ο χρήστης (βλέπε Πίνακα 2.11). Οι πιθανές απαντήσεις που μπορούσε να δώσει ο χρήστης στη συγκεκριμένη συμπληρωματική ερώτηση είναι οι ακόλουθες:
 1. Το κόμμα είναι ικανότερο από τα υπόλοιπα
 2. Η ιδεολογία του κόμματος είναι πιο κοντά στη δική μου
 3. Το κόμμα βοηθάει ανθρώπους σαν εμένα
 4. Η οικογένεια ή οι φίλοι μου υποστηρίζουν το κόμμα αυτό
 5. Είμαι ευχαριστημένος με την ηγεσία του κόμματος
- Q_8 : Το φύλο του χρήστη.
- Q_9 : Η ημερομηνία γέννησης του χρήστη. Ανάλογα με το έτος γέννησης του καθενός, οι χρήστες χωρίστηκαν στις ακόλουθες ηλικιακές κατηγορίες [42] (βλέπε Πίνακα 2.9)
 1. 18-24 χρονών
 2. 25-34 χρονών
 3. 35-44 χρονών
 4. 45-54 χρονών
 5. 55-64 χρονών
 6. 65 και άνω χρονών
- Q_{10} : Το επίπεδο μόρφωσης του χρήστη (βλέπε Πίνακα 2.10).
- Q_{11} : Το πόσο πολύ ενδιαφέρεται για την πολιτική (βλέπε Πίνακα 2.12)
- x -τοποθέτηση: Η αυτοτοποθέτηση του χρήστη στην Οικονομική κλίμακα όπου 0 σημαίνει αριστερά και 10 σημαίνει δεξιά.
- y -τοποθέτηση: Η αυτοτοποθέτηση του χρήστη στην Φιλελεύθερη-Συντηρητική κλίμακα όπου 0 σημαίνει Φιλελεύθερος και 10 σημαίνει Συντηρητικός.
- z -τοποθέτηση: Η αυτοτοποθέτηση του χρήστη στην κλίμακα της Ευρωπαϊκής Ένωσης όπου το 0 σημαίνει λιγότερη ΕΕ και το 10 σημαίνει περισσότερη ΕΕ.

Σκοπός της συγκεκριμένης Ενότητας είναι να εξεταστεί κατά πόσο οι συμπληρωματικές ερωτήσεις ενισχύουν τη γνώση για κάθε χρήστη και οδηγούν σε καλύτερα αποτελέσματα από αυτά που παρουσιάζονται όταν το σύστημα παίρνει ως είσοδο μόνο τις απαντήσεις στις δηλώσεις πολιτικής. Για να συμβεί αυτό, αρχικά εξαιρέθηκαν από το δείγμα όσοι χρήστες δεν απάντησαν στις συμπληρωματικές ερωτήσεις. Στον Πίνακα 2.13 βρίσκεται το ποσοστό των χρηστών, από το δείγμα που χρησιμοποιήθηκε για κάθε χώρα (βλέπε Ενότητα 1.7.3), που απάντησαν στις συμπληρωματικές ερωτήσεις. Στην περίπτωση της Αυστρίας μόνο γύρω στο 57% των χρηστών απάντησαν στις συμπληρωματικές ερωτήσεις, ενώ κατά μέσο όρο το 70% των χρηστών των υπολοίπων χωρών ανταποκρίθηκε σε αυτές.

Πίνακας 2.12: Το ενδιαφέρον των χρηστών των VAAς για την πολιτική σε κάθε χώρα

A/A	Σύνολο Δεδομένων	Καθόλου	Ελάχιστα	Αρκετά	Πολύ
1	Αυστρία	0.0073	0.11	0.37	0.52
2	Βουλγαρία	0.012	0.074	0.43	0.48
3	Κύπρος	0.013	0.10	0.40	0.49
4	Τσεχία	0.0098	0.16	0.5	0.33
5	Γερμανία	0.0058	0.093	0.31	0.59
6	Δανία	0.018	0.19	0.46	0.33
7	Εσθονία	0.013	0.12	0.59	0.28
8	ΗΒ-Αγγλία	0.0087	0.087	0.45	0.45
9	ΗΒ-Βόρεια Ιρλανδία	0.013	0.083	0.41	0.50
10	ΗΒ-Σκωτία	0.0093	0.08	0.43	0.48
11	ΗΒ-Ουαλία	0.011	0.091	0.45	0.45
12	Ισπανία	0.0089	0.056	0.32	0.62
13	Φινλανδία	0.017	0.11	0.50	0.38
14	Γαλλία	0.0062	0.10	0.41	0.48
15	Ελλάδα	0.0089	0.095	0.46	0.43
16	Κροατία	0.0097	0.069	0.40	0.52
17	Ουγγαρία	0.02	0.09	0.45	0.44
18	Ιρλανδία	0.0045	0.049	0.34	0.60
19	Ιταλία	0.0072	0.098	0.48	0.41
20	Λιθουανία	0.0088	0.09	0.54	0.36
21	Λετονία	0.011	0.24	0.57	0.17
22	Ολλανδία	0.0087	0.12	0.47	0.39
23	Πολωνία	0.013	0.09	0.48	0.42
24	Πορτογαλία	0.0072	0.082	0.47	0.44
25	Ρουμανία	0.0088	0.08	0.36	0.55
26	Σουηδία	0.015	0.091	0.40	0.49
27	Σλοβακία	0.0058	0.057	0.47	0.46

Το σύστημα έπαιρνε ως είσοδο μαζί με τις απαντήσεις που έδωσαν οι χρήστες στις δηλώσεις πολιτικής και κάποιες από τις απαντήσεις τους στις συμπληρωματικές ερωτήσεις. Όλοι οι συνδυασμοί που έγιναν μαζί με τα αποτελέσματα που έδωσαν φαίνονται στους Πίνακες IV.1 - IV.25 για κάθε χώρα. Για τη διεκπεραίωση των πειραμάτων χρησιμοποιήθηκε μόνο η μέθοδος SVM και συγκεκριμένα η υλοποίηση *libsvm*², μιας και παρουσίασε την καλύτερη απόδοση σε όλες τις χώρες, εξαιρουμένης της Ολλανδίας, στα αποτελέσματα της Ενότητας 2.1.

Συνοψίζοντας, στην παρούσα ενότητα διερευνούμε κατά πόσο η μέθοδος με την καλύτερη απόδοση όταν χρησιμοποιούνται μόνο οι απαντήσεις στις δηλώσεις πολιτικής, μπορεί να βελτιωθεί ακόμα περισσότερο όταν στο σύστημα εισάγονται και οι απαντήσεις σε κάποιες από τις συμπληρωματικές ερωτήσεις. Αξίζει να σημειωθεί πως στην περίπτωση της Ολλανδίας στην Ενότητα 2.1, τα αποτελέσματα της μεθόδου SVM αν και δεν ήταν τα καλύτερα, είχαν ελάχιστη διαφορά από την πρώτη σε απόδοση μέθοδο.

²<https://en.wikipedia.org/wiki/LIBSVM>

Πίνακας 2.13: Το ποσοστό των χρηστών που απάντησαν σε όλες τις συμπληρωματικές ερωτήσεις για κάθε χώρα

Σύνολο Δεδομένων	Ποσοστό χρηστών που απάντησαν στις συμπληρωματικές ερωτήσεις
Αυστρία	0.5690
Βουλγαρία	0.7230
Κύπρος	0.7074
Τσεχία	0.7182
Γερμανία	0.6674
Δανία	0.7354
Εσθονία	0.7266
ΗΒ-Αγγλία	0.7209
ΗΒ-Βόρεια Ιρλανδία	0.6887
ΗΒ-Σκωτία	0.7245
ΗΒ-Ουαλία	0.6790
Ισπανία	0.6100
Φινλανδία	0.7217
Γαλλία	0.7603
Ελλάδα	0.6544
Κροατία	0.6955
Ουγγαρία	0.7757
Ιρλανδία	0.5915
Ιταλία	0.6857
Λιθουανία	0.8661
Λετονία	0.6983
Ολλανδία	0.7031
Πολωνία	0.7672
Πορτογαλία	0.7569
Ρουμανία	0.7393
Σουηδία	0.7371
Σλοβακία	0.7298

Πίνακας 2.14: Η μέση αύξηση στην απόδοση της πρόβλεψης ψήφου που παρατηρήθηκε για όλες τις χώρες ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις

δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Q3	0.046	0.043	0.045	0.033
Q3Q8	0.048	0.045	0.047	0.034
Q3Q8Q9Q10Q11	0.049	0.046	0.048	0.035
Q3Q9	0.049	0.045	0.047	0.035
Q3Q10	0.05	0.049	0.05	0.036
Q3Q11	0.055	0.055	0.056	0.039
Q8	0.055	0.055	0.056	0.039
Q8Q9	0.058	0.058	0.058	0.04
Q8Q9Q10	0.057	0.057	0.058	0.04
Q8Q10	0.06	0.061	0.061	0.042
Q9	0.058	0.06	0.059	0.041
Q9Q10	0.046	0.043	0.045	0.033
Q10	0.047	0.044	0.046	0.034
Q11	0.048	0.047	0.048	0.035
Q11Q8	0.05	0.05	0.051	0.036
Q11Q9	0.048	0.048	0.048	0.035
Q11Q10	0.048	0.045	0.047	0.034
x	0.04	0.041	0.041	0.029
xy	0.042	0.048	0.046	0.031
xyz	0.046	0.051	0.049	0.033
xz	0.044	0.047	0.046	0.032
y	0.036	0.033	0.035	0.026
yz	0.036	0.037	0.037	0.027
z	0.033	0.036	0.035	0.024

Στον Πίνακα 2.14 βρίσκεται η μέση αύξηση στην απόδοση της πρόβλεψης ψήφου για κάθε χώρα ξεχωριστά, όταν μαζί με τις δηλώσεις πολιτικής το σύστημα έλαβε υπόψη και κάποιες συμπληρωματικές ερωτήσεις. Όπως βλέπουμε οι συνδυασμοί που έγιναν με τις συμπληρωματικές ερωτήσεις, αύξησαν κατά μέσο όρο πέντε μονάδες την απόδοση του VAA (σε σχέση με τα αποτελέσματα της προηγούμενης ενότητας όπου γινόταν χρήση μόνο των δηλώσεων πολιτικής. Ωστόσο, οι συνδυασμοί που περιλάμβαναν το ενδιαφέρον του χρήστη για την πολιτική, τα κριτήρια ψήφου του καθώς και τα δημογραφικά του στοιχεία, κατάφεραν να αυξήσουν κατά μέσο όρο γύρω στις έξι μονάδες την απόδοση της πρόβλεψης ψήφου.

Στους Πίνακες IV.1 - IV.25, με κίτρινο χρώμα εμφανίζεται η απόδοση των SVMs στην προηγούμενη Ενότητα (χρησιμοποιώντας δηλαδή μόνο τις απαντήσεις στις δηλώσεις πολιτικής) και με ροζ χρώμα παρουσιάζεται ο συνδυασμός των συμπληρωματικών ερωτήσεων με τις δηλώσεις πολιτικής που επέφεραν τα καλύτερα αποτελέσματα για κάθε χώρα ξεχωριστά. Ενδεικτικά σε αυτήν την ενότητα παρουσιάζονται τα αποτελέσματα της Φινλανδίας και Γαλλίας (βλέπε Πίνακα 2.15 και Πίνακα 2.16 αντίστοιχα).

Οι συμπληρωματικές ερωτήσεις για την αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη βοήθησαν λιγότερο από τις άλλες ερωτήσεις, στις πλείστες χώρες. Αυτό συμβαίνει γιατί στην αυτοτοποθέτηση στον πολιτικό χάρτη, ο χρήστης επιλέγει μια τιμή από το 0-10, κάτι το οποίο, πολλές φορές, δεν είναι τόσο απλό και ξεκάθαρο. Επίσης ο χρήστης δεν μπορεί πάντα να γνωρίζει με σιγουριά πόσο αριστερά

ή δεξιά να τοποθετήσει τον εαυτό του στον πολιτικό χάρτη με αποτέλεσμα να μπερδεύει το σύστημα, που σε κάποιες περιπτώσεις είχε χειρότερη απόδοση όταν συνδύαζε την αυτοτοποθέτηση του χρήστη με τις δηλώσεις πολιτικής από όταν χρησιμοποιούσε μόνο τις δηλώσεις πολιτικής. Κάτι τέτοιο συνέβη στη Βόρεια Ιρλανδία, την Κροατία και την Ιρλανδία (Πίνακας IV.9, Πίνακας IV.14 και Πίνακας IV.16 αντίστοιχα).

Πίνακας 2.15: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Φινλανδία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Φινλανδία		0.6260	0.6221	0.6241	0.7684
	Q ₃	0.6200	0.5954	0.6074	0.7672
	Q ₃ Q ₈	0.6190	0.5950	0.6068	0.7657
	Q ₃ Q ₈ Q ₉ Q ₁₀ Q ₁₁	0.6200	0.5957	0.6076	0.7659
	Q ₃ Q ₉	0.6219	0.5953	0.6083	0.7682
	Q ₃ Q ₁₀	0.6276	0.6050	0.6161	0.7718
	Q ₃ Q ₁₁	0.6238	0.5983	0.6108	0.7688
	Q ₈	0.6238	0.6000	0.6117	0.7694
	Q ₈ Q ₉	0.6200	0.5901	0.6047	0.7659
	Q ₈ Q ₉ Q ₁₀	0.6219	0.5980	0.6097	0.7690
	Q ₈ Q ₁₀	0.6305	0.6239	0.6272	0.7744
	Q ₉	0.6324	0.6099	0.6210	0.7750
	Q ₉ Q ₁₀	0.6171	0.5919	0.6042	0.7662
	Q ₁₀	0.6219	0.5965	0.6089	0.7687
	Q ₁₁	0.6257	0.5974	0.6112	0.7710
	Q ₁₁ Q ₈	0.6296	0.6204	0.6250	0.7738
	Q ₁₁ Q ₉	0.6286	0.6092	0.6187	0.7718
	Q ₁₁ Q ₁₀	0.6219	0.5953	0.6083	0.7682
	x	0.6607	0.6377	0.6490	0.7881
	xy	0.6670	0.6402	0.6533	0.7905
	xyz	0.6787	0.6512	0.6646	0.8000
	xz	0.6734	0.6490	0.6609	0.7981
	y	0.6479	0.6209	0.6341	0.7778
	yz	0.6554	0.6257	0.6402	0.7838
	z	0.6543	0.6260	0.6398	0.7836

Παρόλα αυτά στη Φινλανδία και Ουγγαρία (Πίνακας 2.15 και Πίνακας IV.15 αντίστοιχα) οι συνδυασμοί με την αυτοτοποθέτηση στον πολιτικό χάρτη απέδωσαν καλύτερα από τις άλλες συμπληρωματικές ερωτήσεις. Όπως επίσης στη Γαλλία (Πίνακας 2.16) ο συνδυασμός μεταξύ της αυτοτοποθέτησης του χρήστη στην Οικονομική κλίμακα με την αυτοτοποθέτηση του στη Φιλελεύθερη-Συντηρητική κλίμακα απέδωσε πολύ καλά. Αυτό θα μπορούσε να αποδεικνύει, ειδικά στην περίπτωση της Φινλανδίας όπου οι συνδυασμοί με την αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη έδωσαν πάρα πολύ καλά αποτελέσματα, πως οι χρήστες των συγκεκριμένων χωρών έχουν πιο κατασταλαγμένες πολιτικές απόψεις και κατάφεραν να τις εκφράσουν με μεγαλύτερη ακρίβεια. Έτσι βοήθησαν την απόδοση του συστήματος, το οποίο συνδύασε την πολιτική αυτοτοποθέτηση των χρηστών με τις δηλώσεις πολιτικής καταλήγοντας σε περισσότερες σωστές προβλέψεις για την πρόθεση ψήφου.

Είναι ξεκάθαρο πως οι συμπληρωματικές ερωτήσεις βελτίωσαν την απόδοση που είχε η μέθοδος SVM

Πίνακας 2.16: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Γαλλία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Γαλλία		0.5704	0.5641	0.5673	0.7433
	Q_3	0.6124	0.6112	0.6118	0.7698
	Q_3Q_8	0.6133	0.6097	0.6115	0.7698
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6150	0.6150	0.6150	0.7723
	Q_3Q_9	0.6184	0.6150	0.6167	0.7722
	Q_3Q_{10}	0.6202	0.6157	0.6179	0.7755
	Q_3Q_{11}	0.6115	0.6096	0.6106	0.7684
	Q_8	0.6107	0.6088	0.6097	0.7695
	Q_8Q_9	0.6115	0.6093	0.6104	0.7694
	$Q_8Q_9Q_{10}$	0.6141	0.6112	0.6127	0.7705
	Q_8Q_{10}	0.6184	0.6146	0.6165	0.7761
	Q_9	0.6133	0.6091	0.6112	0.7712
	Q_9Q_{10}	0.6090	0.6045	0.6067	0.7660
	Q_{10}	0.6141	0.6120	0.6131	0.7710
	Q_{11}	0.6210	0.6190	0.6200	0.7756
	$Q_{11}Q_8$	0.6219	0.6193	0.6206	0.7767
	$Q_{11}Q_9$	0.6158	0.6131	0.6145	0.7719
	$Q_{11}Q_{10}$	0.6184	0.6150	0.6167	0.7722
	x	0.6063	0.6070	0.6067	0.7710
	xy	0.6145	0.6159	0.6152	0.7759
	xyz	0.6063	0.6043	0.6053	0.7722
	xz	0.5920	0.5885	0.5903	0.7638
	y	0.5777	0.5816	0.5797	0.7519
	yz	0.5849	0.5834	0.5841	0.7541
	z	0.5757	0.5746	0.5751	0.7502

με τη χρήση μόνο των δηλώσεων πολιτικής. Αυτό που διαφέρει από χώρα σε χώρα είναι ο συνδυασμός που επιφέρει τη μέγιστη βελτίωση. Επίσης σε κάποιες χώρες η βελτίωση ήταν μεγαλύτερη από κάποιες άλλες και υπήρχαν περιπτώσεις που μερικοί συνδυασμοί κατάφεραν χειρότερα αποτελέσματα από αυτά που παρουσίασε η μέθοδος SVM μόνο με τις δηλώσεις πολιτικής.

Εντούτοις, δεν μπορούν να εξαχθούν ασφαλή συμπεράσματα στην Ενότητα αυτή, μιας και οι χρήστες που χρησιμοποιήθηκαν εδώ είναι λιγότεροι από αυτούς στην Ενότητα 2.1, αφού δεν απάντησαν όλοι οι χρήστες στις συμπληρωματικές ερωτήσεις. Συγκεκριμένα κατά μέσο όρο το 70% των χρηστών όλων των χωρών απάντησαν στις συμπληρωματικές ερωτήσεις (Πίνακας 2.13). Γύρω στο 30% των χρηστών από την Ενότητα 2.1, δηλαδή, αφαιρέθηκαν και αγνοήθηκαν από το δείγμα. Θα ήταν πιο φρόνιμο η σύγκριση να γίνεται με ίσους όρους. Για το λόγο αυτό, στην επόμενη Ενότητα, θα εκτιμηθούν οι απαντήσεις των χρηστών που δεν ανταποκρίθηκαν στις συμπληρωματικές ερωτήσεις και θα διεξαχθεί ξανά η πειραματική διαδικασία της παρούσας ενότητας.

2.3 Εκτίμηση των απαντήσεων στις συμπληρωματικές ερωτήσεις του VAA

Η Ενότητα 2.1 ανέδειξε τη μέθοδο SVM ως την τεχνική που βοηθά περισσότερο στις προβλέψεις ψήφου που γίνονται από το σύστημα, όταν η πληροφορία που λαμβάνεται είναι οι απαντήσεις των χρηστών στις δηλώσεις πολιτικής. Στην Ενότητα 2.2 έγινε διερεύνηση του κατά πόσο τα αποτελέσματα από τη μέθοδο αυτή μπορούν να γίνουν ακόμα καλύτερα, όταν μαζί με τις δηλώσεις πολιτικής λαμβάνονται υπόψη και κάποιες συμπληρωματικές ερωτήσεις. Όμως κάποιοι από τους χρήστες δεν απάντησαν σε όλες τις συμπληρωματικές ερωτήσεις. Έτσι, όσοι δεν απάντησαν έστω και σε μια από τις συμπληρωματικές ερωτήσεις αγνοήθηκαν από το δείγμα και η πειραματική διαδικασία συνεχίστηκε με λιγότερο δείγμα από αυτό που χρησιμοποιήθηκε στην Ενότητα 2.1. Τα αποτελέσματα έδειξαν πως όντως η απόδοση της μεθόδου SVM αυξάνεται όταν το σύστημα υπολογίζει και τις απαντήσεις στις συμπληρωματικές ερωτήσεις, πράγμα που ήταν αναμενόμενο μιας και με αυτόν τον τρόπο ενισχύεται η γνώση για το χρήστη από το σύστημα. Ωστόσο, το να αγνοείται μέρος του δείγματος δεν είναι πάντα η καλύτερη επιλογή μιας και ενδέχεται να επηρεάσει το συνολικό δείγμα και να οδηγήσει σε ένα λανθασμένο αποτέλεσμα. Σε κάθε περίπτωση θα ήταν προτιμότερο το δείγμα να χρησιμοποιείται ολόκληρο. Στην παρούσα Ενότητα γίνεται προσπάθεια υπολογισμού των ελλειπουσών τιμών του δείγματος, για να μπορέσει η πειραματική διαδικασία της Ενότητας 2.2 να επαναληφθεί, έχοντας αυτήν τη φορά όλο το δείγμα των χρηστών διαθέσιμο.

2.3.1 Τεχνικές για τον υπολογισμό των ελλειπουσών τιμών

Το πρόβλημα σύστασης στα ΣΣ μπορεί να διατυπωθεί ως εξής: Έστω ότι υπάρχει ένα σύνολο χρηστών $C_u = \{cu_1, cu_2, \dots, cu_N\}$ και ένα σύνολο αντικειμένων ή υπηρεσιών $I_t = \{it_1, it_2, \dots, it_M\}$ που οι χρήστες μπορούν να αξιολογούν, όπως βιβλία, ταινίες, εστιατόρια. Η u_t είναι μια συνάρτηση χρησιμότητας που μέτρα τη χρησιμότητα του αντικειμένου it στο χρήστη cu , δηλαδή, $u_t: C_u \times I_t \rightarrow R$. Κάθε κελί του πίνακα U_t αντιστοιχεί στο βαθμό που έδωσε ο χρήστης cu για το αντικείμενο it για να δείξει κατά πόσο τον άφησε ευχαριστημένο. Ένα κενό κελί στον πίνακα αντιστοιχεί σε ένα συγκεκριμένο αντικείμενο που δεν έχει αξιολογηθεί από ένα συγκεκριμένο χρήστη. Το ΣΣ σε αυτήν την περίπτωση καλείται να υπολογίσει την τιμή που θα έδινε ο χρήστης στο αντικείμενο, συμπληρώνοντας το άδειο κελί. Για να συμβεί αυτό, το σύστημα μπορεί να εκμεταλλευτεί τις αξιολογήσεις που είναι ήδη καταχωρημένες στη βάση δεδομένων του, για να υπολογίσει τις πιθανές τιμές που θα έδιναν οι χρήστες σε αντικείμενα, τα οποία δεν έχουν βαθμολογήσει ακόμη και αντιστοιχούν στα άδεια κελιά του πίνακα U_t . Έτσι ο πίνακας γεμίζει με πιθανές αξιολογήσεις και συστήνονται στους χρήστες τα αντικείμενα με την υψηλότερη πιθανή ανατροφοδότηση.

Για την εκτίμηση των άδειων κελιών του πίνακα U_t ή αλλιώς των ελλειπουσών τιμών, μπορούν να χρησιμοποιηθούν τεχνικές παραγοντοποίησης πινάκων όπως είναι τα εναλλασσόμενα ελάχιστα τετράγωνα (Alternating Least Squares³ - ALS), η στοχαστική μέθοδος καθοδικής κλίσης (Stochastic Gradient Descent⁴ - SGD) και η ανάλυση σε ιδιάζουσες τιμές (Singular Value Decomposition⁵ - SVD) καθώς και άλλες προσεγγίσεις του συνεργατικού φίλτραρίσματος, βρίσκοντας ομοιότητες μεταξύ των χρηστών (Agathokleous & Tsaratsoulis [86]). Υπό αυτές τις συνθήκες, το σύστημα ψάχνει να βρει τους χρήστες που φαίνεται να προτιμούν τα ίδια προϊόντα. Με αυτό τον τρόπο εάν κάποιος από αυτούς δεν έχει βαθμολογήσει ένα προϊόν, τότε το σύστημα μπορεί να χρησιμοποιήσει τις αξιολογήσεις των άλλων χρηστών για να εκτιμήσει

³<https://datasciencemadesimpler.wordpress.com/tag/alternating-least-squares/>

⁴https://en.wikipedia.org/wiki/Stochastic_gradient_descent

⁵https://en.wikipedia.org/wiki/Singular-value_decomposition

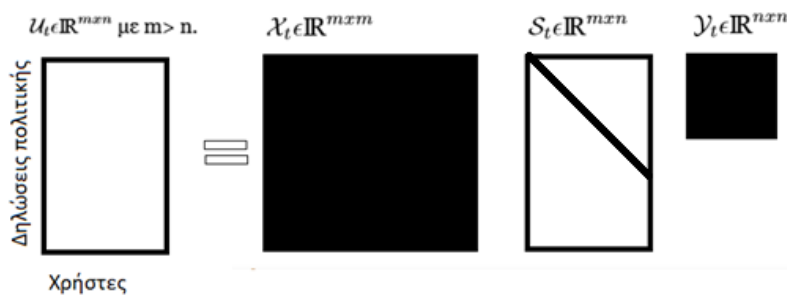
την πιθανή βαθμολογία που θα έδινε ο συγκεκριμένος χρήστης σε αυτό το προϊόν.

Όταν ο πίνακας U_t είναι μεγάλος και αραιός, πράγμα που σημαίνει πως πολλά από τα αντικείμενα στη βάση δεδομένων δεν έχουν αξιολογηθεί από τους χρήστες και υπάρχει ένα μεγάλο ποσοστό ελλιπή στοιχείων, τότε το σύστημα αντιμετωπίζει δυσκολία στην εύρεση σχέσεων μεταξύ των χρηστών. Κάτι τέτοιο επηρεάζει σε μεγάλο βαθμό την αποτελεσματικότητα της διαδικασίας σύστασης, καθόσον ένας μεγάλος πίνακας αυξάνει το χρόνο που απαιτείται για την επεξεργασία των δεδομένων και τη δημιουργία της σύστασης. Επιπρόσθετα κάποιες τεχνικές παραγοντοποίησης πινάκων δεν είναι αποτελεσματικές σε αραιούς πίνακες, όπως συμβαίνει για παράδειγμα με την SVD [87]. Σε αυτές τις περιπτώσεις προτιμάται η χρήση εναλλακτικών τεχνικών παραγοντοποίησης πινάκων, όπως είναι οι ALS (Salakhutdinov & Mnih [88]) και SGD (Zhou et al. [89]). Η αποτελεσματικότητα της κάθε προσέγγισης που εφαρμόζεται για την εκτίμηση των ελλειπουσών τιμών και η μέθοδος που θα προτιμηθεί στο πρόβλημα σύστασης εξαρτάται από το πόσο μεγάλο είναι το μέγεθος του πίνακα U_t και πόσο αραιά είναι τα δεδομένα σε αυτόν. Στην περίπτωση που δεν υπάρχουν πολλές ελλειπούσες τιμές, όπως συμβαίνει στην έρευνα των Agathokleous και Tsapatsoulis [86] η τεχνική SVD έχει αποδειχθεί να υπερτερεί των κλασικών μεθόδων συνεργατικού φιλτραρίσματος και άλλων τεχνικών παραγοντοποίησης.

Η τεχνική SVD ή αλλιώς η ανάλυση σε ιδιάζουσες τιμές, είναι μια γνωστή μέθοδος παραγοντοποίησης πινάκων η οποία παραγοντοποιεί τον πίνακα U_t σε τρεις πίνακες: $U_t = X_t \cdot S_t \cdot Y_t'$, όπου

- $U_t \in \mathbb{R}^{m \times n}$ με $m > n$.
- $X_t \in \mathbb{R}^{m \times m}$. Οι m στήλες του X_t ονομάζονται αριστερά ιδιάζοντα διανύσματα και ικανοποιούν την εξής συνθήκη: $X_t' \cdot X_t = X_t \cdot X_t' = I$.
- $Y_t \in \mathbb{R}^{n \times n}$. Οι n στήλες του Y_t ονομάζονται αριστερά ιδιάζοντα διανύσματα και ικανοποιούν τη συνθήκη: $Y_t' \cdot Y_t = Y_t \cdot Y_t' = I$.
- $S_t \in \mathbb{R}^{m \times n}$. Οι διαγώνιες τιμές του S_t είναι μη αρνητικές ιδιάζουσες τιμές σε φθίνουσα σειρά έτσι ώστε $\sigma_1 \geq \sigma_2 \geq \dots \sigma_{\min(m, n)}$.

Με τη μέθοδο SVD γίνεται αναπαράσταση του αρχικού δείγματος χρησιμοποιώντας ένα μικρότερο αριθμό μεταβλητών. Έτσι δημιουργείται ένας χώρος μικρότερων διαστάσεων, ο οποίος αντιπροσωπεύει το αρχικό δείγμα που τώρα εμφανίζεται με λιγότερα ελλιπή δεδομένα και ως εκ τούτου είναι πιο εύκολη η εύρεση των χρηστών με παρόμοιες προτιμήσεις. Οι νέοι πίνακες που δημιουργούνται (X_t, Y_t, S_t), όταν πολλαπλασιάζονται μεταξύ τους δίνουν πίσω τον αρχικό πίνακα (U_t).



Σχήμα 2.6: Αναπαράσταση της μεθόδου SVD

2.3.2 Η εκτίμηση των ελλειπουσών τιμών των χρηστών

Στην προηγούμενη Ενότητα παρατηρήθηκε πως οι συμπληρωματικές ερωτήσεις ενισχύουν τη γνώση του συστήματος για το χρήστη και συντελούν στην αύξηση της απόδοσης της πρόβλεψης ψήφου. Το συμπέρασμα αυτό διεξήχθη από τη σύγκριση μεταξύ των αποτελεσμάτων στην Ενότητα 2.1 (όπου το σύστημα χρησιμοποίησε μόνο τις απαντήσεις των χρηστών στις δηλώσεις πολιτικής) και των αποτελεσμάτων στην Ενότητα 2.2 (όπου οι απαντήσεις στις δηλώσεις πολιτικής συνδυάστηκαν με τις απαντήσεις σε συμπληρωματικές ερωτήσεις). Τα δεδομένα, όμως, της Ενότητας 2.2 ήταν μειωμένα, αφού αγνοήθηκαν όσοι χρήστες δεν απάντησαν σε όλες ανεξαιρέτως τις συμπληρωματικές ερωτήσεις. Ωστόσο, κάτι τέτοιο μπορεί να αλλοιώσει σημαντικά το δείγμα και να επηρεάσει τα αποτελέσματα. Έτσι, αυτή τη φορά επαναλαμβάνεται το πείραμα της Ενότητας 2.2, εφαρμόζοντας τεχνικές από την έρευνα των Agathokleous και Tsapatsoulis (2013) για να εκτιμηθούν και όχι να αγνοηθούν οι ελλείπουσες τιμές του δείγματος. Στη συνέχεια οι εκτιμημένες απαντήσεις χρησιμοποιούνται για να συμπληρώσουν τις απαντήσεις των χρηστών που δεν υπάρχουν και ακολουθεί η ίδια πειραματική διαδικασία, έχοντας τώρα όλο το δείγμα διαθέσιμο.

Αρχικά στην παρούσα ενότητα διερευνάται κατά πόσο το σύστημα είναι ικανό να προβλέψει τις απαντήσεις του χρήστη σε κάποιες από τις συμπληρωματικές ερωτήσεις, λαμβάνοντας υπόψη τις απαντήσεις του στις δηλώσεις πολιτικής. Η αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη, η οποία όπως φαίνεται από την Ενότητα 2.2 βοηθά αρκετά στην πρόβλεψη της πρόθεσης ψήφου, είναι δύσκολο να εκτιμηθεί αφού είναι κάτι υποκειμενικό, έχει πολλές πιθανές απαντήσεις και ο χρήστης δεν ξέρει πάντα με ακρίβεια πόσο δεξιά ή αριστερά να τοποθετήσει τον εαυτό του στον πολιτικό χάρτη. Έτσι θεωρήθηκε προτιμότερο όπως εκτιμηθούν οι απαντήσεις στις συμπληρωματικές ερωτήσεις για τα δημογραφικά στοιχεία του χρήστη, το ενδιαφέρον του για την πολιτική και τα κριτήρια με τα οποία επιλέγει τον υποψήφιο που πρόκειται να υποστηρίξει στις εκλογές.

Η μέθοδος που είναι καλύτερο να ακολουθηθεί για την εκτίμηση των ελλιπή στοιχείων εξαρτάται από το πόσο μεγάλο είναι το μέγεθος του δείγματος και πόσο αραιά είναι τα δεδομένα σε αυτό. Στην προκειμένη περίπτωση προτιμήθηκε η χρήση της τεχνικής SVD, αφού όπως φαίνεται από τον Πίνακα 2.17 οι ελλείπουσες τιμές (αυτές που βρίσκονται κάτω από τη στήλη sparsity) δεν ξεπερνούν το 4% του δείγματος και οι Agathokleous και Tsapatsoulis (2013) έχουν αποδείξει πως η τεχνική SVD υπερτερεί των κλασικών μεθόδων συνεργατικού φιλτραρίσματος και άλλων τεχνικών παραγοντοποίησης, όταν οι ελλείπουσες τιμές του δείγματος είναι ελάχιστες.

Με τη μέθοδο SVD δημιουργούνται τρεις καινούργιοι πίνακες από τον πίνακα που περιέχει το αρχικό δείγμα, οι οποίοι όταν πολλαπλασιάζονται μεταξύ τους δίνουν πίσω ένα καινούργιο πίνακα που είναι όμοιος με τον αρχικό, αντικαθιστώντας, όμως, όλα τα άδεια κελιά με εκτιμημένες τιμές. Για τον υπολογισμό της απόδοσης της συγκεκριμένης τεχνικής, κατά πόσο δηλαδή ο καινούργιος πίνακας μοιάζει πραγματικά με τον αρχικό, χρησιμοποιήθηκε μια παραλλαγή του μέσου απόλυτου σφάλματος MAE (mean absolute error), που υπολογίζεται με τη βοήθεια της νόρμας Frobenius. Εάν \hat{U}_t είναι η εκτίμηση του αρχικού πίνακα U_t , τότε η ακρίβεια της μεθόδου υπολογίζεται ως ακολούθως:

$$A = \frac{\|U_t - \hat{U}_t\|}{\|U_t\| + \|\hat{U}_t\|}, \quad (2.3)$$

όπου $\|U\|$ είναι η νόρμα Frobenius του πίνακα U .

Στον Πίνακα 2.18 βρίσκονται τα αποτελέσματα της εκτίμησης των ελλειπουσών τιμών με τη χρήση της τεχνικής SVD. Το MAE δείχνει το σφάλμα που υπάρχει μεταξύ του πίνακα που αναπαριστά το αρχικό

δείγμα και του πίνακα που δημιουργήθηκε χρησιμοποιώντας την τεχνική SVD. Όπως ήταν αναμενόμενο τα αποτελέσματα είναι πάρα πολύ καλά και το σφάλμα μεταξύ του αρχικού και εκτιμημένου πίνακα δεν ξεπερνά το 5% για καμία χώρα.

Αφού υπολογίστηκαν όλες οι ελλείπουσες τιμές στις συμπληρωματικές ερωτήσεις και δημιουργήθηκε ο πίνακας \hat{U}_t που δεν έχει ελλιπή στοιχεία και δε διαφέρει περισσότερο από 5% από τον αρχικό πίνακα του δείγματος, η επόμενη φάση της πειραματικής διαδικασίας είναι η εκτίμηση της πρόβλεψης ψήφου του χρήστη, χρησιμοποιώντας τα δεδομένα του πίνακα \hat{U}_t . Με αυτό τον τρόπο γίνεται εφικτή η μελέτη της απόδοσης της πρόβλεψης ψήφου όταν λαμβάνονται υπόψη οι απαντήσεις στις συμπληρωματικές ερωτήσεις, κάτι που έγινε και στην Ενότητα 2.2, έχοντας όμως τώρα όλο το δείγμα των χρηστών για σύγκριση.

Για την εκτίμηση της πρόβλεψης ψήφου εφαρμόστηκε η μέθοδος SVM, παίρνοντας ως είσοδο τις απαντήσεις των χρηστών στις δηλώσεις πολιτικής μαζί με τις απαντήσεις τους στις συμπληρωματικές ερωτήσεις για το φύλο, την ηλικία, το μορφωτικό επίπεδο, το ενδιαφέρον για την πολιτική και τα κριτήρια ψήφου, όπως βρίσκονται στον πίνακα \hat{U}_t .

Πίνακας 2.17: Το ποσοστό των ελλειπουσών τιμών στο δείγμα

Σύνολο Δεδομένων	Sparsity
Αυστρία	0.0351
Βουλγαρία	0.0284
Κύπρος	0.0246
Τσεχία	0.0185
Γερμανία	0.0211
Δανία	0.0239
Εσθονία	0.0131
ΗΒ-Αγγλία	0.0154
ΗΒ-Βόρεια Ιρλανδία	0.0251
ΗΒ-Σκωτία	0.0177
ΗΒ-Ουαλία	0.0200
Ισπανία	0.0314
Φινλανδία	0.0159
Γαλλία	0.0139
Ελλάδα	0.0264
Κροατία	0.0209
Ουγγαρία	0.0112
Ιρλανδία	0.0315
Ιταλία	0.0220
Λιθουανία	0.0144
Λετονία	0.0152
Ολλανδία	0.0255
Πολωνία	0.0122
Πορτογαλία	0.0147
Ρουμανία	0.0134
Σουηδία	0.0185
Σλοβακία	0.0167

Ο Πίνακας 2.19 στη δεύτερη στήλη αναγράφει τα αποτελέσματα του F-measure που καταγράφηκαν

στην Ενότητα 2.1 με τη μέθοδο SVM, όταν δηλαδή ως είσοδο χρησιμοποιήθηκαν μόνο οι απαντήσεις του κάθε χρήστη στις δηλώσεις πολιτικής. Στην τρίτη στήλη είναι η απόδοση του F-measure από την Ενότητα 2.2 όταν το σύστημα πήρε ως είσοδο τις απαντήσεις του χρήστη στις δηλώσεις πολιτικής μαζί με τις απαντήσεις στις συμπληρωματικές ερωτήσεις για το φύλο, την ηλικία, το μορφωτικό επίπεδο, το ενδιαφέρον για την πολιτική και τα κριτήρια ψήφου. Στην τελευταία στήλη αναγράφονται τα αποτελέσματα του F-measure από την πειραματική διαδικασία αυτής της ενότητας. Η δεύτερη και τελευταία στήλη απευθύνεται στο ίδιο ακριβώς δείγμα, ενώ η τρίτη στήλη σε ένα ποσοστό αυτού του δείγματος.

Πίνακας 2.18: Η απόδοση της SVD για την εκτίμηση των ελλειπουσών τιμών των χρηστών

Σύνολο Δεδομένων	MAE
Αυστρία	0.0379
Βουλγαρία	0.0413
Κύπρος	0.0364
Τσεχία	0.0389
Γερμανία	0.0389
Δανία	0.0407
Εσθονία	0.0396
ΗΒ-Αγγλία	0.0363
ΗΒ-Βόρεια Ιρλανδία	0.0425
ΗΒ-Σκωτία	0.0366
ΗΒ-Ουαλία	0.0422
Ισπανία	0.0365
Φινλανδία	0.0392
Γαλλία	0.0389
Ελλάδα	0.0358
Κροατία	0.0408
Ουγγαρία	0.0388
Ιρλανδία	0.0374
Ιταλία	0.0373
Λιθουανία	0.0368
Λετονία	0.0401
Ολλανδία	0.0369
Πολωνία	0.0378
Πορτογαλία	0.0356
Ρουμανία	0.0428
Σουηδία	0.0425
Σλοβακία	0.0415

Τα αποτελέσματα στην Ενότητα 2.1 συμφωνούν με τα αποτελέσματα αυτής της Ενότητας (βλέπε τις δύο τελευταίες στήλες του Πίνακα 2.19). Οι δύο αυτές Ενότητες αναφέρονται στο ίδιο θέμα, με τη διαφορά πως η παρούσα Ενότητα χρησιμοποιεί ολόκληρο το δείγμα ενώ η προηγούμενη Ενότητα εφαρμόζει μέρος του δείγματος αυτού (απομονώνει το δείγμα που δεν έχει ελλιπή στοιχεία). Και στις δύο περιπτώσεις η απόδοση της πρόβλεψης ψήφου είναι καλύτερη από αυτής της Ενότητας 2.1 που δεν έλαβε υπόψη τις απαντήσεις των χρηστών στις συμπληρωματικές ερωτήσεις. Οπότε οι συγκεκριμένες παρατηρήσεις οδηγούν στο συμπέρασμα πως όντως οι συμπληρωματικές ερωτήσεις βοηθούν στην πρόβλεψη της πρόθεσης ψήφου και θα ήταν καλό να λαμβάνονται υπόψη. Επίσης, στην περίπτωση που το δείγμα χρησιμο-

Πίνακας 2.19: Σύγκριση μεταξύ των αποτελεσμάτων στην Ενότητα 2.1 και 2.2

Χώρα	F1 μόνο με δηλώσεις πολιτικής (από τους Πίνακες 2.1 - 2.7 στην Ενότητα 2.1)	F1 με δηλώσεις πολιτικής + συμπληρωματικές ερωτήσεις (από τους Πίνακες IV.1 - IV.25 στην Ενότητα 2.2)	F1 με δηλώσεις πολιτικής + πραγματικές και εκτιμώμενες τιμές
Αυστρία	0.5012	0.5656	0.5321
Βουλγαρία	0.4735	0.5789	0.5829
Κύπρος	0.6399	0.6590	0.6598
Τσεχία	0.5102	0.5777	0.5626
Γερμανία	0.5917	0.5956	0.5873
Δανία	0.5540	0.5806	0.5926
Εσθονία	0.4401	0.4590	0.4762
ΗΒ-Αγγλία	0.5821	0.6646	0.6619
ΗΒ-Βόρεια Ιρλανδία	0.5129	0.5481	0.5171
ΗΒ-Σκωτία	0.5743	0.6330	0.6513
ΗΒ-Ουαλία	0.5008	0.6218	0.5953
Ισπανία	0.3739	0.4229	0.4290
Φινλανδία	0.6241	0.6076	0.6098
Γαλλία	0.5673	0.6150	0.6047
Ελλάδα	0.5508	0.5897	0.5768
Κροατία	0.5714	0.6242	0.6231
Ουγγαρία	0.6060	0.5875	0.6266
Ιρλανδία	0.3832	0.4699	0.4417
Ιταλία	0.5998	0.6536	0.6339
Λιθουανία	0.6046	0.6467	0.6662
Λετονία	0.7223	0.7429	0.7499
Ολλανδία	0.6832	0.6763	0.6916
Πολωνία	0.6260	0.6809	0.6728
Πορτογαλία	0.4325	0.5321	0.5351
Ρουμανία	0.4960	0.5657	0.5447
Σουηδία	0.4544	0.4904	0.4841
Σλοβακία	0.3993	0.4210	0.4614

ποιείται ολόκληρο αλλά και όταν λαμβάνεται υπόψη μέρος από αυτό, τα αποτελέσματα είναι παρόμοια. Οι μικρές διαφορές που υπάρχουν και δεν ξεπερνούν το 4%, οφείλονται στην απόκλιση μεταξύ των εκτιμώμενων τιμών και των πραγματικών που θα έδιναν οι χρήστες εάν απαντούσαν στις συγκεκριμένες συμπληρωματικές ερωτήσεις καθώς και επίσης στο γεγονός πως το δείγμα διαφέρει σε μέγεθος.

Κεφάλαιο 3

Πολιτικές - Κοινωνικές επιπτώσεις των VAAs

3.1 Ιδεολογική ταύτιση κόμματος - χρήστη

Τον τελευταίο καιρό, έχει παρατηρηθεί σε μεγάλο βαθμό η χαμηλή προσέλευση των ψηφοφόρων στις κάλπες, που ίσως να οφείλεται εν μέρει στο χρόνο και τον κόπο που σπαταλά ο πολίτης για να εισέλθει σε αυτές. Όμως, στην πραγματικότητα αυτό που κοστίζει περισσότερο χρόνο και κόπο σε ένα ψηφοφόρο είναι η διαδικασία που περνά για να ενημερωθεί σωστά σχετικά με τα κόμματα και τους υποψηφίους, έτσι ώστε να καταλήξει σε αυτόν που θα κερδίσει την ψήφο εμπιστοσύνης του [37]. Η συγκεκριμένη διαδικασία γίνεται ακόμα δυσκολότερη όταν η “υπερπροσφορά” πληροφοριών μπερδεύει αντί να βοηθά τον πολίτη, ο οποίος δεν ξέρει πώς να χειριστεί σωστά την ενημέρωση που λαμβάνει και να εστιάσει την προσοχή του στις σωστές και ακριβείς πληροφορίες [38, 37, 39].

Υπάρχουν πολλοί τρόποι για να ενημερωθεί ο κόσμος σχετικά με τις θέσεις των πολιτικών κομμάτων που λαμβάνουν μέρος σε επερχόμενες εκλογές. Τα μέσα μαζικής ενημέρωσης συνήθως παρέχουν αυτές τις πληροφορίες, καθώς επίσης ο πολίτης μπορεί να μάθει για τις οποιεσδήποτε πολιτικές εξελίξεις από τις καθημερινές συζητήσεις με τους συνανθρώπους του. Επίσης τα ίδια τα πολιτικά κόμματα κρατούν ενημερους τους πολίτες για τα έργα και τους στόχους τους μέσα από πολιτικές αφίσες, φυλλάδια, τηλεοπτικές διαφημίσεις, συνεντεύξεις, συζητήσεις σε τηλεοπτικά πάνελ πολιτικών εκπομπών, τηλεφωνική επικοινωνία με ψηφοφόρους.

Παρόλα αυτά, κάθε πηγή ενημέρωσης προβάλλει με αλλιώτικο τρόπο τα διάφορα πολιτικά κόμματα, κάτι που μπερδεύει πολλές φορές τον πολίτη, ο οποίος συχνά δεν είναι σωστά ενημερωμένος ακόμα και για τις πολιτικές θέσεις των κομμάτων που υποστηρίζει. Αυτό έχει ως αποτέλεσμα οι απαντήσεις που δίνει ένας χρήστης στο online ερωτηματολόγιο του VAA να μην συμπίπτουν με αυτές που συμπληρώνονται εκ μέρους των πολιτικών κομμάτων, πράγμα που μειώνει την απόδοση της παραδοσιακής προσέγγισης του VAA (βλέπε Ενότητα 2.1.3).

Η σωστή γνώση της ιδεολογίας των κομμάτων εκ μέρους των πολιτών είναι απαραίτητη, μιας και είναι το κριτήριο με το οποίο επιλέγουν οι περισσότεροι το κόμμα που πρόκειται να υποστηρίξουν στις επερχόμενες εκλογές. Αυτό φάνηκε μέσα από την Ενότητα 2.2 και τον Πίνακα 2.11, όπου κατά μέσο όρο το 70% των χρηστών δήλωσε πως ψηφίζει ένα κόμμα εξαιτίας της ιδεολογίας του και σχεδόν όλο το υπόλοιπο ποσοστό επειδή το κόμμα είναι ικανότερο από τα υπόλοιπα. Η παρούσα ενότητα εξετάζει την ιδεολογική ταύτιση μεταξύ των κομμάτων και των χρηστών, για να καταλήξει εάν όντως οι χρήστες

γνωρίζουν τις θέσεις των κομμάτων που προτίθενται να ψηφίσουν.

Η ιδεολογική ταύτιση μπορεί να μετρηθεί από την απόδοση της παραδοσιακής προσέγγισης του VAA, η οποία στην ουσία υπολογίζει κατά πόσο οι χρήστες έδωσαν όμοιες απαντήσεις με το πολιτικό κόμμα που υποστηρίζουν. Έτσι, εάν η ομοιότητα είναι μεγάλη τότε ενδέχεται να υπάρχει και ιδεολογική ταύτιση μεταξύ κόμματος και χρήστη. Στην αντίθετη περίπτωση, πολύ πιθανόν ο χρήστης να υποστηρίζει ένα κόμμα χωρίς να συμφωνεί με τις πολιτικές του θέσεις (ή χωρίς να γνωρίζει καν για αυτές) αλλά επειδή επηρεάζεται από την κοινότητα στην οποία ζει (αυτή είναι και η λογική του SVAA).

Στην Ενότητα 2.1 εξετάστηκε η παραδοσιακή προσέγγιση του VAA καθώς και η λογική του SVAA. Αν και η απόδοση της παραδοσιακής προσέγγισης δεν ήταν ίδια σε όλες τις χώρες, η λογική του SVAA φάνηκε να αποδίδει καλύτερα σε όλες τις περιπτώσεις. Επιπρόσθετα τα αποτελέσματα της συγκεκριμένης Ενότητας έδωσαν ενδείξεις για πολυμορφία στο δείγμα των κομμάτων, δηλαδή οι χρήστες που υποστήριξαν το ίδιο κόμμα δεν απάντησαν με παρόμοιο τρόπο στο ερωτηματολόγιο του VAA κάτι που πιθανό να εξυπονοεί πως δεν έχουν όλοι οι χρήστες το ίδιο πολιτικό υπόβαθρο. Η πολυμορφία στο δείγμα είναι και μια αιτία που επηρεάζει την απόδοση της παραδοσιακής προσέγγισης του VAA και δεν αφήνει να φανεί ξεκάθαρα εάν όντως υπάρχει και σε ποιο βαθμό ιδεολογική ταύτιση μεταξύ του κόμματος και των χρηστών που το υποστηρίζουν.

Στην παρούσα Ενότητα η ιδεολογική ταύτιση μεταξύ κόμματος και χρήστη διερευνάται με διαφορετικό τρόπο από τον παραδοσιακό. Όλο το δείγμα χωρίζεται σε ομάδες, ανάλογα με την ομοιότητα στις απαντήσεις τους στις δηλώσεις πολιτικής. Στη συνέχεια εξετάζεται κατά πόσο χρήστες που δήλωσαν πως θα υποστηρίξουν το ίδιο πολιτικό κόμμα, βρίσκονται στην ίδια ομάδα. Ταυτόχρονα μελετάται σε πόσες καινούργιες ομάδες χωρίζονται οι χρήστες του κάθε κόμματος, πράγμα που απαντά ξεκάθαρα στο εάν υπάρχει όντως πολυμορφία στο δείγμα των χρηστών που υποστηρίζουν το ίδιο κόμμα.

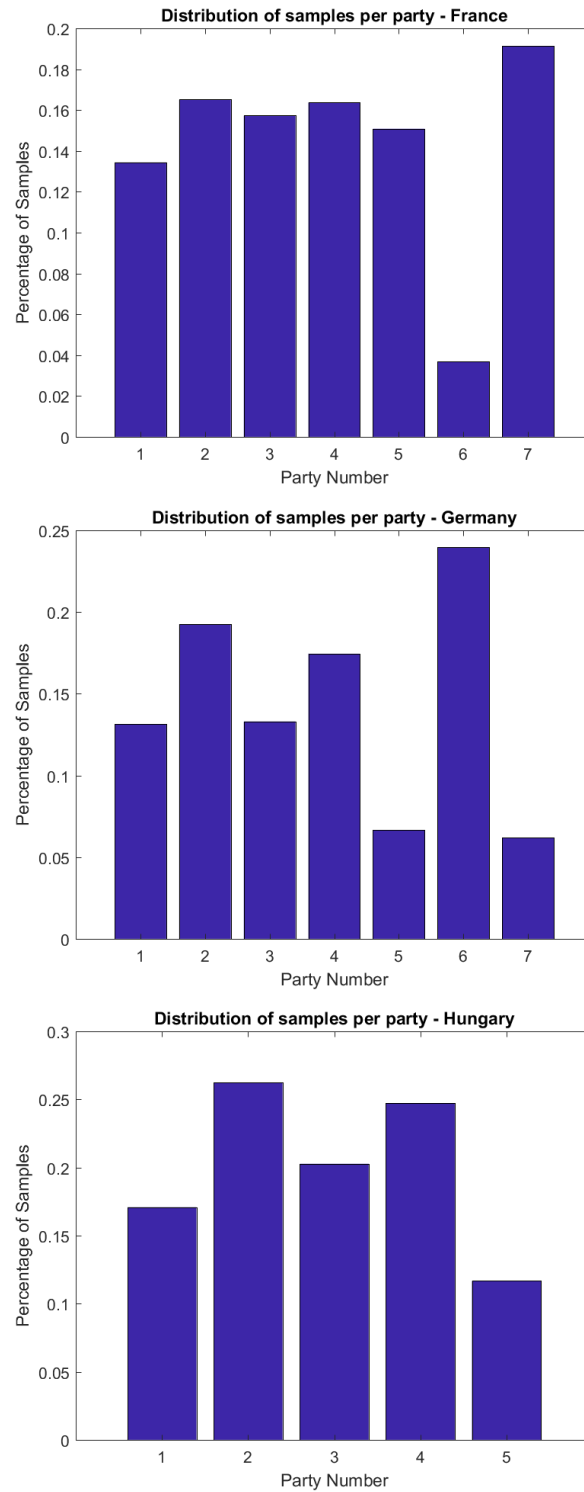
3.1.1 Αναζήτηση της ιδεολογικής ταύτισης

Στην παρούσα ενότητα διερευνάται κατά πόσο υπάρχει ιδεολογική ταύτιση μεταξύ των κομμάτων και των χρηστών που τα υποστηρίζουν. Θεωρείται πως ιδεολογική ταύτιση υπάρχει στην περίπτωση που οι χρήστες ενός κόμματος, απαντούν όλοι με παρόμοιο τρόπο στις δηλώσεις πολιτικής του VAA. Αυτό υποδηλώνει πως οι συγκεκριμένοι χρήστες ανήκουν στην ίδια ομάδα, έχουν την ίδια ιδεολογία και είναι συσπειρωμένοι.

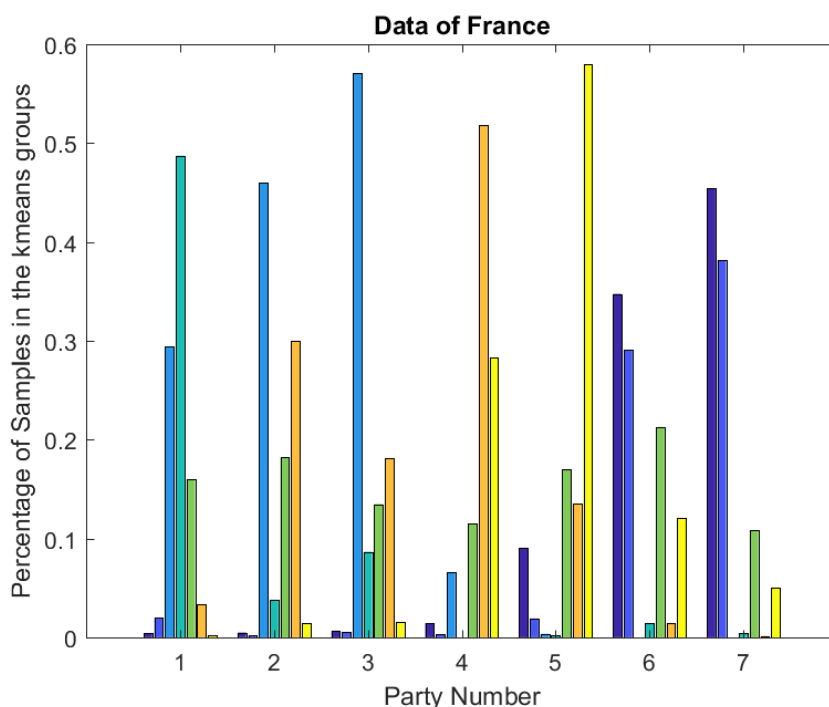
Η πειραματική διαδικασία αυτής της Ενότητας εξετάζει πόσοι από τους χρήστες που υποστηρίζουν το ίδιο κόμμα δίνουν στην πραγματικότητα και τις ίδιες απαντήσεις στις δηλώσεις πολιτικής. Όσο πιο μεγάλο είναι το συγκεκριμένο ποσοστό, τόσο πιο μεγάλη είναι και η ιδεολογική ταύτιση μεταξύ των χρηστών και του κόμματος. Επιπρόσθετα, μέσα από τα πειράματα που διεξάγονται, μας δίνεται η ευκαιρία να διερευνήσουμε τα ιδεολογικά χαρακτηριστικά ενός κόμματος βάσει των ψηφοφόρων του. Επίσης μπορεί να μελετηθεί εάν όντως υπάρχει πολυμορφία στα κόμματα, βλέποντας σε πόσες ομάδες ομοίων χρηστών μπορούν να χωριστούν οι χρήστες που ανήκουν στο ίδιο κόμμα, ανάλογα με τις απαντήσεις τους στις δηλώσεις πολιτικής.

Η πειραματική διαδικασία που ακολουθείται περιγράφεται ως εξής:

- Αρχικά οι χρήστες της κάθε χώρας χωρίζονται σε ομάδες σύμφωνα με την ομοιότητα στον τρόπο που απάντησαν στις δηλώσεις πολιτικής. Για την ομαδοποίηση (clustering) τους χρησιμοποιείται ο αλγόριθμος k -means [90]. Πρόκειται για ένα επαναληπτικό αλγόριθμο διαμερισμού δεδομένων, ο οποίος ταξινομεί τον κάθε χρήστη στην ομάδα, της οποίας το μέσο διάνυσμα απέχει λιγότερο από το διάνυσμα των απαντήσεων του χρήστη.



Σχήμα 3.1: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Γαλλία (β) Γερμανία (γ) Ουγγαρία



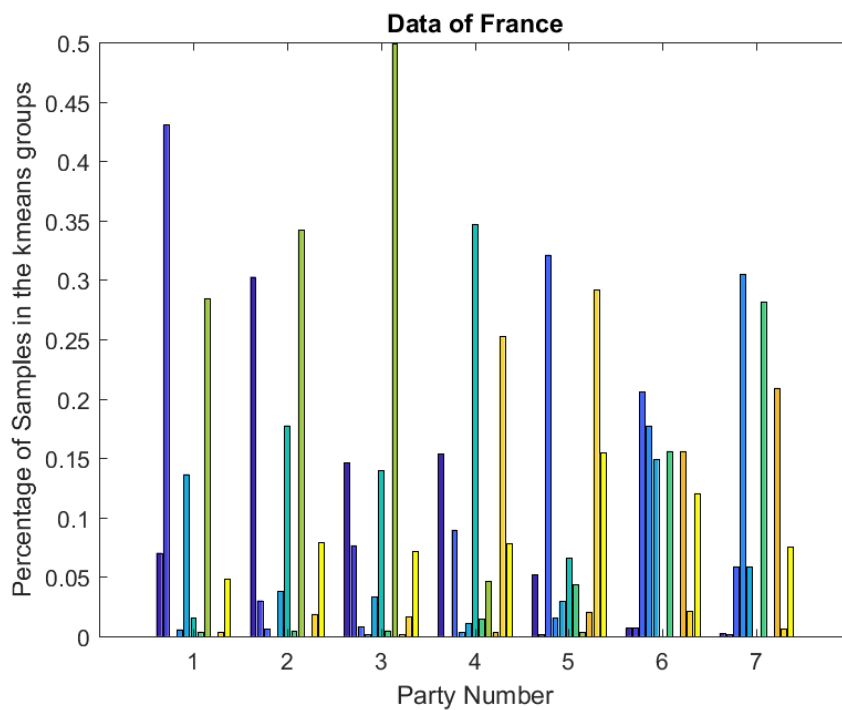
Σχήμα 3.2: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (7 ομάδες)

- Ο αριθμός των ομάδων επιλέγεται από την αρχή. Στην προκειμένη περίπτωση δημιουργούνται ομάδες ίσες με τον αριθμό των κομμάτων που λαμβάνουν μέρος στο *EUvox* και στη συνέχεια η διαδικασία επαναλαμβάνεται ακόμα τρεις φορές για να κατασκευαστούν ομάδες διπλάσιες, τριπλάσιες και ενάμιση φορά μεγαλύτερες από τον αριθμό αυτό.
- Το δείγμα που κατανέμεται σε κάθε ομάδα μελετάται για να εξαχθούν συμπεράσματα σχετικά με τη συσπείρωση που υπάρχει σε κάθε κόμμα. Επιπρόσθετα γίνεται διερεύνηση των ιδεολογικών χαρακτηριστικών των κομμάτων βάσει των χρηστών που τα υποστηρίζουν.

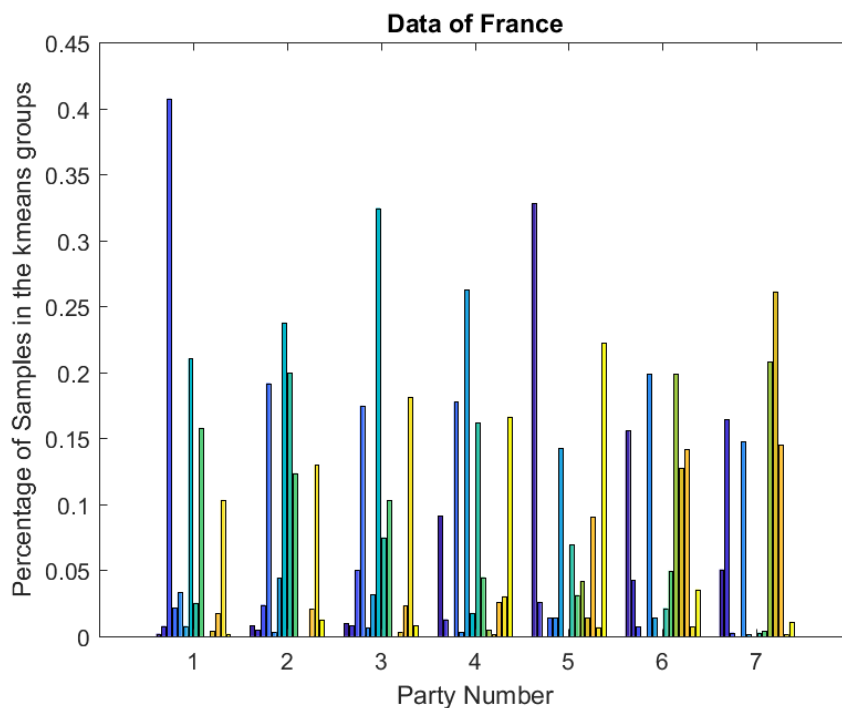
Για την περιγραφή των αποτελεσμάτων που προκύπτουν από την ομαδοποίηση, χρησιμοποιούνται δύο ειδών γραφικές μέθοδοι: Το ραβδόγραμμα (bar chart) και το κυκλικό διάγραμμα (pie chart). Το ραβδόγραμμα δείχνει σε πόσες επιμέρους ομάδες χωρίζονται οι χρήστες του κάθε κόμματος σύμφωνα με την ομοιότητα στις απαντήσεις τους. Επίσης φαίνεται το ποσοστό των χρηστών που κατανέμεται στην κάθε ομάδα. Το κυκλικό διάγραμμα παρουσιάζει τις ομάδες που δημιουργήθηκαν από τους όμοιους χρήστες και από ποια κόμματα προέρχονται οι χρήστες αυτοί.

Σε αυτήν την ενότητα αναλύονται τα αποτελέσματα για τις χώρες Γαλλία, Γερμανία και Ουγγαρία. Επιλέχθηκαν οι συγκεκριμένες χώρες μιας και σε αυτές τις περιπτώσεις δεν υπάρχει μεγάλη διακύμανση μεταξύ του αριθμού των χρηστών που βρίσκονται σε κάθε κόμμα, εκτός από τα μικρά κόμματα στα οποία η διαφορά είναι εμφανής. Στο Παράρτημα V, Παράρτημα VI και Παράρτημα VII βρίσκονται τα αποτελέσματα της ομαδοποίησης για τις υπόλοιπες χώρες.

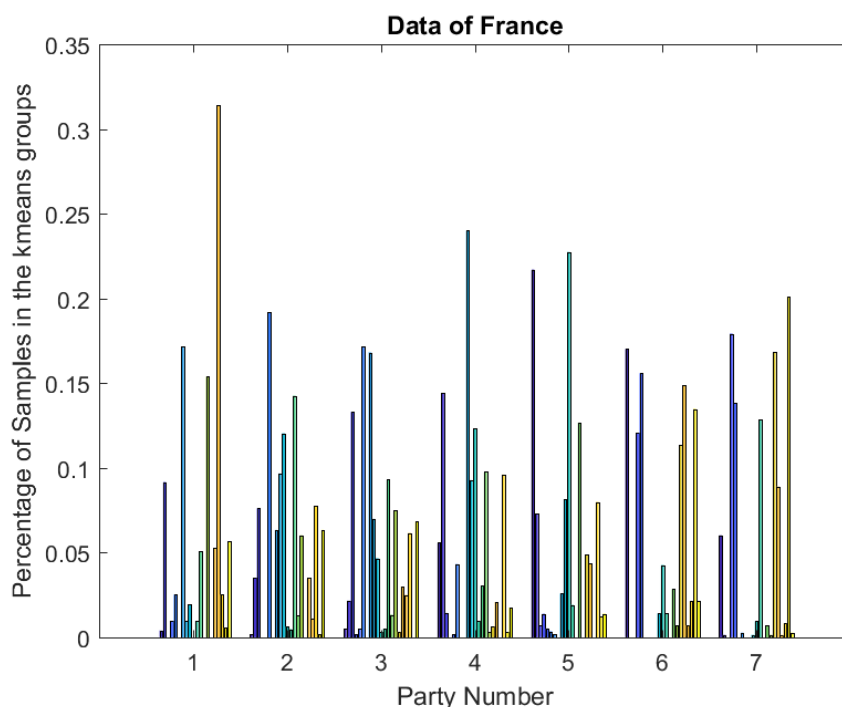
Στο Διάγραμμα III.3 φαίνεται η κατανομή των χρηστών σε κάθε κόμμα, σύμφωνα με την πρόθεση ψήφου τους, για τις χώρες Γαλλία, Γερμανία και Ουγγαρία. Υπό ιδανικές συνθήκες, οι χρήστες που βρίσκονται στα κόμματα αυτά θα έπρεπε να απαντούν με παρόμοιο τρόπο στις δηλώσεις πολιτικής του VAA, μιας και προτιθενται να ψηφίσουν το ίδιο κόμμα στις επερχόμενες εκλογές. Αυτό θα σήμαινε πως οι χρήστες μοιράζονται την ίδια πολιτική ιδεολογία, κάτι που είναι λογικό αφού υποστηρίζουν το ίδιο κόμμα. Ωστόσο τα αποτελέσματα από την ομαδοποίηση δε συμφωνούν απόλυτα με αυτή την υπόθεση.



Σχήμα 3.3: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (11 ομάδες)



Σχήμα 3.4: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (14 ομάδες)

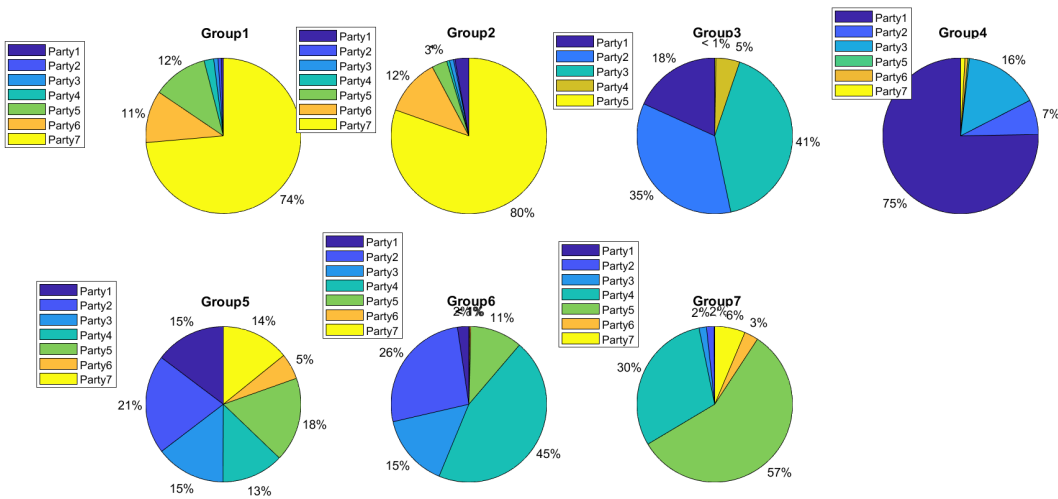


Σχήμα 3.5: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γαλλία (21 ομάδες)

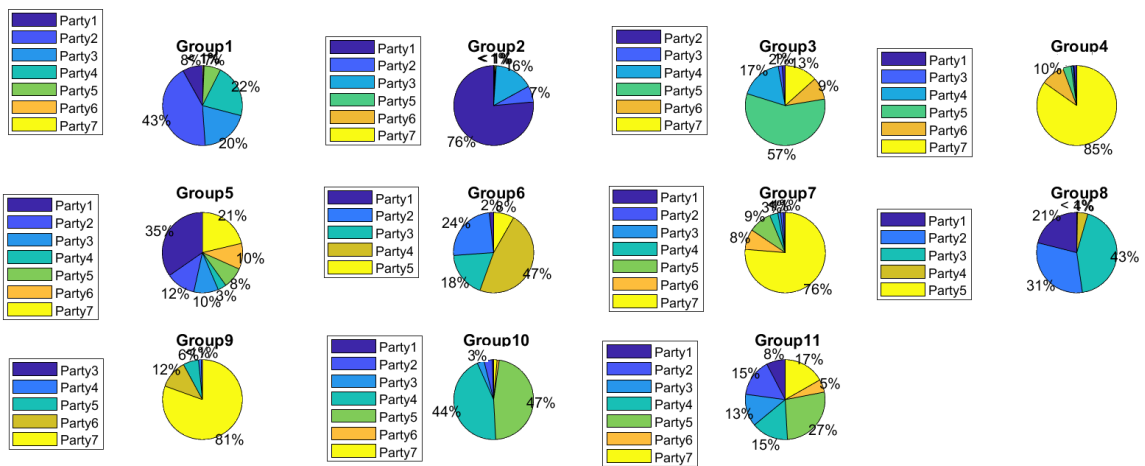
Κατά την πειραματική διαδικασία οι χρήστες της Γαλλίας και Γερμανίας μοιράστηκαν στην αρχή σε 7 ομάδες ομοίων χρηστών (οι χώρες αυτές αποτελούνται από 7 κόμματα), μετά σε 11 ομάδες, στη συνέχεια σε 14 ομάδες και τέλος σε 21 ομάδες, ενώ αυτοί της Ουγγαρίας χωρίστηκαν σε 5 ομάδες ομοίων χρηστών (η χώρα περιέχει 5 κόμματα), 8 ομάδες, 10 ομάδες και στο τέλος σε 15 ομάδες. Τα Διαγράμματα 3.2 – 3.5, τα Διαγράμματα 3.10 – 3.13 και τα Διαγράμματα 3.18 – 3.21 αναφέρονται στα ραβδογράμματα που δημιουργήθηκαν για τη Γαλλία, Γερμανία και Ουγγαρία αντίστοιχα. Στα διαγράμματα αυτά παρουσιάζεται η κατανομή των χρηστών σε ομάδες, σύμφωνα με την ομοιότητα στις απαντήσεις τους. Στον άξονα x βρίσκεται ο αριθμός των κομμάτων της κάθε χώρας και ο άξονας y δείχνει το ποσοστό των χρηστών που μοιράζεται σε κάθε ομάδα. Όσο πιο λίγες ομάδες δημιουργούνται στο κάθε κόμμα, τόσο πιο μεγάλη είναι η συσπείρωση του κόμματος και η ιδεολογική ταύτιση μεταξύ των χρηστών.

Τα Διαγράμματα που δημιουργούνται προδιαγράφουν την απόδοση του Mahalanobis ταξινομητή (βλέπε Ενότητα 1.7.1 για αυτή τη μέθοδο) όταν εφαρμόζεται στο SVAA (βλέπε Ενότητα 2.1). Όσον αφορά τη Γαλλία, από τα Διαγράμματα 3.2 – 3.5 φαίνεται πως υπάρχει ποικιλομορφία στα κόμματα. Μάλιστα κάποια κόμματα μοιάζουν μεταξύ τους σε σημείο που να μοιράζονται τους χρήστες τους (βλέπε Διαγράμματα 3.6 – 3.9). Τα κόμματα 2 και 3 φαίνεται να έχουν χρήστες που δίνουν όμοιες απαντήσεις, κάτι που συμβαίνει και μεταξύ των κομμάτων 4 και 5, καθώς και με τα κόμματα 6 και 7. Αυτό αποδίδεται στο γεγονός πως οι χρήστες στις συγκεκριμένες περιπτώσεις χωρίζονται στις ίδιες ομάδες ομοίων χρηστών. Άρα, σύμφωνα με τους χρήστες, αυτά τα κόμματα έχουν παρόμοια ιδεολογικά χαρακτηριστικά.

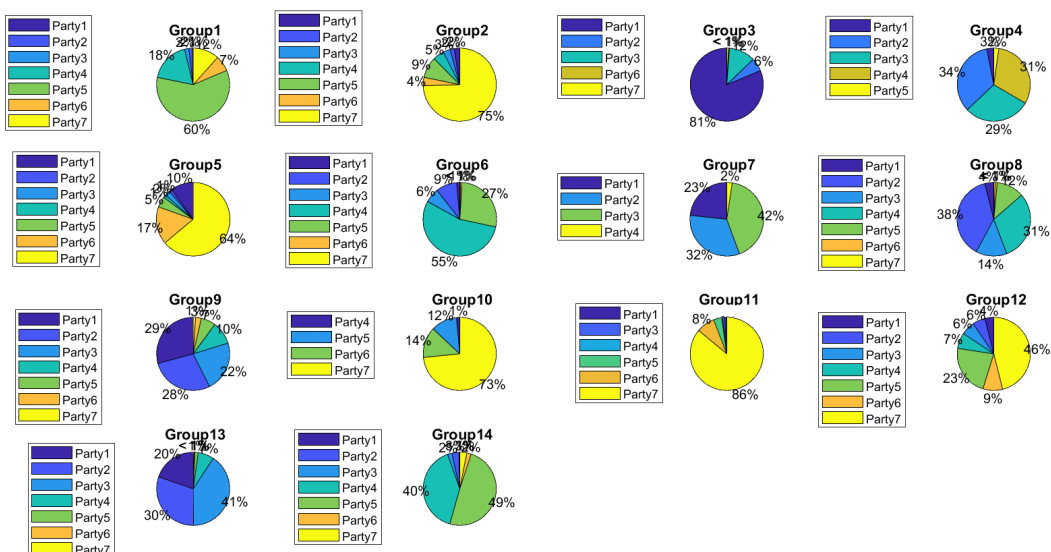
Ο ταξινομητής Mahalanobis μπερδεύεται όταν δεν υπάρχουν ξεκάθαρα διαχωρίσιμες κατηγορίες, μιας και όταν η μια κατηγορία μοιάζει με την άλλη δεν μπορεί να προβλέψει καθαρά σε ποια από αυτές να ταξινομήσει το δείγμα. Έτσι στην προκειμένη περίπτωση δεν μπορεί να προβλέψει με ακρίβεια σε ποιο από τα παρόμοια κόμματα ανήκει ένας χρήστης, ανάλογα με τις δηλώσεις πολιτικής που έδωσε. Υπό αυτές τις συνθήκες είναι πιο πιθανόν, όταν επιλέγει μεταξύ δύο κομμάτων, να ταξινομήσει το χρήστη στο κόμμα με το μεγαλύτερο δείγμα. Όπως δείχνει το Διάγραμμα III.3, το κόμμα 6 είναι αυτό με το μικρότερο δείγμα. Ακολουθώντας, αφού μοιάζει ιδεολογικά με το 7, ο ταξινομητής κατανέμει τους περισσό-



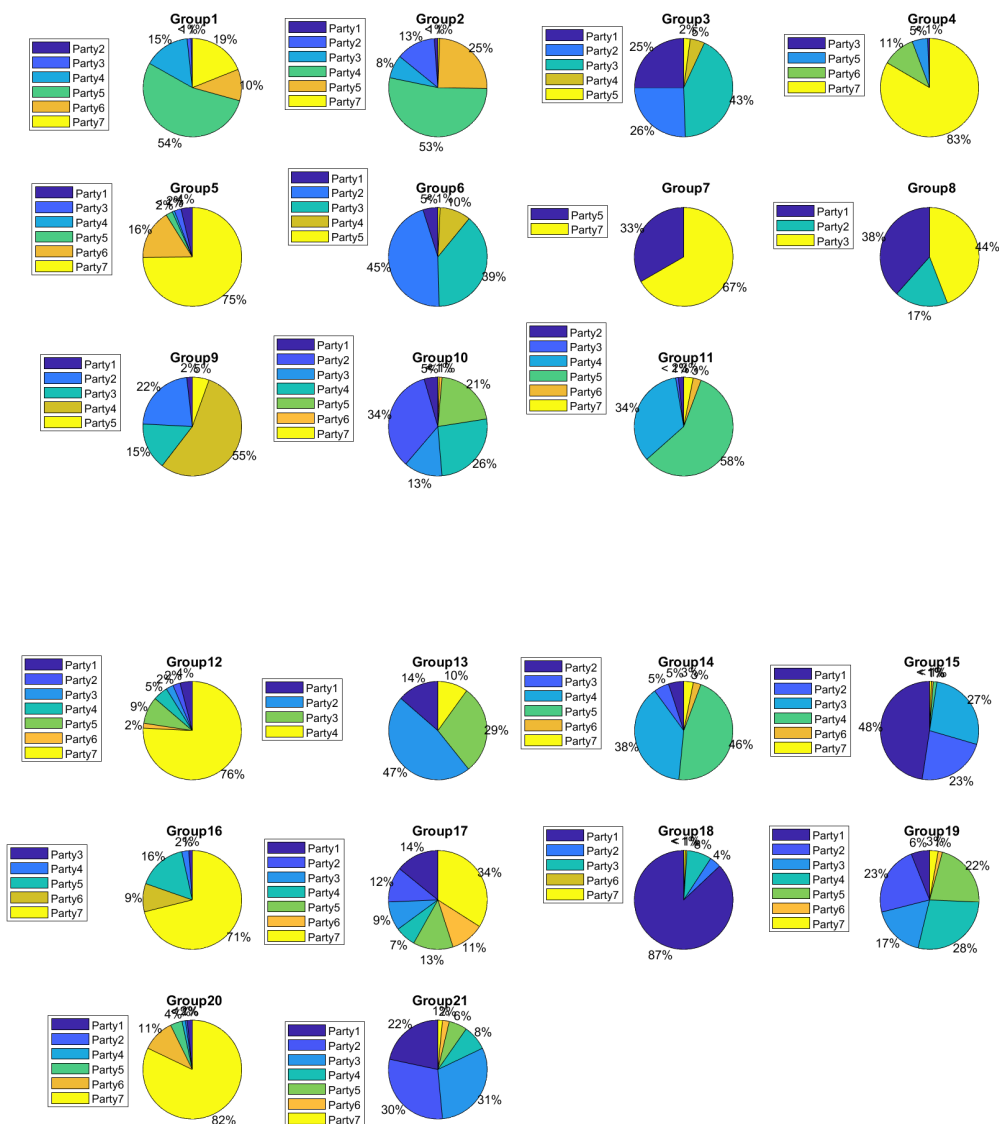
Σχήμα 3.6: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 7 ομάδες



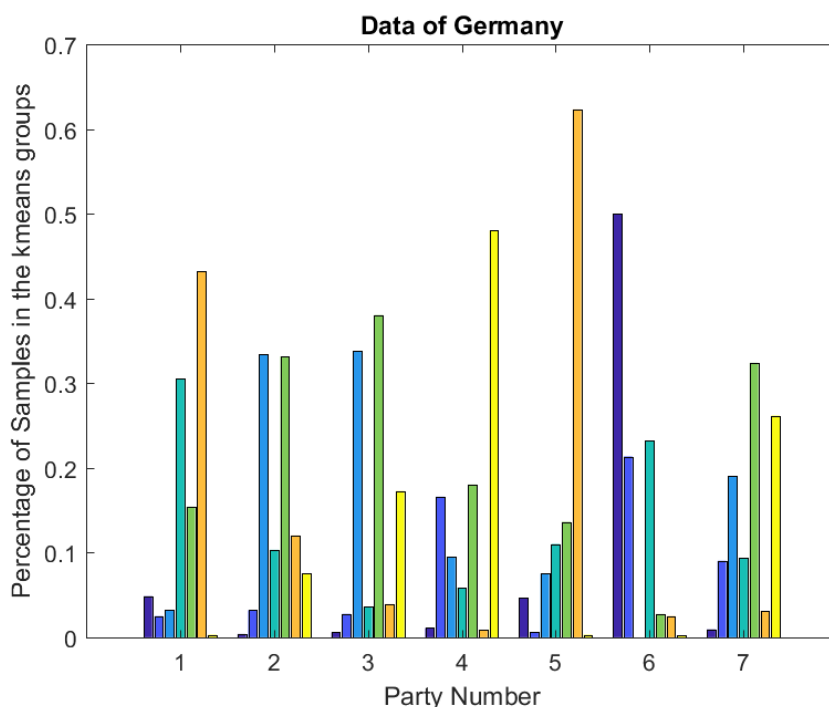
Σχήμα 3.7: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 11 ομάδες



Σχήμα 3.8: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 14 ομάδες



Σχήμα 3.9: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γαλλία – 21 ομάδες



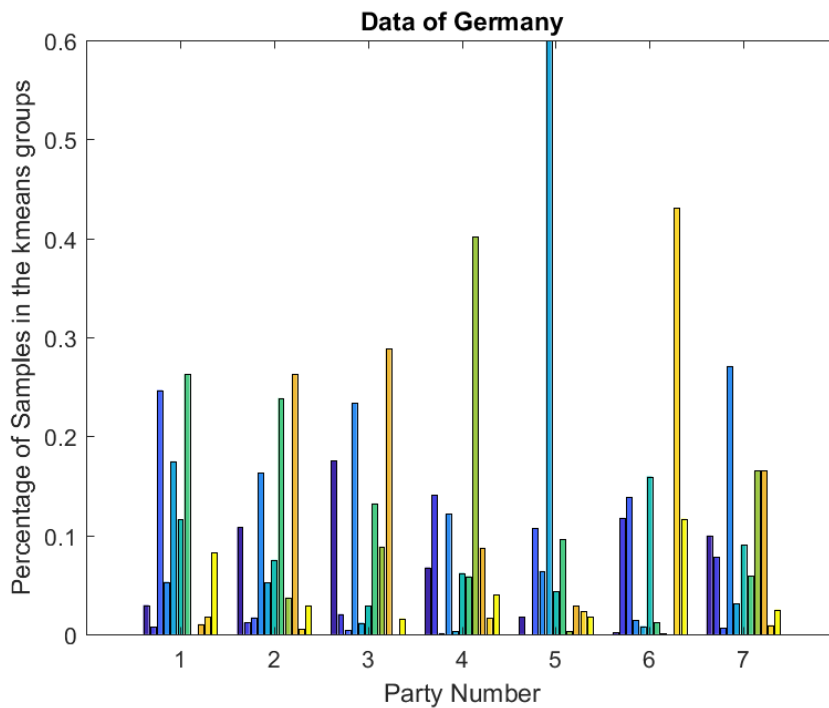
Σχήμα 3.10: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 7 ομάδες

τερους χρήστες στο κόμμα 7, που είναι πολύ μεγαλύτερο από το κόμμα 6, παρουσιάζοντας την καλύτερη του επίδοση σε σχέση με όλα τα κόμματα (βλέπε τα αποτελέσματα του Mahalanobis για τη Γαλλία στο Παράρτημα II.14).

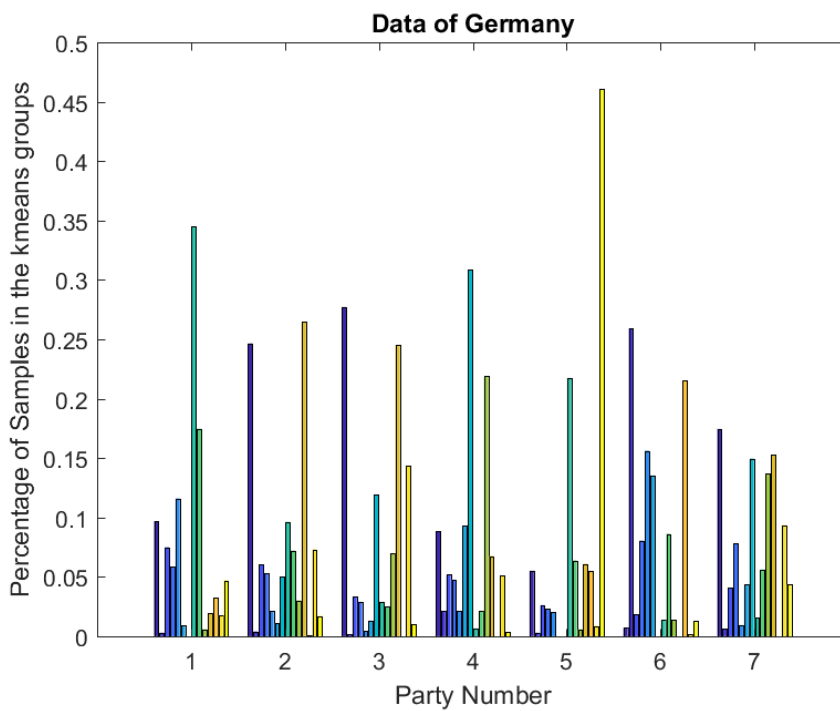
Στα κόμματα 2 με 3 καθώς και στα κόμματα 4 με 5 της Γαλλίας που είναι παρόμοια, ο ταξινομητής δεν μπορεί να ακολουθήσει την ίδια τακτική με αυτήν που εφάρμοσε στα κόμματα 6 με 7, να ταξινομήσει δηλαδή τους χρήστες στο κόμμα με το μεγαλύτερο δείγμα. Σε αυτές τις περιπτώσεις τα κόμματα έχουν σχεδόν τον ίδιο αριθμό χρηστών (βλέπε Διάγραμμα III.3). Έτσι, το κόμμα που ' ' κερδίζει" είναι αυτό που παρουσιάζει τη μεγαλύτερη συσπείρωση χρηστών και σύμφωνα με τα Διαγράμματα 3.2 – 3.5 αυτά είναι τα κόμματα 3 και 5. Ως εκ τούτου ο Mahalanobis ταξινομητής έχει καλύτερα αποτελέσματα στο κόμμα 3 αντί στο κόμμα 2 και στο κόμμα 5 αντί στο κόμμα 4 (βλέπε Παράρτημα II.14)

Ο Mahalanobis ταξινομητής είχε τη δεύτερη καλύτερη επίδοση στο κόμμα 1 της Γαλλίας (βλέπε Παράρτημα II.14). Αυτό συμβαίνει γιατί το συγκεκριμένο κόμμα είναι πιο διαχωρίσιμο από τα υπόλοιπα. Επίσης, όσο αυξανόταν ο αριθμός των ομοίων ομάδων, το 30% των χρηστών αυτού του κόμματος παρέμεινε στην ίδια ομάδα. Κάθε ομάδα δηλώνει το βαθμό ομοιότητας μεταξύ των χρηστών. Όσο αυξάνεται ο αριθμός των ομάδων, τα κριτήρια για να μπορέσει να εισέλθει ένας χρήστης σε κάποια ομάδα γίνονται πιο αυστηρά. Στη δεδομένη περίπτωση, καθώς αυξάνονται οι ομάδες αυξάνεται και ο βαθμός ομοιότητας που πρέπει να τηρείται μεταξύ των απαντήσεων των χρηστών, για να μπορέσει ένας χρήστης να καταταχθεί σε μια από αυτές. Κάθε χρήστης μπορεί να βρίσκεται μόνο σε μια ομάδα. Το πιο πάνω ποσοστό απευθύνεται στο ποσοστό των χρηστών που έμεινε συσπειρωμένο ακόμα και στην περίπτωση που υπήρχαν ομάδες τριπλάσιες των κομμάτων και θα ήταν πολύ εύκολο οι χρήστες να διασπαστούν σε αυτές. Όταν οι ομάδες ήταν ίσες με τον αριθμό των κομμάτων, δηλαδή 7 ομάδες (βλέπε Διάγραμμα 3.2) οι όμοιοι χρήστες στο κόμμα 1 ήταν γύρω στο 50%.

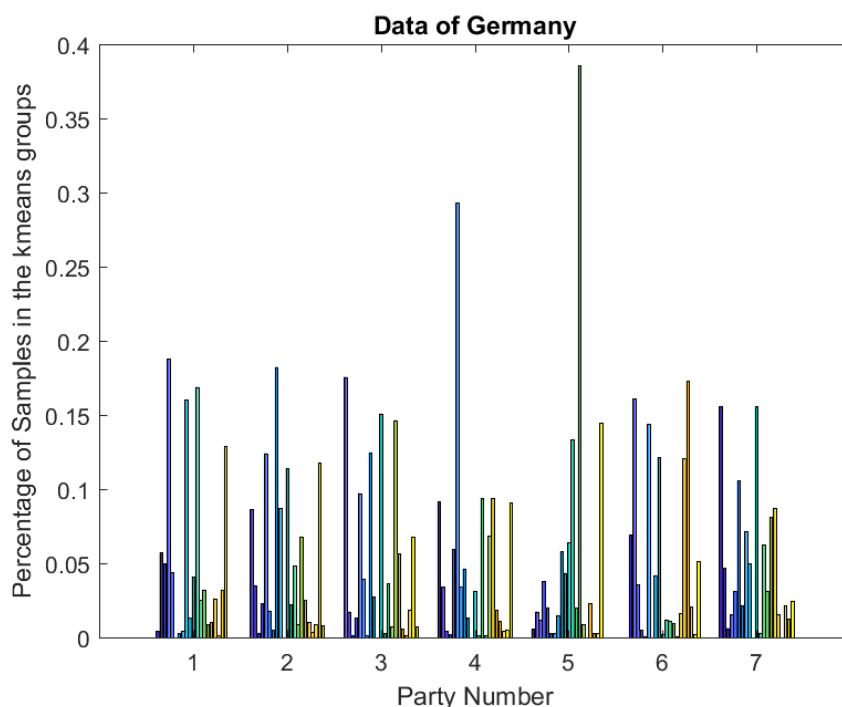
Σε αντίθεση με τον Mahalanobis ταξινομητή, ο SVM καταφέρνει να αναγνωρίσει τις πολλές κατηγορίες στις οποίες μπορούν να χωριστούν οι χρήστες του κάθε κόμματος και δίνει πιο ομοιόμορφα αποτελέσματα (βλέπε Παράρτημα II.14). Αυτός είναι και ο βασικός λόγος που η μέθοδος SVM αποδίδει καλύτερα



Σχήμα 3.11: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 11 ομάδες



Σχήμα 3.12: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 14 ομάδες



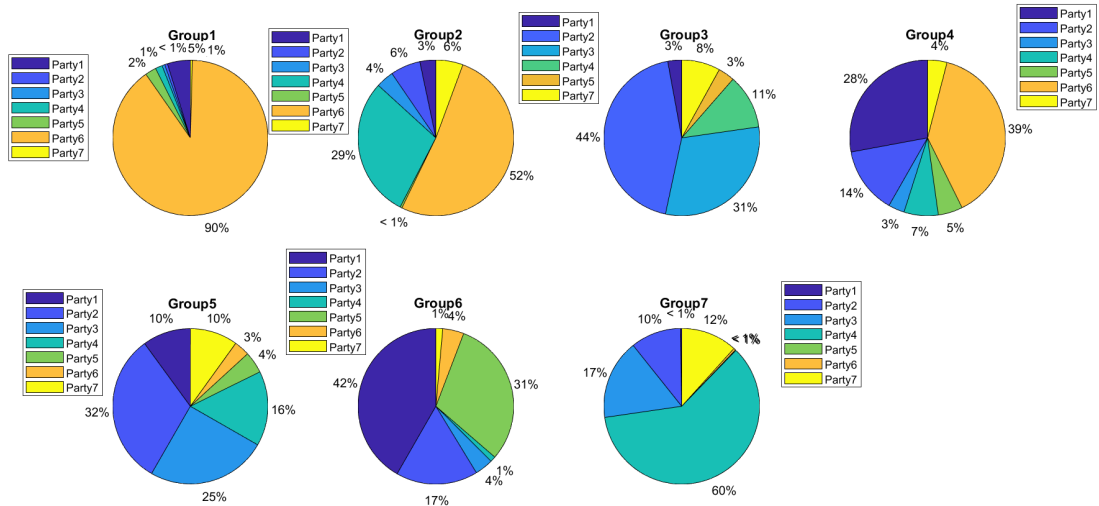
Σχήμα 3.13: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Γερμανία – 21 ομάδες

από όλες τις άλλες μεθόδους που εφαρμόστηκαν για την πρόβλεψη ψήφου των χρηστών.

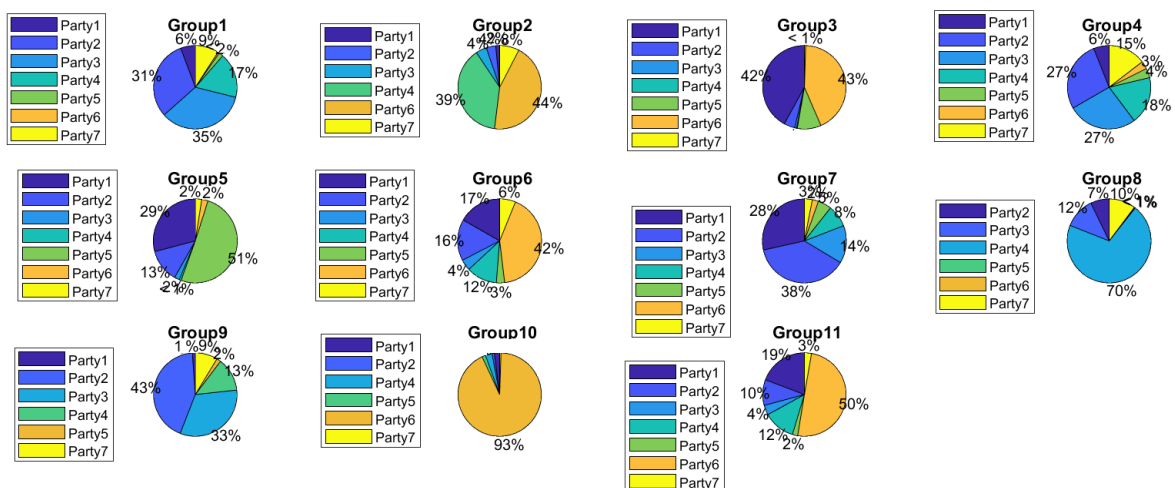
Σχετικά με τα κόμματα της Γερμανίας, για ακόμα μια φορά παρατηρείται ποικιλομορφία (βλέπε Διαγράμματα 3.10 – 3.13) και κόμματα που μοιάζουν ιδεολογικά μεταξύ τους. Τα κόμματα 2, 3 και 7 φαίνεται να έχουν χρήστες που μοιράζονται την ίδια ιδεολογία, όπως συμβαίνει και με τα κόμματα 1 και 5 (βλέπε Διαγράμματα 3.14 – 3.17). Και πάλι ο Mahalanobis ταξινομητής μπερδεύεται με τα όμοια κόμματα 2,3 και 7, αποδίδοντας με παρόμοιο τρόπο στα κόμματα 2 και 3, τα οποία δεν έχουν μεγάλη διαφορά στο μέγεθος και χειρότερα στο κόμμα 7 που είναι το μικρότερο από όλα τα κόμματα (βλέπε τα αποτελέσματα του Mahalanobis για τη Γερμανία στο Παράρτημα II.5 και το μέγεθος των κομμάτων στο Διάγραμμα III.3). Το κόμμα 7 'δυσκολεύει' ακόμα και τη μέθοδο SVM, αφού ενώ στα υπόλοιπα κόμματα παρουσιάζει μια αρκετά ισορροπημένη απόδοση, στο κόμμα 7 ο SVM ταξινομητής αποδίδει λιγότερο από τον Mahalanobis.

Τα κόμματα 1 και 5 της Γερμανίας που δείχνουν όμοια ιδεολογικά μεταξύ τους στην αρχή (βλέπε Διάγραμμα 3.10), καθώς αυξάνεται ο αριθμός των ομοίων ομάδων αρχίζουν να ξεχωρίζουν (βλέπε Διαγράμματα 3.11 – 3.13). Έτσι, εάν και διαφέρουν κατά πολύ σε μέγεθος, με το κόμμα 5 να είναι πολύ μικρότερο από το κόμμα 1, στο κόμμα 5 ο Mahalanobis ταξινομητής παρουσιάζει καλύτερα αποτελέσματα από το κόμμα 1. Αυτό συμβαίνει γιατί οι χρήστες στο κόμμα 5 παρουσιάζουν συσπείρωση, με το 40% από αυτούς να παραμένουν στην ίδια ομάδα, ακόμα και όταν οι ομάδες ήταν ίσες με τον τριπλάσιο αριθμό των κομμάτων (βλέπε Διάγραμμα 3.13). Στην περίπτωση που οι ομάδες ήταν 7, ίσες δηλαδή με τον αριθμό των κομμάτων, το ποσοστό της συσπείρωσης ήταν πάνω από 60% (βλέπε Διάγραμμα 3.10).

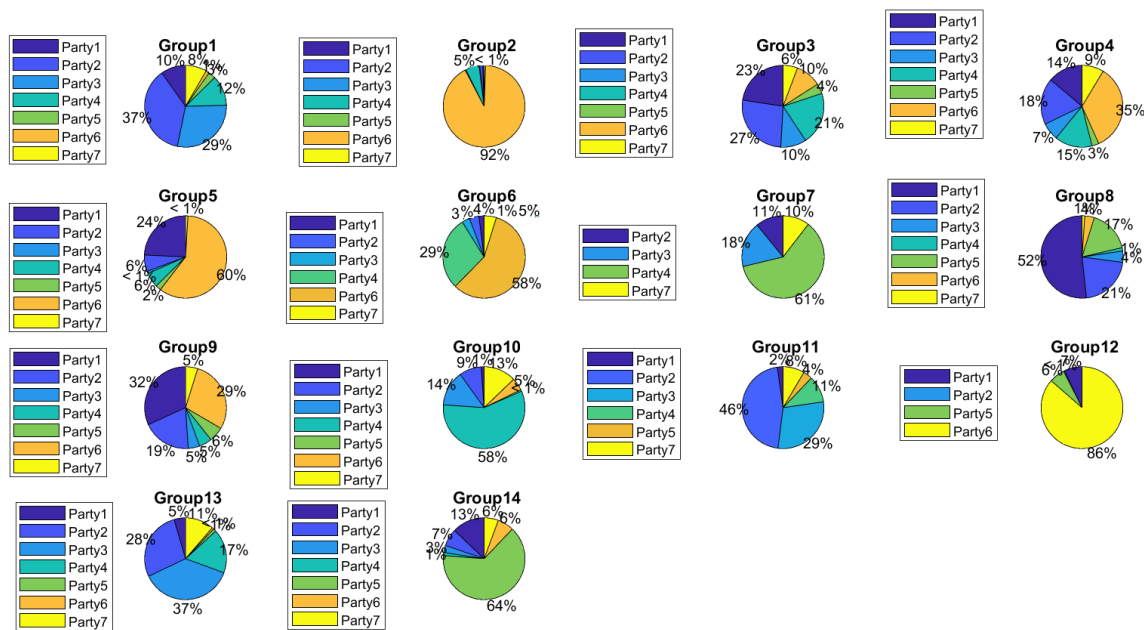
Τα κόμματα που διαφέρουν περισσότερο από τα υπόλοιπα είναι το 4 και το 6 για τη Γερμανία. Ως εκ τούτου ο Mahalanobis ταξινομητής έχει τις δύο καλύτερες επιδόσεις σε αυτά τα δύο κόμματα. Παρόλα αυτά, τα αποτελέσματα στο κόμμα 6 είναι μακράν καλύτερα από τα υπόλοιπα (βλέπε Παράρτημα II.5), μιας και είναι το μεγαλύτερο κόμμα και στις όμοιες ομάδες που κατανέμονται οι χρήστες του καταλαμβάνουν την πλειοψηφία (βλέπε Διαγράμματα 3.14 – 3.17).



Σχήμα 3.14: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 7 ομάδες



Σχήμα 3.15: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 11 ομάδες



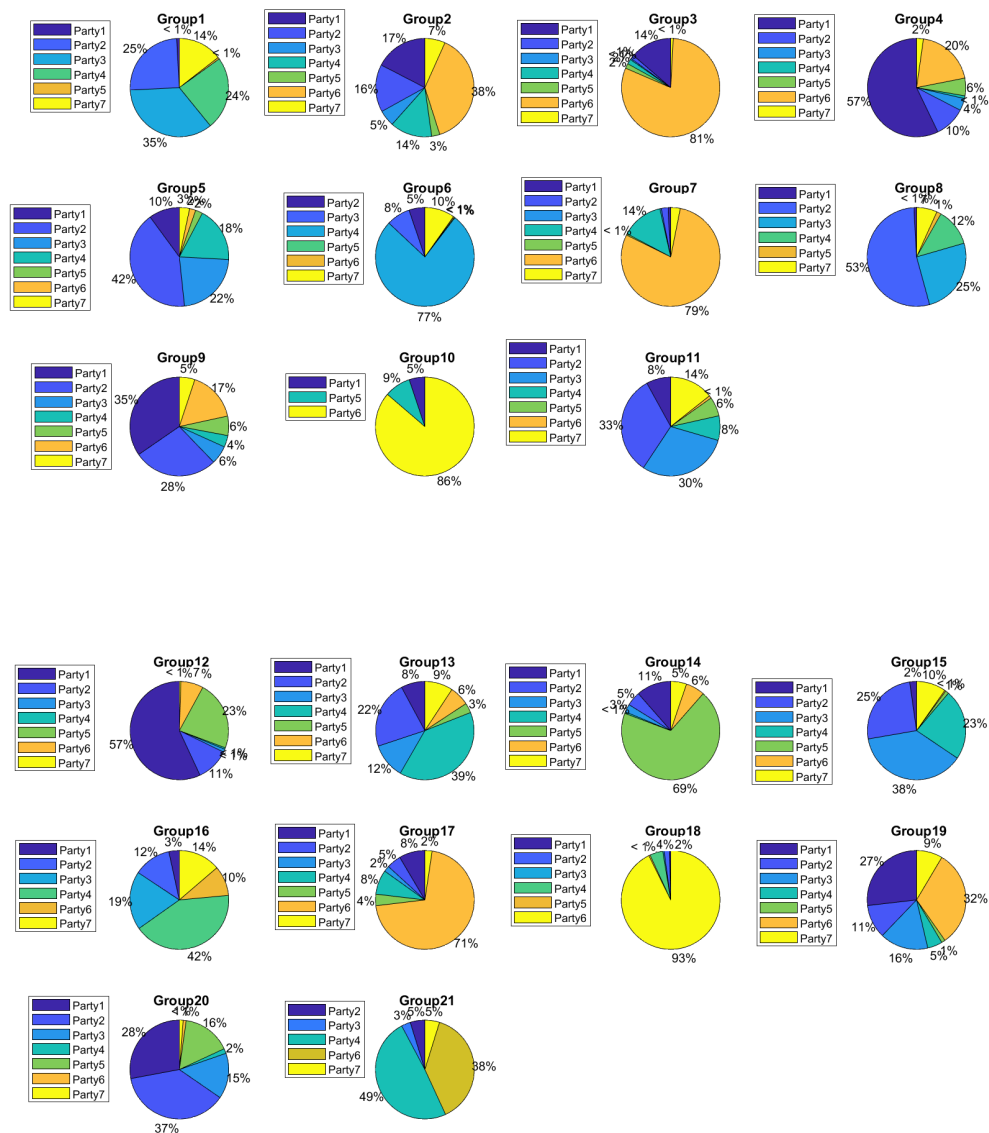
Σχήμα 3.16: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Γερμανία – 14 ομάδες

Για την Ουγγαρία, τα κόμματα που φαίνεται να έχουν όμοιους χρήστες είναι το κόμμα 4 με το 5 και το κόμμα 1 με το 2. Εδώ είναι ξεκάθαρο πως η πλειοψηφία κερδίζει στη χρήση του Mahalanobis ταξινομητή. Έτσι τα αποτελέσματα στο κόμμα 4 υπερτερούν αυτά του κόμματος 5 (το κόμμα 4 είναι μεγαλύτερο από το κόμμα 5) και η επίδοση στο κόμμα 2 είναι καλύτερη από ότι στο κόμμα 1 (το κόμμα 2 είναι μεγαλύτερο από το κόμμα 1), όπως δείχνουν τα αποτελέσματα του Mahalanobis για την Ουγγαρία στο Παράρτημα II.17 και το μέγεθος των κομμάτων στο Διάγραμμα III.3.

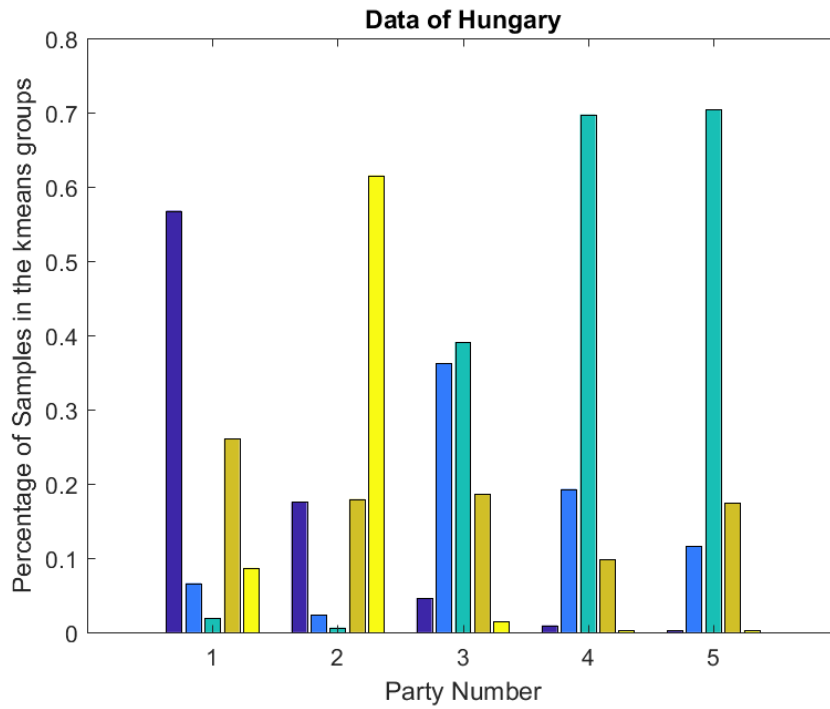
Αξιοσημείωτο είναι πως οι χρήστες στο κόμμα 1 και στο κόμμα 5 παρουσιάζουν μεγαλύτερη συσπείρωση από τα άλλα δύο κόμματα, μιας και το 40% και λίγο περισσότερο από το 30% των χρηστών αντίστοιχα παραμένει στην ίδια ομάδα όσο και αν αυξάνεται ο αριθμός των ομάδων στις οποίες μπορούν να διαχωριστούν (βλέπε Διαγράμματα 3.18 – 3.21). Παρόλα αυτά στο κόμμα 2 και στο κόμμα 4 οι περισσότεροι ταξινομητές αποδίδουν καλύτερα. Εκτός από το μέγεθος των κομμάτων, η καλή αυτή επίδοση πιθανόν να οφείλεται και στο γεγονός πως όταν οι χρήστες χωρίστηκαν σε ομάδες ίσες με τον αριθμό των κομμάτων, σε αυτά τα δύο κόμματα γύρω στο 60% και 70% αντίστοιχα των χρηστών παρουσίασαν ιδεολογική ταύτιση (βλέπε Διάγραμμα 3.18).

Το μεγαλύτερο ποσοστό των χρηστών στο κόμμα 3 της Ουγγαρίας, χωρίζεται αρχικά σε τρεις κατηγορίες (βλέπε Διάγραμμα 3.18). Οι δύο από αυτές περιέχουν τον ίδιο αριθμό χρηστών και είναι οι ίδιες κατηγορίες στις οποίες βρίσκεται το μεγαλύτερο ποσοστό των κομμάτων 4 και 5. Οι υπόλοιποι χρήστες από το κόμμα 3 κατανέμονται κυρίως σε μια τρίτη ομάδα, μέσα στην οποία βρίσκονται χρήστες από όλα τα κόμματα, καταλαμβάνοντας σχεδόν το ίδιο ποσοστό για κάθε κόμμα (βλέπε την ομάδα 4 στο Διάγραμμα 3.22). Έτσι, ο Mahalanobis ταξινομητής για αυτό το κόμμα παρουσιάζει τη χειρότερη συνολική επίδοση.

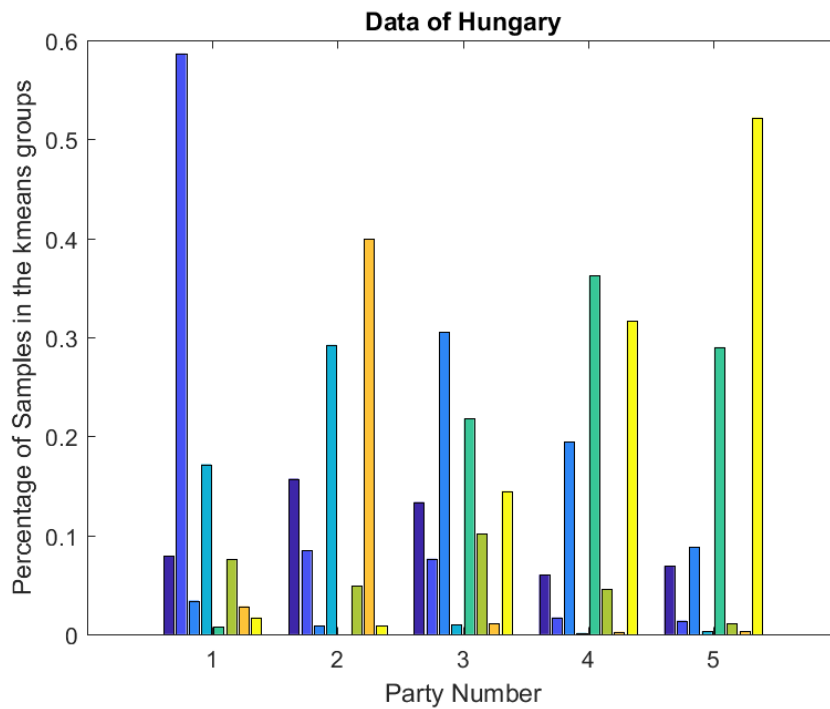
Για ακόμα μια φορά ο ταξινομητής SVM καταφέρνει να δώσει ισορροπημένα αποτελέσματα, με εξαίρεση την επίδοση του στο κόμμα 5 που δεν είναι καλή. Αυτό οφείλεται στο μικρό μέγεθος του κόμματος 5 και στη μεγάλη ομοιότητα μεταξύ των χρηστών σε αυτό το κόμμα και το κόμμα 4, το οποίο είναι ένα από τα μεγαλύτερα κόμματα. Ως εκ τούτου ο ταξινομητής οδηγείται σε λανθασμένες εκτιμήσεις, αφού δυσκολεύεται να θέσει τα όρια μεταξύ των δύο αυτών κομμάτων.



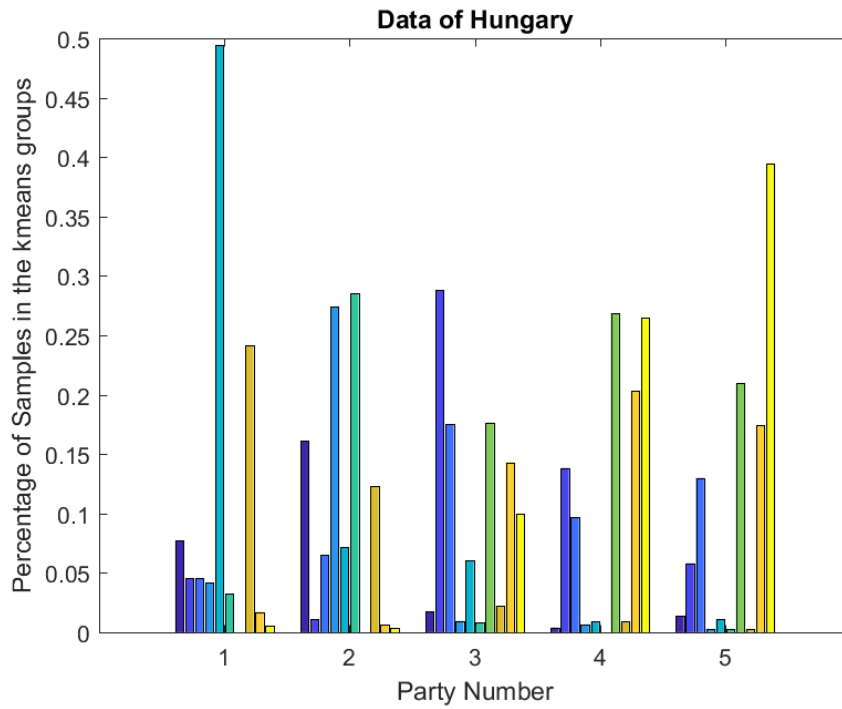
Σχήμα 3.17: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα –Γερμανία – 21 ομάδες



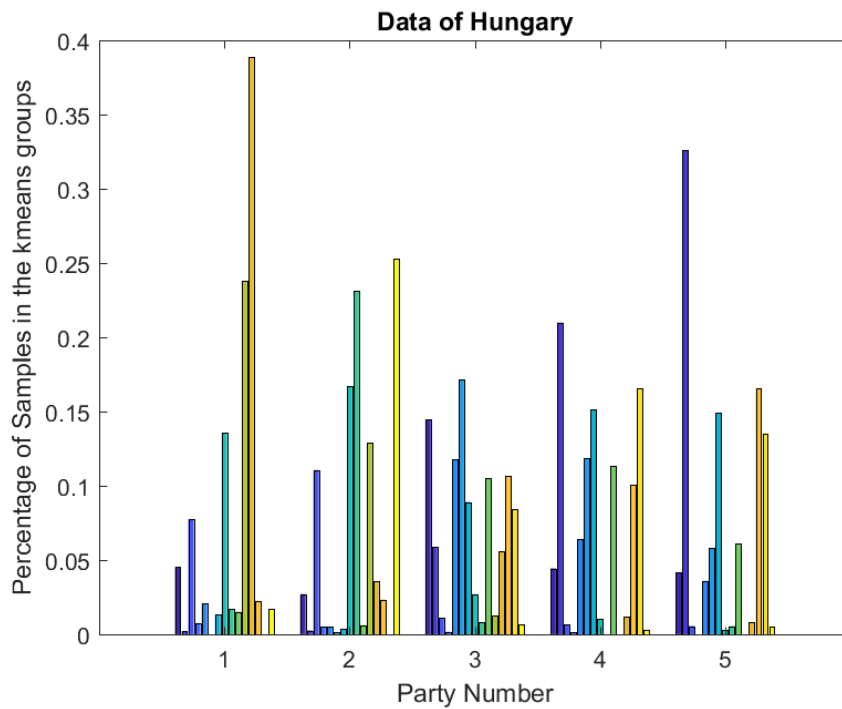
Σχήμα 3.18: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 5 ομάδες



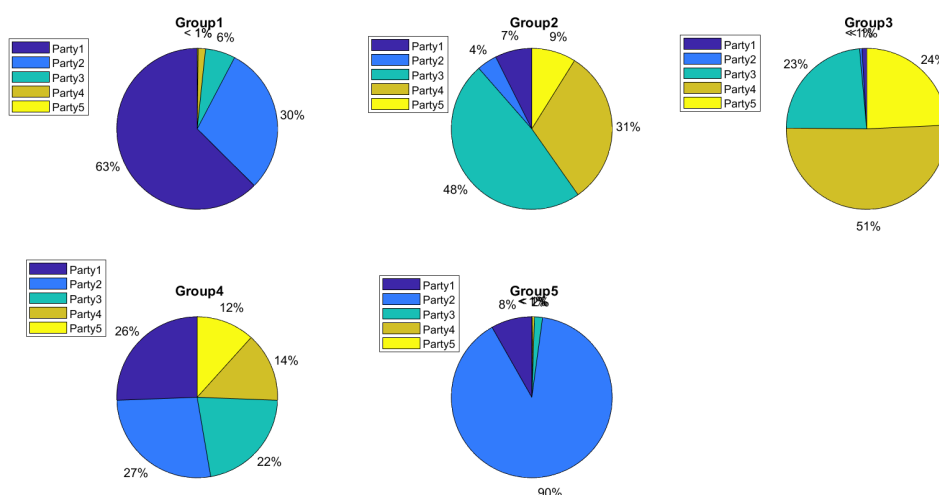
Σχήμα 3.19: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 8 ομάδες



Σχήμα 3.20: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 10 ομάδες



Σχήμα 3.21: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ουγγαρία – 15 ομάδες



Σχήμα 3.22: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 5 ομάδες

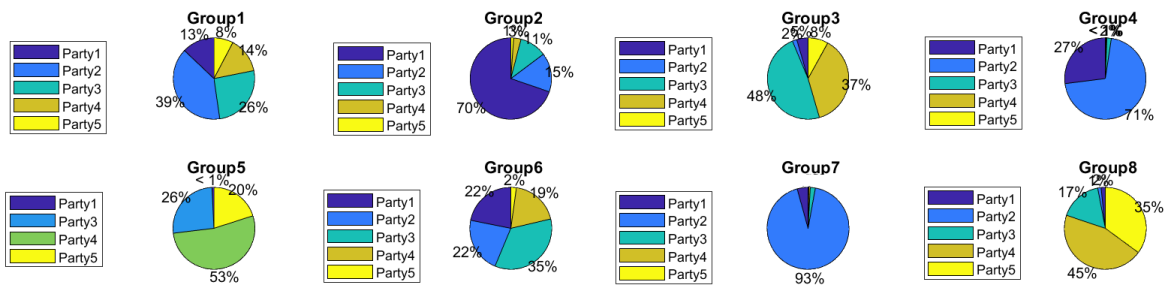
Τα Διαγράμματα 3.6 – 3.8, Διαγράμματα 3.14 – 3.16 και Διαγράμματα 3.22 – 3.24 δείχνουν τις ομάδες που δημιουργούνται κάθε φορά από τους όμοιους χρήστες και από ποια κόμματα προέρχονται οι χρήστες αυτοί για τη Γαλλία, Γερμανία και Ουγγαρία αντίστοιχα. Οι περισσότερες ομάδες κυριαρχούνται από συγκεκριμένα κόμματα και αυτή η παρατήρηση γίνεται πιο εμφανής όταν ο αριθμός των κομμάτων αυξάνεται.

Στο Παράρτημα V, Παράρτημα VI και Παράρτημα VII βρίσκονται τα αποτελέσματα της ομαδοποίησης για τις υπόλοιπες χώρες.

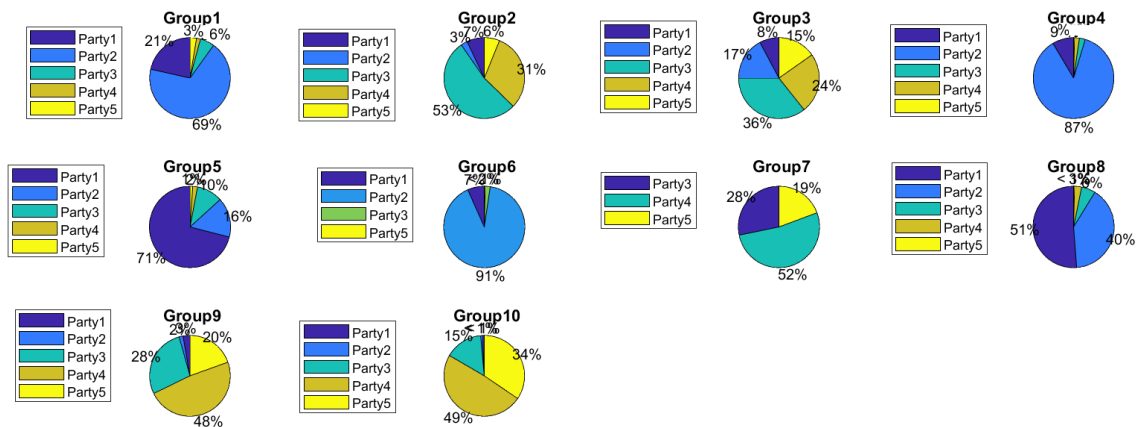
Η πειραματική διαδικασία που εφαρμόστηκε αποδεικνύει πως όντως υπάρχει πολυμορφία στο δείγμα της κάθε χώρας. Για αυτό υπάρχουν πολλοί χρήστες που δίνουν όμοιες απαντήσεις στις δηλώσεις πολιτικής αλλά δεν υποστηρίζουν το ίδιο κόμμα. Σε όλες τις περιπτώσεις σε κάθε κόμμα δημιουργούνται πολλές μικρές ομάδες από όμοιους χρήστες που δημιουργούν “θόρυβο” και αλλοιώνουν την επίδοση του SVAA, η οποία φαίνεται στο Παράρτημα II. Αυτός ο “θόρυβος” μπερδεύει το σύστημα, το οποίο, στην περίπτωση κυρίως που εφαρμόζεται η μέθοδος Mahalanobis, τείνει να ταξινομή τους χρήστες στο κόμμα με το μεγαλύτερο μέγεθος.

Ωστόσο σε κάποια κόμματα, ένα μεγάλο ποσοστό των χρηστών είναι συσπειρωμένο. Αυτό δείχνει πως υπάρχει ιδεολογική ταύτιση μεταξύ των χρηστών και των κομμάτων, αλλά σε διαφορετικό βαθμό από χώρα σε χώρα και από κόμμα σε κόμμα. Όταν υπάρχει συσπείρωση μεταξύ των χρηστών, ακόμα και στα μικρά κόμματα το σύστημα μπορεί να δώσει πολύ ικανοποιητικά αποτελέσματα. Ενώ στα μεγάλα κόμματα, όταν δεν υπάρχει ιδεολογική ταύτιση μεταξύ των χρηστών και των κομμάτων που υποστηρίζουν, η επίδοση του συστήματος μειώνεται.

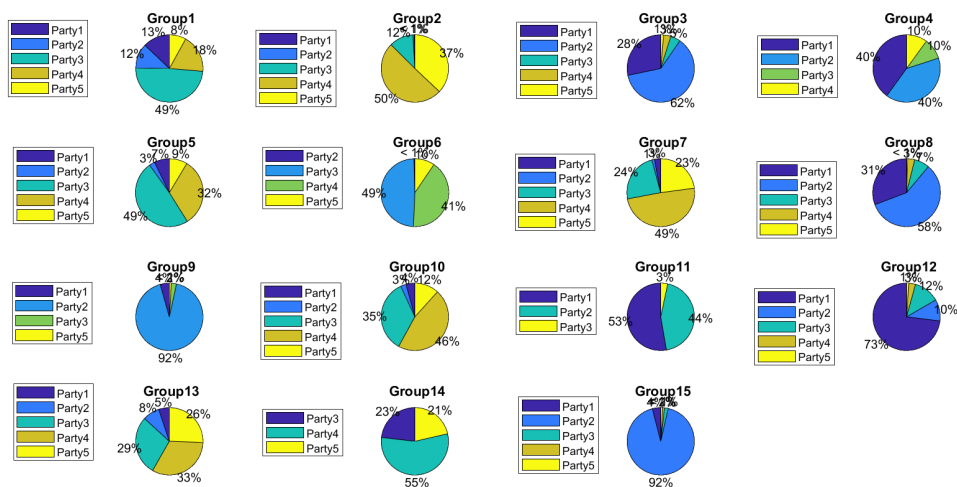
Οι διάφορες ομάδες από όμοιους χρήστες που δημιουργούνται σε κάθε χώρα οφείλονται πιθανόν στο διαφορετικό πολιτικό υπόβαθρο μεταξύ των χρηστών καθώς και στα διαφορετικά τους δημογραφικά χαρακτηριστικά.



Σχήμα 3.23: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 8 ομάδες



Σχήμα 3.24: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 10 ομάδες



Σχήμα 3.25: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ουγγαρία – 15 ομάδες

3.2 Το ενδιαφέρον του Ευρωπαίου χρήστη των VAAs για την πολιτική

Στη σημερινή εποχή οι πολίτες τείνουν να μην ενημερώνονται για θέματα σχετικά με την πολιτική και οι λόγοι ποικίλουν: υπάρχει ο κόσμος που απλά αδιαφορεί για οτιδήποτε περιστρέφεται γύρω από την πολιτική, οι πολίτες που λόγω ανεπάρκειας ελεύθερου χρόνου δεν μπορούν να εμπλακούν με θέματα πολιτικής καθώς και αυτοί που εξαιτίας της δυσaréσκειας τους προς την υφιστάμενη κυβέρνηση ή τις παλαιότερες κυβερνήσεις της χώρας σταμάτησαν να ασχολούνται με την όλη εκλογική διαδικασία [37]. Τα άτομα που απέχουν από τις εκλογικές διαδικασίες συνήθως έχουν τα ακόλουθα χαρακτηριστικά:

1. δυσκολεύονται να αντιληφθούν με ποιόν υποψήφιο ταιριάζουν οι απόψεις τους,
2. έχουν μικρά εισοδήματα, χαμηλό επίπεδο μόρφωσης και ανήκουν στα χαμηλά κοινωνικά στρώματα της χώρας,
3. δεν νιώθουν αρκετά αρμόδιοι για να ψηφίσουν και αυτό συμβαίνει κυρίως λόγω της έλλειψης γνώσης σχετικά με τα πολιτικά θέματα. Έχει σημειωθεί πως οι πολίτες που επιλέγουν να μην ψηφίζουν είναι λιγότερο πληροφορημένοι για τα πολιτικά θέματα από αυτούς που ασκούν το δικαίωμα ψήφου [91].

Αυτό που απασχολεί πολλούς ερευνητές, οι οποίοι ασχολούνται με την εκλογική διαδικασία είναι κατά πόσο τα VAAs δίνουν ώθηση στους πολίτες για να ψηφίσουν και εάν η σύσταση που τους γίνεται από τα συστήματα αυτά επηρεάζει την τελική τους ψήφο [50, 42, 39, 49]. Έρευνες έχουν δείξει πως οι προτάσεις που γίνονται μέσω των VAAs έχουν επίδραση σε ένα σημαντικό κομμάτι των ψηφοφόρων για την τελική επιλογή του υποψηφίου ή του κόμματος που θα υποστηρίξουν στις εκλογές, ειδικά εάν οι συγκεκριμένοι είναι αναποφάσιστοι, γυναίκες, άτομα κάτω των 34 χρόνων και άτομα που ψηφίζουν για πρώτη φορά [42, 39]. Μάλιστα οι χρήστες των VAAs που ανέφεραν ότι επηρεάστηκαν ή ξαφνιάστηκαν από τη σύσταση που τους έγινε, ήταν πολύ πιθανόν να αλλάξουν την αρχική ιδέα για την ψήφο τους, ιδιαίτερα εάν αυτοί ήταν νεαροί [92]. Επίσης σε αρκετές περιπτώσεις τα VAAs ευθύνονταν για την αύξηση της συμμετοχής των πολιτών στις εκλογές [92, 39].

Οι προηγούμενες ενότητες έδειξαν πως η επιπλέον πληροφορία που δίνεται για το χρήστη, όπως είναι για παράδειγμα τα δημογραφικά του στοιχεία ή το ενδιαφέρον του για την πολιτική, βοηθά το VAA να τον γνωρίσει καλύτερα. Έτσι παρατηρήθηκε πως η απόδοση του συστήματος αυξάνεται όταν μαζί με τις δηλώσεις πολιτικής λαμβάνονται υπόψη και οι απαντήσεις του χρήστη σε κάποιες από τις συμπληρωματικές ερωτήσεις (βλέπε Ενότητα 2.2). Στην παρούσα Ενότητα εξετάζεται κατά πόσο η ηλικία, το φύλο και το μορφωτικό επίπεδο επηρεάζουν και το ενδιαφέρον για την πολιτική. Επιπρόσθετα διερευνάται η σχέση μεταξύ του ενδιαφέροντος για την πολιτική και του λόγου για τον οποίο ο χρήστης επιλέγει να υποστηρίξει ένα κόμμα.

Η μεγάλη πλειοψηφία των χρηστών (κατά μέσο όρο για κάθε χώρα πάνω από το 70%) απάντησε πως ο λόγος που ψηφίζει ένα κόμμα είναι η ιδεολογία του. Οι υπόλοιποι χρήστες εξέφρασαν πως υποστηρίζουν ένα κόμμα για διάφορους λόγους, όπως για παράδειγμα εξαιτίας της οικογένειας και των φίλων τους. Αυτοί οι χρήστες, πιθανόν να μην ενδιαφέρονται τόσο για την πολιτική, όσο αυτοί που δήλωσαν πως ψηφίζουν ένα κόμμα λόγω της ιδεολογίας του. Οι χρήστες που γνωρίζουν την ιδεολογία των κομμάτων, αναμένεται να ασχολούνται με την πολιτική πιο πολύ από αυτούς που επιλέγουν ένα κόμμα επειδή το υποστηρίζει κάποιος άλλος.

Για τη διερεύνηση της σχέσης που έχουν τα δημογραφικά χαρακτηριστικά του χρήστη με το ενδιαφέρον του για την πολιτική καθώς και της σχέσης μεταξύ των κριτηρίων ψήφου και του ενδιαφέροντος

για την πολιτική, χρησιμοποιήθηκε η γραμμική παλινδρόμηση. Γραμμική παλινδρόμηση είναι η μέθοδος κατά την οποία η σχέση μεταξύ μιας εξαρτημένης μεταβλητής (στη συγκεκριμένη περίπτωση ως εξαρτημένη μεταβλητή ορίζεται το ενδιαφέρον του χρήστη για την πολιτική) και μιας η περισσότερων ανεξάρτητων μεταβλητών (ανεξάρτητες μεταβλητές είναι το φύλο, η ηλικία και τα κριτήρια ψήφου του χρήστη) εκφράζεται ως η γραμμική συνάρτηση που δείχνει το σταθμισμένο άθροισμα των μεταβλητών εισόδου και θορύβου [26]. Εξαρτημένη μεταβλητή είναι η έξοδος του συστήματος, δηλαδή το αποτέλεσμα που προκύπτει από τη διαχείριση κάποιων παραγόντων, οι οποίοι συνήθως αναφέρονται στις ανεξάρτητες μεταβλητές. Ανεξάρτητη μεταβλητή είναι η είσοδος του συστήματος που ορίζεται ως η υποτιθέμενη αιτία που προκαλεί τη μεταβολή στην κατάσταση της εξαρτημένης μεταβλητής.

3.2.1 Το ενδιαφέρον του χρήστη για την πολιτική, σύμφωνα με τα δημογραφικά του χαρακτηριστικά

Ο Πίνακας 3.1 παρουσιάζει την εξίσωση της γραμμικής παλινδρόμησης για κάθε χώρα ξεχωριστά, η οποία εκφράζει τη σχέση μεταξύ των δημογραφικών χαρακτηριστικών των χρηστών και του ενδιαφέροντος τους για την πολιτική. Το ενδιαφέρον για την πολιτική συμβολίζεται με τη μεταβλητή Q_{11} (εξαρτημένη μεταβλητή), ενώ τα δημογραφικά χαρακτηριστικά που αντιστοιχούν στις ανεξάρτητες μεταβλητές του προβλήματος συμβολίζονται με τις μεταβλητές Q_8 για το φύλο, Q_9 για την ηλικία και Q_{10} για το μορφωτικό επίπεδο. Η μεταβλητή Q_8 έχει την τιμή 0 όταν αναφέρεται στις γυναίκες και την τιμή 1 όταν πρόκειται για τους άντρες. Οι ηλικιακές ομάδες που η αντιπροσωπεύει η Q_9 αντιστοιχούν στις τιμές 1 (18-24 χρονών), 2 (25-34 χρονών), 3 (35-44 χρονών), 4 (45-54 χρονών), 5 (55-64 χρονών) και 6 (65 και άνω χρονών). Τα μορφωτικά επίπεδα της Q_{10} ορίζονται ως ακολούθως:

1. Δημοτικό
2. Γυμνάσιο
3. Λύκειο, Τεχνικό Λύκειο/Επαγγελματική σχολή
4. Πανεπιστήμιο, Μεταπτυχιακή εκπαίδευση

Τα αστεράκια στον Πίνακα 3.1 συμβολίζουν το στατιστικό επίπεδο σημαντικότητας (p -value) της κάθε μεταβλητής. Τα δύο αστεράκια (**) δηλώνουν πως το p -value είναι μικρότερο του 0.01 και η πιθανότητα η σχέση μεταξύ των μεταβλητών που βρίσκονται στο δείγμα να είναι “ψευδής”, είναι μικρότερη του 1%. Το ένα αστεράκι (*) αντιστοιχεί σε p -value μικρότερο του 0.05, ενώ όταν δεν υπάρχει κανένα αστεράκι τότε το p -value είναι μεγαλύτερο του 0.05. Το R^2 δείχνει το ποσοστό που εξηγείται από τις ανεξάρτητες μεταβλητές στη διακύμανση στο ενδιαφέρον για την πολιτική.

Σε όλες τις περιπτώσεις, εκτός από τη Ρουμανία, είναι τουλάχιστον 99% σίγουρο πως οι άντρες ενδιαφέρονται περισσότερο για την πολιτική από τις γυναίκες (μεταβλητή Q_8), με τις πλείστες από αυτές η διαφορά τους να κυμαίνεται γύρω στις 0.2-0.3 μονάδες. Εξαιρέση αποτελεί η Αυστρία, όπου οι άντρες φαίνεται να ενδιαφέρονται κατά 0.4 μονάδες περισσότερο από τις γυναίκες και η Βουλγαρία με τη Λετονία, όπου η διαφορά αυτή φτάνει τις 0.13 μονάδες. Όμως τόσο στη Λετονία όσο και στη Ρουμανία το $R^2 = 0.017$, πράγμα που δείχνει πως οι ανεξάρτητες μεταβλητές καταφέρνουν να εξηγήσουν μόλις το 1.7% της διακύμανσης στο ενδιαφέρον για την πολιτική που είναι και το πιο χαμηλό ποσοστό από όλες τις χώρες.

Όσον αφορά την ηλικία (μεταβλητή Q_9) και το μορφωτικό επίπεδο (μεταβλητή Q_{10}) παρατηρείται να επηρεάζουν διαφορετικά το ενδιαφέρον του χρήστη. Το μορφωτικό επίπεδο, σε όλες σχεδόν τις χώρες, εκτός της Λιθουανίας, αυξάνει το ενδιαφέρον του χρήστη για την πολιτική. Στη Λιθουανία φαίνεται

Πίνακας 3.1: Η εξίσωση της γραμμικής παλινδρόμησης για κάθε χώρα (η σχέση μεταξύ δημογραφικών χαρακτηριστικών και ενδιαφέροντος για την πολιτική)

A/A	Σύνολο Δεδομένων	Εξίσωση παλινδρόμησης	R ²
1	Αυστρία	$Q_{11} = 2.5 + 0.4 \cdot Q_8^{**} + 0.062 \cdot Q_9^{**} + 0.127 \cdot Q_{10}^{**}$	0.105
2	Βουλγαρία	$Q_{11} = 2.45 + 0.126 \cdot Q_8^{**} + 0.042 \cdot Q_9^{**} + 0.200 \cdot Q_{10}^{**}$	0.047
3	Κύπρος	$Q_{11} = 2.26 + 0.265 \cdot Q_8^{**} + 0.019 \cdot Q_9 + 0.223 \cdot Q_{10}^{**}$	0.042
4	Τσεχία	$Q_{11} = 2.5 + 0.245 \cdot Q_8^{**} + 0.028 \cdot Q_9^{**} + 0.103 \cdot Q_{10}^{**}$	0.031
5	Γερμανία	$Q_{11} = 2.6 + 0.314 \cdot Q_8^{**} + 0.051 \cdot Q_9^{**} + 0.144 \cdot Q_{10}^{**}$	0.053
6	Δανία	$Q_{11} = 2.27 + 0.22 \cdot Q_8^{**} + 0.048 \cdot Q_9^{**} + 0.163 \cdot Q_{10}^{**}$	0.050
7	Εσθονία	$Q_{11} = 2.58 + 0.183 \cdot Q_8^{**} + 0.072 \cdot Q_9^{**} + 0.067 \cdot Q_{10}^{**}$	0.042
8	ΗΒ-Αγγλία	$Q_{11} = 2.7 + 0.263 \cdot Q_8^{**} + 0.029 \cdot Q_9^{**} + 0.109 \cdot Q_{10}^{**}$	0.037
9	ΗΒ-Β. Ιρλανδία	$Q_{11} = 3.04 + 0.31 \cdot Q_8^{**} - 0.013 \cdot Q_9 + 0.036 \cdot Q_{10}$	0.039
10	ΗΒ-Σκωτία	$Q_{11} = 2.71 + 0.27 \cdot Q_8^{**} + 0.004 \cdot Q_9 + 0.128 \cdot Q_{10}^{**}$	0.047
11	ΗΒ-Ουαλία	$Q_{11} = 2.52 + 0.244 \cdot Q_8^{**} + 0.040 \cdot Q_9^{**} + 0.146 \cdot Q_{10}^{**}$	0.050
12	Ισπανία	$Q_{11} = 2.98 + 0.190 \cdot Q_8^{**} + 0.018 \cdot Q_9 + 0.108 \cdot Q_{10}^{**}$	0.030
13	Φινλανδία	$Q_{11} = 2.44 + 0.176 \cdot Q_8^{**} + 0.009 \cdot Q_9^{**} + 0.184 \cdot Q_{10}^{**}$	0.031
14	Γαλλία	$Q_{11} = 2.84 + 0.291 \cdot Q_8^{**} - 0.015 \cdot Q_9^{**} + 0.103 \cdot Q_{10}^{**}$	0.048
15	Ελλάδα	$Q_{11} = 2.23 + 0.195 \cdot Q_8^{**} + 0.053 \cdot Q_9^{**} + 0.210 \cdot Q_{10}^{**}$	0.043
16	Κροατία	$Q_{11} = 3.087 + 0.227 \cdot Q_8^{**} + 0.003 \cdot Q_9 + 0.043 \cdot Q_{10}^{**}$	0.028
17	Ουγγαρία	$Q_{11} = 2.45 + 0.311 \cdot Q_8^{**} + 0.007 \cdot Q_9 + 0.010 \cdot Q_{10}$	0.048
18	Ιρλανδία	$Q_{11} = 3.067 + 0.218 \cdot Q_8^{**} + 0.003 \cdot Q_9 + 0.078 \cdot Q_{10}^{**}$	0.025
19	Ιταλία	$Q_{11} = 2.62 + 0.232 \cdot Q_8^{**} + 0.037 \cdot Q_9^{**} + 0.122 \cdot Q_{10}^{**}$	0.041
20	Λιθουανία	$Q_{11} = 3.1 + 0.243 \cdot Q_8^{**} + 0.026 \cdot Q_9 - 0.012 \cdot Q_{10}$	0.035
21	Λετονία	$Q_{11} = 2.57 + 0.133 \cdot Q_8^{**} + 0.028 \cdot Q_9 + 0.068 \cdot Q_{10}$	0.017
22	Ολλανδία	$Q_{11} = 2.32 + 0.350 \cdot Q_8^{**} - 0.034 \cdot Q_9 + 0.198 \cdot Q_{10}^{**}$	0.062
23	Πολωνία	$Q_{11} = 2.7 + 0.342 \cdot Q_8^{**} - 0.012 \cdot Q_9^{**} + 0.098 \cdot Q_{10}^{**}$	0.043
24	Πορτογαλία	$Q_{11} = 2.61 + 0.243 \cdot Q_8^{**} + 0.039 \cdot Q_9^{**} + 0.126 \cdot Q_{10}^{**}$	0.043
25	Ρουμανία	$Q_{11} = 3.052 + 0.01 \cdot Q_8 + 0.052 \cdot Q_9^{**} + 0.071 \cdot Q_{10}^{**}$	0.017
26	Σουηδία	$Q_{11} = 2.43 + 0.261 \cdot Q_8^{**} + 0.012 \cdot Q_9 + 0.204 \cdot Q_{10}^{**}$	0.057
27	Σλοβακία	$Q_{11} = 2.46 + 0.283 \cdot Q_8^{**} + 0.048 \cdot Q_9^{**} + 0.084 \cdot Q_{10}^{**}$	0.039

να το μειώνει, η μεταβλητή όμως αυτή δεν παράγει στατιστικά σημαντικό συντελεστή, έτσι μπορεί να αφαιρεθεί από την εξίσωση. Στατιστικά σημαντική επίδραση δείχνει να μην έχει ούτε στη Βόρεια Ιρλανδία, Ουγγαρία και Λετονία. Για όλες τις υπόλοιπες χώρες, είναι τουλάχιστον 99% σίγουρο πως για κάθε μονάδα αύξησης στην κλίμακα του μορφωτικού επιπέδου, αυξάνεται το ενδιαφέρον για την πολιτική από 0.1 έως 0.2 μονάδες.

Η ηλικία σχεδόν για τις μισές χώρες δεν έχει σημαντική επίδραση στο ενδιαφέρον του χρήστη (βλέπε τα αποτελέσματα για την Κύπρο, Βόρεια Ιρλανδία, Σκωτία, Ισπανία, Κροατία, Ουγγαρία, Ιρλανδία, Λετονία, Ολλανδία και Σουηδία). Για τη Γαλλία και Πολωνία είναι τουλάχιστον 95% και 99% αντίστοιχα σίγουρο πως για κάθε μονάδα αύξησης στην ηλικιακή μονάδα του χρήστη, μειώνεται το ενδιαφέρον του για την πολιτική κατά 0.01 μονάδα. Είναι τουλάχιστον 95% σίγουρο πως το ενδιαφέρον των χρηστών της Λιθουανίας για την πολιτική αυξάνεται κατά 0.03 μονάδες για κάθε μονάδα αύξησης στην ηλικιακή ομάδα του χρήστη, ενώ για τις υπόλοιπες χώρες είναι τουλάχιστον 99% βέβαιο πως για κάθε μονάδα αύξησης στην ηλικιακή ομάδα του χρήστη, το ενδιαφέρον του για την πολιτική μεγαλώνει από 0.03 μέχρι 0.07 μονάδες. Εξαιρεση αποτελεί η Φινλανδία, όπου το ενδιαφέρον για την πολιτική αυξάνεται μόλις 0.009 μονάδες, δηλαδή σχεδόν καθόλου. Γενικά το μέγιστο που μπορεί να επηρεάσει η ηλικία του χρήστη το ενδιαφέρον του για την πολιτική, παρατηρείται στην περίπτωση της Εσθονίας όπου οι ηλικίες από 65 χρονών και άνω (που δίνουν δηλαδή την τιμή 6 στη μεταβλητή Q9), παρουσιάζουν αυξημένο

ενδιαφέρον για την πολιτική κατά 0.42 μονάδες (6 φορές την τιμή 0.07).

Άρα γενικά, οι άντρες χρήστες που έχουν υψηλή μόρφωση τείνουν να ενδιαφέρονται περισσότερο για την πολιτική. Κάτι που η ηλικία φαίνεται να μην επηρεάζει σε μεγάλο βαθμό, αφού τις πιο πολλές φορές η αύξηση που παρατηρείται στο ενδιαφέρον του χρήστη καθώς αυξάνεται η ηλικιακή του ομάδα, είναι πολύ μικρή.

3.2.2 Τι δηλώνουν τα κριτήρια ψήφου του χρήστη για το ενδιαφέρον του για την πολιτική

Ο Πίνακας 3.2 δείχνει την εξίσωση της γραμμικής παλινδρόμησης για κάθε χώρα ξεχωριστά, μέσα από την οποία μπορούν να εξαχθούν συμπεράσματα σχετικά με το εάν ο λόγος που επιλέγει ένας χρήστης να υποστηρίξει κάποιο συγκεκριμένο κόμμα, δηλώνει και το ενδιαφέρον του για την πολιτική. Το ενδιαφέρον για την πολιτική συμβολίζεται με τη μεταβλητή Q_{11} που είναι η εξαρτημένη μεταβλητή αυτού του προβλήματος, ενώ τα κριτήρια ψήφου αντιστοιχούν στη μεταβλητή Q_3 που είναι η ανεξάρτητη μεταβλητή και είναι ονομαστική με 5 κατηγορίες:

- Q_{31} : Το κόμμα είναι ικανότερο από τα υπόλοιπα
- Q_{32} : Η ιδεολογία του κόμματος είναι πιο κοντά στη δική μου
- Q_{33} : Το κόμμα βοηθάει ανθρώπους σαν εμένα
- Q_{34} : Η οικογένεια ή οι φίλοι μου υποστηρίζουν το κόμμα αυτό
- Q_{35} : Είμαι ευχαριστημένος με την ηγεσία του κόμματος

Η κατηγορία Q_{32} επιλέχθηκε κατά μέσον όρο από το 70% των χρηστών όλων των χωρών (βλέπε Ενότητα 5, Υποενότητα 5.1), έτσι αποφασίστηκε πως θα είναι η κατηγορία αναφοράς, αυτή δηλαδή που θα ισχύει όταν όλες οι άλλες κατηγορίες παίρνουν μηδενικές τιμές. Ουσιαστικά, η κάθε μια από τις κατηγορίες Q_{31} , Q_{33} , Q_{34} και Q_{35} παίρνει την τιμή 1 όταν η απάντηση του χρήστη για το λόγο που ψηφίζει ένα κόμμα είναι αυτή που κρύβεται πίσω από την αντίστοιχη κατηγορία, ενώ όλες μαζί παίρνουν την τιμή 0 όταν ο χρήστης δηλώνει πως ψηφίζει εξαιτίας της ιδεολογίας του κόμματος, δηλαδή δίνει την απάντηση της κατηγορίας Q_{32} .

Τα αστεράκια στον Πίνακα 3.2 που υπάρχουν ή όχι στις μεταβλητές, δηλώνουν το στατιστικό επίπεδο σημαντικότητας (p-value) της κάθε μεταβλητής. Τα δύο αστεράκια (**) δηλώνουν πως το p-value είναι μικρότερο του 0.01 και η πιθανότητα η σχέση μεταξύ των μεταβλητών που βρίσκονται στο δείγμα να είναι "ψευδής", είναι μικρότερη του 1%. Το ένα αστεράκι (*) αντιστοιχεί σε p-value μικρότερο του 0.05, ενώ όταν δεν υπάρχει κανένα αστεράκι τότε το p-value είναι μεγαλύτερο του 0.05. Το R^2 εξηγεί το ποσοστό στη διακύμανση του ενδιαφέροντος για την πολιτική που εξηγείται από τις ανεξάρτητες μεταβλητές.

Από τα αποτελέσματα παρατηρείται πως σε όλες τις χώρες, εκτός της Ολλανδίας και της Σλοβακίας, είναι τουλάχιστον 99% σίγουρο πως οι χρήστες που προτιμούν ένα κόμμα εξαιτίας της οικογένειας ή των φίλων τους (μεταβλητή Q_{34}) ενδιαφέρονται λιγότερο για την πολιτική, από αυτούς που δήλωσαν πως ψηφίζουν ένα κόμμα εξαιτίας της ιδεολογίας του. Για τη Σλοβακία, η συγκεκριμένη παρατήρηση είναι τουλάχιστον 95% σίγουρο πως συμβαίνει ενώ για την Ολλανδία, αν και οι χρήστες που ψηφίζουν σύμφωνα με τους φίλους ή την οικογένεια τους έχουν κατά 0.4 μονάδες λιγότερο ενδιαφέρον για την πολιτική από αυτούς που επιλέγουν ένα κόμμα ανάλογα με την ιδεολογία του, αυτό δεν είναι τουλάχιστον 95% σίγουρο πως συμβαίνει στον πληθυσμό. Ωστόσο, παρατηρείται ότι το $R^2=0.016$ για την Ολλανδία

Πίνακας 3.2: Η εξίσωση της γραμμικής παλινδρόμησης για κάθε χώρα (η σχέση μεταξύ κριτηρίων ψήφου και ενδιαφέροντος για την πολιτική)

A/A	Σύνολο Δεδομένων	Εξίσωση παλινδρόμησης	R ²
1	Αυστρία	$Q_{11}=3.4+0.112Q_{31}^{**}-0.230Q_{33}^{**}-0.893Q_{34}^{**}-0.084Q_{35}$	0.031
2	Βουλγαρία	$Q_{11}=3.4-0.019Q_{31}^{**}-0.152Q_{33}^{**}-0.339Q_{34}^{**}-0.293Q_{35}^{**}$	0.021
3	Κύπρος	$Q_{11}=3.4-0.145Q_{31}^{**}-0.309Q_{33}^{**}-0.891Q_{34}^{**}-0.259Q_{35}^{**}$	0.049
4	Τσεχία	$Q_{11}=3.2+0.026Q_{31}^{**}-0.210Q_{33}^{**}-0.267Q_{34}^{**}-0.112Q_{35}^{**}$	0.009
5	Γερμανία	$Q_{11}=3.5-0.071Q_{31}^{**}-0.455Q_{33}^{**}-0.813Q_{34}^{**}-0.135Q_{35}^{**}$	0.059
6	Δανία	$Q_{11}=3.1+0.032Q_{31}^{**}-0.378Q_{33}^{**}-0.811Q_{34}^{**}-0.230Q_{35}^{**}$	0.021
7	Εσθονία	$Q_{11}=3.2-0.072Q_{31}^{**}-0.271Q_{33}^{**}-0.359Q_{34}^{**}-0.227Q_{35}^{**}$	0.042
8	ΗΒ-Αγγλία	$Q_{11}=3.4-0.102Q_{31}^{**}-0.261Q_{33}^{**}-0.737Q_{34}^{**}-0.129Q_{35}^{**}$	0.031
9	ΗΒ-Βόρεια Ιρλανδία	$Q_{11}=3.4-0.045Q_{31}^{**}-0.166Q_{33}^{**}-0.495Q_{34}^{**}-0.198Q_{35}$	0.025
10	ΗΒ-Σκωτία	$Q_{11}=3.4-0.011Q_{31}^{**}-0.207Q_{33}^{**}-0.853Q_{34}^{**}-0.130Q_{35}$	0.029
11	ΗΒ-Ουαλία	$Q_{11}=3.4-0.153Q_{31}^{**}-0.355Q_{33}^{**}-0.650Q_{34}^{**}-0.147Q_{35}$	0.039
12	Ισπανία	$Q_{11}=3.5-0.028Q_{31}^{**}-0.139Q_{33}^{**}-0.749Q_{34}^{**}-0.215Q_{35}^{**}$	0.017
13	Φινλανδία	$Q_{11}=3.3+0.006Q_{31}^{**}-0.302Q_{33}^{**}-0.632Q_{34}^{**}-0.181Q_{35}$	0.033
14	Γαλλία	$Q_{11}=3.4+0.066Q_{31}^{**}-0.201Q_{33}^{**}-0.629Q_{34}^{**}-0.092Q_{35}$	0.016
15	Ελλάδα	$Q_{11}=3.4-0.111Q_{31}^{**}-0.254Q_{33}^{**}-0.839Q_{34}^{**}-0.218Q_{35}^{**}$	0.026
16	Κροατία	$Q_{11}=3.5+0.014Q_{31}^{**}-0.234Q_{33}^{**}-0.376Q_{34}^{**}-0.282Q_{35}^{**}$	0.029
17	Ουγγαρία	$Q_{11}=3.3+0.368Q_{31}^{**}-0.257Q_{33}^{**}-0.182Q_{34}^{**}-0.102Q_{35}^{**}$	0.048
18	Ιρλανδία	$Q_{11}=3.6-0.079Q_{31}^{**}-0.192Q_{33}^{**}-0.507Q_{34}^{**}-0.150Q_{35}$	0.033
19	Ιταλία	$Q_{11}=3.4-0.020Q_{31}^{**}-0.289Q_{33}^{**}-0.811Q_{34}^{**}-0.168Q_{35}^{**}$	0.024
20	Λιθουανία	$Q_{11}=3.2+0.089Q_{31}^{**}-0.079Q_{33}^{**}-0.469Q_{34}^{**}-0.260Q_{35}^{**}$	0.026
21	Λετονία	$Q_{11}=3.0-0.049Q_{31}^{**}-0.018Q_{33}^{**}-0.496Q_{34}^{**}-0.191Q_{35}^{**}$	0.026
22	Ολλανδία	$Q_{11}=3.3-0.085Q_{31}^{**}-0.449Q_{33}^{**}-0.449Q_{34}^{**}-0.223Q_{35}$	0.016
23	Πολωνία	$Q_{11}=3.3-0.040Q_{31}^{**}-0.246Q_{33}^{**}-0.944Q_{34}^{**}-0.216Q_{35}^{**}$	0.037
24	Πορτογαλία	$Q_{11}=3.4+0.020Q_{31}^{**}-0.200Q_{33}^{**}-0.717Q_{34}^{**}-0.176Q_{35}^{**}$	0.031
25	Ρουμανία	$Q_{11}=3.5-0.078Q_{31}^{**}-0.235Q_{33}^{**}-0.580Q_{34}^{**}-0.199Q_{35}^{**}$	0.029
26	Σουηδία	$Q_{11}=3.4+0.015Q_{31}^{**}-0.316Q_{33}^{**}-0.585Q_{34}^{**}-0.151Q_{35}$	0.011
27	Σλοβακία	$Q_{11}=3.1+0.144Q_{31}^{**}+0.247Q_{33}^{**}-0.226Q_{34}^{**}-0.048Q_{35}$	0.010

και R²=0.010 για τη Σλοβακία, που σημαίνει πως τα κριτήρια ψήφου εξηγούν μόνο το 1.6% και 1% αντίστοιχα της διακύμανσης στο ενδιαφέρον για την πολιτική των χρηστών αυτών των δύο χωρών.

Όσον αφορά τις υπόλοιπες χώρες, που η πιο πάνω παρατήρηση είναι τουλάχιστον 99% σίγουρο πως συμβαίνει στον πληθυσμό, τα αποτελέσματα αναλύονται ως ακολούθως: Οι χρήστες της Αυστρίας, Κύπρου και Πολωνίας που ψηφίζουν σύμφωνα με την οικογένεια ή τους φίλους τους ενδιαφέρονται κατά 0.9 μονάδες λιγότερο από αυτούς που επηρεάζονται από την ιδεολογία του κόμματος. Οι χρήστες της Γερμανίας, Δανίας, Σκωτίας, Ελλάδας και Ιταλίας που δήλωσαν πως υποστηρίζουν το ίδιο κόμμα με τους φίλους ή την οικογένεια τους ενδιαφέρονται για την πολιτική κατά 0.8 μονάδες λιγότερο, αυτοί της Αγγλίας, Ισπανίας και Πορτογαλίας κατά 0.7 μονάδες λιγότερο και αυτοί της Ουαλίας, Φινλανδίας, Γαλλίας, Ρουμανίας και Σουηδίας κατά 0.6 μονάδες λιγότερο. Όλοι οι χρήστες των άλλων χωρών ενδιαφέρονται για την πολιτική κάτω από 0.5 μονάδες λιγότερο. Συγκεκριμένα, σε όλες τις περιπτώσεις το ενδιαφέρον για την πολιτική όσων υποστήριξαν ένα κόμμα για την ιδεολογία του, είναι κατά μέσον όρο από αρκετό (αντίστοιχη τιμή 3) έως πολύ (αντίστοιχη τιμή 4), κυμαίνεται δηλαδή στις τιμές 3.0-3.6. Υπό αυτές τις περιπτώσεις, το ενδιαφέρον αυτών που ψηφίζουν σύμφωνα με την οικογένεια και τους φίλους τους διαμορφώνεται ως εξής:

- χρήστες Αυστρίας, Κύπρου και Πολωνίας: από ελάχιστο (αντίστοιχη τιμή 2) έως αρκετό (αντί-

στοιχη τιμή 3), τιμή μεταξύ 2.3-2.5 και $R^2=0.031$, $R^2=0.049$ και $R^2=0.037$ αντίστοιχα.

- χρήστες Γερμανίας, Δανίας, Σκωτίας, Ελλάδας και Ιταλίας: από ελάχιστο (αντίστοιχη τιμή 2) έως αρκετό (αντίστοιχη τιμή 3), τιμή μεταξύ 2.3-2.7 και $R^2=0.059$, $R^2=0.021$, $R^2=0.029$, $R^2=0.026$ και $R^2=0.024$ αντίστοιχα.
- χρήστες Αγγλίας, Ισπανίας και Πορτογαλίας: από ελάχιστο (αντίστοιχη τιμή 2) έως αρκετό (αντίστοιχη τιμή 3), τιμή μεταξύ 2.3-2.7 και $R^2=0.031$, $R^2=0.017$ και $R^2=0.031$ αντίστοιχα.
- χρήστες Ουαλίας, Φινλανδίας, Γαλλίας, Ρουμανίας και Σουηδίας: από ελάχιστο (αντίστοιχη τιμή 2) έως αρκετό (αντίστοιχη τιμή 3), τιμή γύρω στο 2.7 και $R^2=0.039$, $R^2=0.033$, $R^2=0.016$, $R^2=0.029$ και $R^2=0.011$ αντίστοιχα.
- χρήστες Λιθουανίας και Λετονίας: από ελάχιστο (αντίστοιχη τιμή 2) έως αρκετό (αντίστοιχη τιμή 3), τιμή μεταξύ στο 2.5 και 2.7 και $R^2=0.026$ και στις δύο.
- χρήστες Βουλγαρίας, Τσεχίας, Εσθονίας, Βόρειας Ιρλανδίας, Κροατίας, Ουγγαρίας και Ιρλανδίας: αρκετό (αντίστοιχη τιμή 3), τιμή γύρω στο 3 και $R^2=0.021$, $R^2=0.009$, $R^2=0.042$, $R^2=0.025$, $R^2=0.029$, $R^2=0.048$ και $R^2=0.033$ αντίστοιχα.

Όσον αφορά τους χρήστες που υποστηρίζουν το κόμμα, το οποίο βοηθάει ανθρώπους σαν αυτούς (μεταβλητή Q_{33}), είναι τουλάχιστον 99% σίγουρο σχεδόν σε όλες τις χώρες, πως ενδιαφέρονται κατά μέσον όρο αρκετά για την πολιτική (τιμή γύρω στο 3). Στην Κύπρο, Κροατία και Ολλανδία, η συγκεκριμένη παρατήρηση είναι τουλάχιστον 95% βέβαιο ότι συμβαίνει στον πληθυσμό. Για τη Βόρεια Ιρλανδία, Λιθουανία, Λετονία και Σλοβακία δεν έχει εξασφαλιστεί ικανοποιητική στατιστική σημαντικότητα, αλλά αυτό δεν αποδεικνύει κάτι μιας και σε αυτές τις χώρες οι λόγοι για τους οποίους επιλέγει ο χρήστης το κόμμα που θα ψηφίσει δεν εξηγούν περισσότερο από το 2.6% της διακύμανσης στο ενδιαφέρον για την πολιτική.

Οι χρήστες που επιλέγουν να ψηφίσουν ένα κόμμα γιατί είναι ευχαριστημένοι με την ηγεσία του κόμματος (μεταβλητή Q_{35}) ή επειδή το κόμμα αυτό είναι ικανότερο από τα υπόλοιπα (μεταβλητή Q_{31}) ενδιαφέρονται από αρκετά έως πολύ για την πολιτική στις πλείστες περιπτώσεις, χωρίς να είναι πάντα τουλάχιστον 95% σίγουρο πως αυτό συμβαίνει στον πληθυσμό.

Επίσης το σχετικά χαμηλό R^2 στα αποτελέσματα αυτής και της προηγούμενης υποενοότητας, είναι δικαιολογημένο αφού το δείγμα που επεξεργαστήκαμε για κάθε χώρα ήταν αρκετά μεγάλο. Επιπρόσθετα το ενδιαφέρον του χρήστη για την πολιτική επηρεάζεται και από πολλούς άλλους παράγοντες, όχι μόνο αυτούς που εξετάσαμε στις συγκεκριμένες υποενοότητες, που αν τους προσθέσουμε στην εξίσωση μας θα αυξηθεί ταυτόχρονα και το R^2 .

3.3 Το πρόβλημα σύστασης στα VAAs ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων

Οι δηλώσεις πολιτικής στο ερωτηματολόγιο του EUnox χωρίζονται σε τρεις θεματικές κατηγορίες:

- Ευρωπαϊκή Ένωση
- Οικονομία
- Πολιτική - Κουλτούρα

Οι δηλώσεις πολιτικής μπορούν να ομαδοποιηθούν κάτω από κάθε κατηγορία, ανάλογα με το θέμα στο οποίο αναφέρονται, και να αθροιστούν, έτσι ώστε οι απαντήσεις που έδωσε ο κάθε χρήστης να αντικατασταθούν με τρεις συνολικές βαθμολογίες (μια για κάθε κατηγορία) οι οποίες θα χρησιμοποιούνται αντί των απαντήσεων για την πρόβλεψη ψήφου σε μια μεθοδολογία ανάλογη με αυτή που είδαμε στην Ενότητα 1.2 για την ταύτιση χαμηλών διαστάσεων (βλέπε Ενότητα 1.5.3). Υπό αυτές τις συνθήκες το πρόβλημα σύστασης στο VAA μπορεί να θεωρηθεί ως ένα πρόβλημα ΣΣ πολλαπλών κριτηρίων, όπου η κάθε κατηγορία θα αντιμετωπίζεται ως κριτήριο.

Τα ΣΣ πολλαπλών κριτηρίων μπορούν να χρησιμοποιήσουν τόσο τις βαθμολογίες των κριτηρίων που αφορούν ένα προϊόν (τις βαθμολογίες δηλαδή που έδωσε ο χρήστης σε κριτήρια που αφορούν συγκεκριμένα χαρακτηριστικά του προϊόντος ή της υπηρεσίας, όπως για παράδειγμα το είδος μιας ταινίας, το χρώμα ενός παντελονιού, το συγγραφέα ενός βιβλίου) όσο και τη συνολική βαθμολογία (που δίνεται ως η συνολική εντύπωση που σχημάτισε ο χρήστης για το συγκεκριμένο προϊόν ή υπηρεσία) για να προχωρήσουν σε συστάσεις, εφαρμόζοντας ανάλυση αποφάσεων πολλαπλών κριτηρίων. Η ανάλυση αποφάσεων πολλαπλών κριτηρίων (*Multi-Criteria Decision Analysis*) είναι ένα καλά εδραιωμένο πεδίο στον τομέα της Επιστήμης Αποφάσεων, που θα μπορούσε να χρησιμοποιηθεί για την υποστήριξη των χρηστών στη διαδικασία λήψης αποφάσεων, αναλύοντας τις επιλογές τους και μοντελοποιώντας τον τρόπο με τον οποίο αξιολογούν τα διάφορα κριτήρια για να καταλήξουν στις συγκεκριμένες επιλογές [93]. Προσεγγίζοντας το συγκεκριμένο πρόβλημα με τη χρήση του συνεργατικού φιλτραρίσματος η ομοιότητα μεταξύ των χρηστών μπορεί να υπολογιστεί είτε συγκρίνοντας τις βαθμολογίες (συνολική βαθμολογία και βαθμολογία σε κάθε κριτήριο) που έδωσαν σε ίδια προϊόντα οι χρήστες ή συγκρίνοντας τα μοντέλα των χρηστών μεταξύ τους, τα οποία δείχνουν τον τρόπο με τον οποίο αξιολογεί τα κριτήρια ο κάθε χρήστης για να καταλήξει στις επιλογές του.

Οι Agathokleous και Tsaratsoulis [94] προτείνουν τη χρήση νευρωνικών δικτύων για τη μοντελοποίηση του τρόπου με τον οποίο οι χρήστες συνδυάζουν τα διάφορα κριτήρια για να καταλήξουν στη συνολική τους βαθμολογία, υποθέτοντας πως η διαδικασία λήψης αποφάσεων είναι πολύπλοκη από ότι παρουσιάζεται στην υπάρχουσα βιβλιογραφία. Τα αποτελέσματα τους έδειξαν πως η απόδοση των μη γραμμικών μοντέλων ήταν καλύτερη από των γραμμικών, όμως η διαφορά μεταξύ τους ήταν πολύ μικρή. Επίσης η εύρεση ομοιότητας μεταξύ των χρηστών συγκρίνοντας τα μοντέλα τους, αντί να γίνεται σύγκριση μεταξύ των βαθμολογιών που έδωσαν στα ίδια προϊόντα (όπως γίνεται παραδοσιακά), φαίνεται να λειτουργεί αποδεικνύοντας πως η μοντελοποίηση των χρηστών είναι ένα αξιόλογο εργαλείο για την αύξηση της αποδοτικότητας των ΣΣ πολλαπλών κριτηρίων.

Όταν το πρόβλημα σύστασης στο VAA αντιμετωπίζεται ως ένα πρόβλημα ΣΣ πολλαπλών κριτηρίων, τότε οι ερωτήσεις κάτω από κάθε κατηγορία μπορούν να αθροιστούν και η συνολική τους βαθμολογία να θεωρηθεί ως η βαθμολογία του κάθε κριτηρίου. Έτσι δημιουργούνται τρεις συνολικές βαθμολογίες, τις οποίες το σύστημα ερμηνεύει ως τα κριτήρια που οδήγησαν το χρήστη να καταλήξει στο κόμμα που

θα δώσει την ψήφο εμπιστοσύνης του. Υπό αυτές τις συνθήκες, για να προβλέψει το σύστημα την πρόθεση ψήφου του χρήστη λαμβάνοντα ως είσοδο μόνο τις βαθμολογίες των τριών κριτηρίων μπορούν να χρησιμοποιηθούν τεχνικές που αναφέρονται από τους Agathokleous και Tsaratsoulis [94]. Η μέθοδος SVM, όπως φάνηκε στην Ενότητα 2.1, δίνει τα καλύτερα αποτελέσματα, έτσι αποφασίστηκε να υιοθετηθεί αυτή η μέθοδος για τα πειράματα. Η παραγοντοποίηση πινάκων δεν μπορεί να εφαρμοστεί, λόγω της μεγάλης ανισότητας μεταξύ των γνωστών τιμών και αυτών που καλείται το σύστημα να προβλέψει, που στην προκειμένη περίπτωση το σύστημα προβλέπει μόνο την πρόθεση ψήφου του χρήστη.

3.3.1 Η απόδοση του VAA με την ομαδοποίηση των δηλώσεων πολιτικής του VAA

Για να αντιμετωπιστεί το πρόβλημα σύστασης του VAA ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων, λήφθηκαν υπόψη οι τρεις θεματικές ενότητες, στις οποίες ανήκαν οι δηλώσεις πολιτικής. Έτσι αντί να παίρνει ως είσοδο το σύστημα τις απαντήσεις του κάθε χρήστη, με αυτή τη μέθοδο οι είσοδοι γίνονται τρεις που αντιστοιχούν στη συνολική βαθμολογία που προέκυψε από τις απαντήσεις του χρήστη στις δηλώσεις πολιτικής που ανήκουν σε ένα συγκεκριμένο θέμα. Η κάθε θεματική ενότητα με αυτόν τον τρόπο θεωρήθηκε ως ένα κριτήριο και ουσιαστικά το σύστημα κλήθηκε να προβλέψει την πρόθεση ψήφου του χρήστη μέσα από τις συνολικές βαθμολογίες σε κάθε κριτήριο. Στον Πίνακα 3.3 αναφέρεται σε ποια θεματική ενότητα βρίσκεται η κάθε ερώτηση.

Για να υπολογιστεί η συνολική βαθμολογία σε κάθε κριτήριο, αρχικά οι απαντήσεις των χρηστών που έπαιρναν τις τιμές $\mathcal{L} = \{1, 2, 3, 4, 5, 6\}$, αντικαταστάθηκαν με τις τιμές $\mathcal{L} = \{-2, -1, 0, 1, 2\}$, όπου με το πρόσημο $-$ δηλώνεται η αρνητική άποψη του χρήστη, με $+$ η θετική και με 0 η ουδέτερη του στάση (με 0 αντικαταστάθηκαν οι τιμές 3 και 6). Στη συνέχεια οι απαντήσεις της κάθε κατηγορίας πολλαπλασιάστηκαν με τον αριθμό της δήλωσης πολιτικής που αντιπροσώπευαν (εδώ ως ο αριθμός της δήλωσης πολιτικής δηλώθηκε η σειρά με την οποία εμφανίστηκε η κάθε ερώτηση μέσα σε κάθε κατηγορία, για παράδειγμα η ερώτηση 12 για τους χρήστες της Αυστρίας που εμφανίζεται πρώτη στην κατηγορία Θέματα Οικονομίας, αντιστοιχεί στον αριθμό 1) και αθροίστηκαν μεταξύ τους. Με αυτόν τον τρόπο δημιουργήθηκαν τρεις συνολικές βαθμολογίες που αντιστοιχούν σε τρία κριτήρια: το κριτήριο της Ευρωπαϊκής Ένωσης, το κριτήριο της Οικονομίας και το κριτήριο της Πολιτικής-Κουλτούρας. Από το κάθε κριτήριο αφαιρέθηκε η μέγιστη του τιμή και το αποτέλεσμα διαιρέθηκε με τη διαφορά μεταξύ της μέγιστης και ελάχιστης τιμής, έτσι ώστε η συνολική βαθμολογία για κάθε κριτήριο να κυμαίνεται στο διάστημα $[0, 1]$.

Στον Πίνακα 3.4 παρουσιάζονται τα αποτελέσματα για την απόδοση της πρόβλεψης ψήφου χρησιμοποιώντας την *libsvm*, όταν ως είσοδο το σύστημα πήρε μόνο τα τρία κριτήρια που δημιουργήθηκαν για κάθε χρήστη με την πιο πάνω μέθοδο (αναφερόμενη ως Μέθοδος Α εφεξής). Στην τελευταία στήλη βρίσκεται το F-measure για κάθε χώρα από τα αποτελέσματα της Ενότητας 2.1, όταν το σύστημα είχε ως είσοδο όλες τις απαντήσεις του κάθε χρήστη στις δηλώσεις πολιτικής. Παρατηρούμε πως με την ομαδοποίηση των ερωτήσεων και τη χρήση των συνολικών βαθμολογιών κάθε ομάδας για την πρόβλεψη ψήφου, μειώνονται οι διαστάσεις του προβλήματος αλλά αυξάνεται η διακύμανση των τιμών της εισόδου, με αποτέλεσμα να επηρεάζεται αρνητικά (μειώνεται) η απόδοση της σύστασης.

Καθώς η μεγάλη διακύμανση μεταξύ των τιμών της εισόδου φαίνεται να επηρεάζει αρνητικά την απόδοση του συστήματος, ακολουθήθηκε μια δεύτερη μέθοδος (αναφερόμενη ως Μέθοδος Β εφεξής) στην προσπάθεια να γίνει πιο μικρή η διακύμανση στις τιμές της εισόδου, η οποία περιγράφεται ως εξής:

- Αρχικά οι απαντήσεις των χρηστών με τις τιμές 5 και 4 αντικαταστάθηκαν με την τιμή -1 , δείχνοντας έτσι τη διαφωνία του χρήστη με τις συγκεκριμένες ερωτήσεις, όσες είχαν τιμές 3 και 6

Πίνακας 3.3: Η θεματική ενότητα για κάθε σύνολο ερωτήσεων

Σύνολο Δεδομένων	Θέματα Ευρωπαϊκής Ένωσης	Θέματα Οικονομίας	Θέματα Πολιτικής -Κουλτούρας
Αυστρία	1-11	12-20	21-30
Βουλγαρία	1-9	10-20	21-30
Κύπρος	1-8	9-19	20-30
Τσεχία	1-11	12-19	20-30
Γερμανία	1-12	13-21	22-30
Δανία	1-13	14-22	23-30
Εσθονία	1-9	10-19	20-30
ΗΒ-Αγγλία	1-10	11-20	21-30
ΗΒ-Βόρεια Ιρλανδία	1-8	9-16	17-30
ΗΒ-Σκωτία	1-9	10-19	20-30
ΗΒ-Ουαλία	1-9	10-19	20-30
Ισπανία	1-8	9-16	17-30
Φινλανδία	1-11	12-21	22-30
Γαλλία	1-10	11-20	21-30
Ελλάδα	1-9	10-20	21-30
Κροατία	1-8	9-20	21-30
Ουγγαρία	1-11	12-21	22-30
Ιρλανδία	1-9	10-22	23-30
Ιταλία	1-7	8-18	19-30
Λιθουανία	1-9	10-20	21-30
Λετονία	1-12	13-21	22-30
Ολλανδία	1-11	12-21	22-30
Πολωνία	1-8	9-16	17-30
Πορτογαλία	1-8	9-21	22-30
Ρουμανία	1-9	10-19	20-30
Σουηδία	1-10	11-20	21-30
Σλοβακία	1-12	13-21	22-30

πήραν την τιμή 0, παρουσιάζοντας την ουδέτερη στάση του χρήστη και με την τιμή 1 αντικαταστάθηκαν όσες απαντήσεις είχαν τιμές 1 και 2, αντιπροσωπεύοντας τους χρήστες που συμφωνούν με τις ερωτήσεις αυτές. Έτσι η τιμή που μπορούσε να δοθεί για κάθε ερώτηση, ήταν -1 , 0 ή 1. Αυτό κρίθηκε αναγκαίο καθώς ο διαχωρισμός μεταξύ του πόσο πολύ συμφωνεί ή διαφωνεί ο χρήστης σχετικά με ένα θέμα είναι συχνά δύσκολο να γίνει. Επιπλέον, οι χρήστες του VAA τείνουν να αποφεύγουν να παίρνουν “ακραίες” θέσεις στην κλίμακα Likert, με αποτέλεσμα οι διαφορές μεταξύ του “Συμφωνώ απόλυτα” και “Συμφωνώ” και του “Διαφωνώ απόλυτα” και “Διαφωνώ” να δημιουργούν θόρυβο [60].

- Ακολούθως χρησιμοποιήθηκε η μέθοδος k -means (βλέπε Ενότητα 3.1) για να χωριστούν οι ερωτήσεις κάθε θεματικής ενότητας σε τρεις επιμέρους ομάδες (οι μισές από το μικρότερο αριθμό ερωτήσεων που βρίσκονται σε θεματική ενότητα καθώς η κάθε θεματική ενότητα περιέχει τουλάχιστον 7 δηλώσεις πολιτικής - βλέπε Πίνακα 3.3, σύμφωνα με την ομοιότητα τους στον τρόπο που απαντήθηκαν από τους χρήστες. Έτσι δημιουργήθηκαν τρεις ομάδες για κάθε θεματική ενότητα, στις οποίες χωρίστηκαν οι δηλώσεις πολιτικής της κάθε ενότητας που απαντήθηκαν όμοια από τους χρήστες.
- Στη συνέχεια βρέθηκε η μέση ανταπόκριση του κάθε χρήστη στις δηλώσεις πολιτικής της κάθε ομά-

δας, η οποία πολλαπλασιάστηκε με τον αριθμό της αντίστοιχης ομάδας όπου βρισκόταν. Δηλαδή, η μέση ανταπόκριση του χρήστη στις δηλώσεις πολιτικής της πρώτης ομάδας πολλαπλασιάστηκε με τον αριθμό 1, η μέση του ανταπόκριση στις δηλώσεις πολιτικής της δεύτερης ομάδας πολλαπλασιάστηκε με 2 και με 3 πολλαπλασιάστηκε η μέση του ανταπόκριση στις δηλώσεις πολιτικής της τρίτης ομάδας. Έτσι δημιουργήθηκαν τρεις βαθμολογίες για κάθε ενότητα, οι οποίες αργότερα αθροίστηκαν αφήνοντας μια συνολική τιμή. Με τον τρόπο αυτό υπολογίστηκαν τρεις συνολικές τιμές, μια για κάθε θεματική ενότητα, που αντιπροσώπευαν τα τρία κριτήρια με τα οποία ο χρήστης κατέληξε στην πρόθεση ψήφου του. Οι τιμές που μπορούσε να πάρει το κάθε κριτήριο ήταν ένας ακέραιος αριθμός από το -6 μέχρι το 6.

Τα αποτελέσματα της μεθόδου B (βλέπε Πίνακα 3.5) είναι καλύτερα από αυτά της μεθόδου A (βλέπε Πίνακα 3.4) και αυτό οφείλεται πιθανόν στη μείωση της διακύμανσης των τιμών της εισόδου. Όμως πάλι η απόδοση της πρόβλεψης ψήφου ήταν χειρότερη από την απόδοση που παρουσίασε το σύστημα χρησιμοποιώντας ως είσοδο όλες τις δηλώσεις πολιτικής, αντί για τις αθροιστικές βαθμολογίες στις τρεις θεματικές ενότητες.

Πίνακας 3.4: Η απόδοση της πρόβλεψης ψήφου ομαδοποιώντας τις δηλώσεις πολιτικής σε κριτήρια εισόδου για κάθε χώρα ξεχωριστά – Μέθοδος A

Σύνολο Δεδομένων	Recall	Precision	F1	MAP	F1 χωρίς ομαδοποίηση
Αυστρία	0.4465	0.3339	0.3821	0.6503	0.5012
Βουλγαρία	0.4216	0.3883	0.4042	0.6429	0.4735
Κύπρος	0.5271	0.3744	0.4378	0.6812	0.6399
Τσεχία	0.3950	0.3665	0.3802	0.5829	0.5102
Γερμανία	0.3861	0.3564	0.3707	0.5851	0.5917
Δανία	0.4191	0.3943	0.4063	0.6162	0.5540
Εσθονία	0.3286	0.3118	0.3200	0.5641	0.4401
ΗΒ-Αγγλία	0.5104	0.4739	0.4915	0.7087	0.5821
ΗΒ-Βόρεια Ιρλανδία	0.2932	0.2914	0.2923	0.5203	0.5129
ΗΒ-Σκωτία	0.4419	0.393	0.4160	0.6574	0.5743
ΗΒ-Ουαλία	0.3932	0.3211	0.3535	0.6106	0.5008
Ισπανία	0.3111	0.2429	0.2728	0.5320	0.3739
Φινλανδία	0.3445	0.2701	0.3028	0.5682	0.6241
Γαλλία	0.3914	0.3592	0.3746	0.6007	0.5673
Ελλάδα	0.4807	0.4159	0.4459	0.6597	0.5508
Κροατία	0.4005	0.2953	0.3399	0.6270	0.5714
Ουγγαρία	0.4764	0.3982	0.4338	0.6732	0.6060
Ιρλανδία	0.2877	0.2366	0.2597	0.5060	0.3832
Ιταλία	0.4791	0.4278	0.4520	0.6707	0.5998
Λιθουανία	0.5619	0.4943	0.5260	0.7593	0.6046
Λετονία	0.6719	0.6420	0.6566	0.8207	0.7223
Ολλανδία	0.4513	0.3159	0.3717	0.6637	0.6832
Πολωνία	0.5394	0.4343	0.4812	0.6969	0.6260
Πορτογαλία	0.3677	0.2897	0.3240	0.5749	0.4325
Ρουμανία	0.3439	0.2859	0.3122	0.5781	0.4960
Σουηδία	0.2749	0.2578	0.2661	0.4886	0.4544
Σλοβακία	0.3432	0.2246	0.2716	0.5388	0.3993

Πίνακας 3.5: Η απόδοση της πρόβλεψης ψήφου ομαδοποιώντας τις δηλώσεις πολιτικής σε κριτήρια εισόδου για κάθε χώρα ξεχωριστά – Μέθοδος B

Σύνολο Δεδομένων	Recall	Precision	F1	MAP	F1 χωρίς ομαδοποίηση
Αυστρία	0.4465	0.3339	0.3821	0.6503	0.5012
Βουλγαρία	0.4216	0.3883	0.4042	0.6429	0.4735
Κύπρος	0.5271	0.3744	0.4378	0.6812	0.6399
Τσεχία	0.3950	0.3665	0.3802	0.5829	0.5102
Γερμανία	0.3861	0.3564	0.3707	0.5851	0.5917
Δανία	0.4191	0.3943	0.4063	0.6162	0.5540
Εσθονία	0.3286	0.3118	0.3200	0.5641	0.4401
ΗΒ-Αγγλία	0.5104	0.4739	0.4915	0.7087	0.5821
ΗΒ-Βόρεια Ιρλανδία	0.2932	0.2914	0.2923	0.5203	0.5129
ΗΒ-Σκωτία	0.4419	0.393	0.4160	0.6574	0.5743
ΗΒ-Ουαλία	0.3932	0.3211	0.3535	0.6106	0.5008
Ισπανία	0.3111	0.2429	0.2728	0.5320	0.3739
Φινλανδία	0.3445	0.2701	0.3028	0.5682	0.6241
Γαλλία	0.3914	0.3592	0.3746	0.6007	0.5673
Ελλάδα	0.4807	0.4159	0.4459	0.6597	0.5508
Κροατία	0.4005	0.2953	0.3399	0.6270	0.5714
Ουγγαρία	0.4764	0.3982	0.4338	0.6732	0.6060
Ιρλανδία	0.2877	0.2366	0.2597	0.5060	0.3832
Ιταλία	0.4791	0.4278	0.4520	0.6707	0.5998
Λιθουανία	0.5619	0.4943	0.5260	0.7593	0.6046
Λετονία	0.6719	0.6420	0.6566	0.8207	0.7223
Ολλανδία	0.4513	0.3159	0.3717	0.6637	0.6832
Πολωνία	0.5394	0.4343	0.4812	0.6969	0.6260
Πορτογαλία	0.3677	0.2897	0.3240	0.5749	0.4325
Ρουμανία	0.3439	0.2859	0.3122	0.5781	0.4960
Σουηδία	0.2749	0.2578	0.2661	0.4886	0.4544
Σλοβακία	0.3432	0.2246	0.2716	0.5388	0.3993

Κεφάλαιο 4

Συζήτηση

4.1 Σχολιασμός Αποτελεσμάτων

Σκοπός της εργασίας αυτής ήταν να διερευνηθεί το πώς συμπεριφέρεται ο Ευρωπαίος χρήστης στο περιβάλλον του ΕΥνοχ και να εξετάσει τρόπους για τη βελτιστοποίηση της απόδοσης της πρόβλεψης ψήφου που γίνεται από το σύστημα. Έτσι διεξήχθησαν συγκεκριμένα πειράματα που καταπιάστηκαν ουσιαστικά με δύο μεγάλα θέματα:

1. Πώς μπορούν να χρησιμοποιηθούν τεχνικές και μέθοδοι για τη βελτίωση της πρόβλεψης ψήφου.
2. Η συμπεριφορά του χρήστη στο VAA και πώς επηρεάζεται από τα ιδιαίτερα χαρακτηριστικά του.

Στην Ενότητα αυτή παρουσιάζονται και ερμηνεύονται συνολικά τα αποτελέσματα από τα πειράματα που διεξήχθησαν, η περιγραφή των οποίων βρίσκεται στις προηγούμενες ενότητες. Αρχικά γίνεται αναφορά στα ερευνητικά ερωτήματα που αναλύθηκαν σε κάθε ενότητα, ενώ παράλληλα σχολιάζεται κατά πόσο τα αποτελέσματα που προέκυψαν ήταν τα αναμενόμενα. Κατά το σχολιασμό συγκρίνονται τα ευρήματα της μελέτης με αυτά άλλων ερευνητών και γίνεται συζήτηση για τις διαφορές και τις ομοιότητες μεταξύ των αποτελεσμάτων και των άλλων ερευνών. Παράλληλα διαφαίνεται τι νέο και σημαντικό προκύπτει από την παρούσα έρευνα.

4.1.1 Η βελτιστοποίηση της πρόβλεψης ψήφου

Για τη βελτιστοποίηση της πρόβλεψης ψήφου εξετάστηκαν διάφορες τεχνικές και μέθοδοι, που χρησιμοποιούνται από τα Συστήματα Συστάσεων. Η απόδοση του συστήματος υπολογίστηκε από το κατά πόσο το σύστημα κατάφερε να προβλέψει σωστά την πρόθεση ψήφου του χρήστη, την οποία δήλωσε στη συμπληρωματική ερώτηση του VAA “Στις Ευρωεκλογές, ποιο κόμμα προτίθεστε να ψηφίσετε;”. Η αποτελεσματικότητα της κάθε τεχνικής μπορεί να μετρηθεί από την απόδοση της στο να προβλέψει την πρόθεση ψήφου του χρήστη, καθώς ο τρόπος αυτός ακολουθήθηκε και σε προηγούμενες έρευνες [95, 96, 57, 65, 81]. Επίσης οι Tsaratsoulis και Mendez [97] που ασχολήθηκαν με την κοινωνική προσέγγιση του VAA, μόνο που αντί για την πρόθεση ψήφου χρησιμοποίησαν την πιθανότητα οι χρήστες να ψηφίσουν ένα συγκεκριμένο κόμμα που συμμετείχε στις Γερμανικές εκλογές του 2013, κατέληξαν πως δεν υπάρχει σημαντικό κέρδος όταν αντί της πρόθεσης ψήφου χρησιμοποιείται η πιθανότητα ψήφου.

Τα πειράματα ξεκίνησαν με τη διερεύνηση του κατά πόσο η κοινωνική προσέγγιση των VAAs δίνει καλύτερα αποτελέσματα από την παραδοσιακή (βλέπε Ενότητα 2.1), κάτι που υποστηρίζεται ότι συμβαί-

νει στις υπάρχουσες έρευνες [95, 57]. Αρχικά το συνολικό δείγμα της κάθε χώρας χωρίστηκε σε σύνολο δοκιμής και σύνολο εκπαίδευσης, όπου το πρώτο χρησιμοποιήθηκε για την εκπαίδευση του προβλήματος και η απόδοση του συστήματος μετρήθηκε με το σύνολο δοκιμής. Οι χρήστες από το σύνολο εκπαίδευσης κατανεμήθηκαν σε ομάδες, σύμφωνα με την πρόθεση ψήφου τους. Όσοι υποστήριξαν το ίδιο κόμμα καταχωρήθηκαν στην ίδια ομάδα. Όταν το πρόβλημα του VAA προσεγγίστηκε κοινωνικά, χρησιμοποιήθηκαν διάφορες τεχνικές μηχανικής μάθησης για να μοντελοποιήσουν τις δηλώσεις πολιτικής των χρηστών που βρίσκονταν στην κάθε ομάδα. Ακολουθώς γινόταν σύγκριση μεταξύ των απαντήσεων που έδωσαν οι χρήστες στο σύνολο δοκιμής και των μοντέλων κόμματος, για να ταξινομηθεί ο κάθε χρήστης από το σύνολο δοκιμής στο κόμμα που ήταν πιο πιθανόν να ανήκει. Στο τέλος έγινε αντιστοίχιση του κόμματος στο οποίο ταξινομήθηκε ο χρήστης, με αυτό που δήλωσε πραγματικά ως πρόθεση ψήφου του και μετρήθηκε η απόδοση της κάθε τεχνικής από το πόσο σωστά πραγματοποίησε την κάθε ταξινόμηση.

Οι τεχνικές που υιοθετήθηκαν για την κοινωνική σύσταση ήταν οι Μηχανές Διανυσμάτων Υποστήριξης (SVM), τα Κρυμμένα Μοντέλα Markov (HMM), ο Naïve Bayes ταξινομητής, τα τεχνητά Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης, ο ταξινομητής ελάχιστης Mahalanobis απόστασης και ο αλγόριθμος των k -κοντινότερων γειτόνων (KNN). Για την παραδοσιακή προσέγγιση εφαρμόστηκε η συσχέτιση Pearson, που ουσιαστικά υπολόγιζε την ομοιότητα μεταξύ των απαντήσεων που έδωσαν οι χρήστες και των απαντήσεων των κομμάτων, για να καταλήξει στο κόμμα με το οποίο ταιριάζει καλύτερα ο χρήστης. Ακολουθώς το κόμμα αυτό παραλληλίστηκε με την πρόθεση ψήφου του χρήστη και η απόδοση της παραδοσιακής προσέγγισης μετρήθηκε από το κατά πόσο το κόμμα που βρήκε ότι ταιριάζει περισσότερο με το χρήστη είναι και αυτό που δήλωσε στην πρόθεση ψήφου του.

Όπως ήταν αναμενόμενο η κοινωνική σύσταση είχε μεγαλύτερη απόδοση από την παραδοσιακή. Αυτό που ήταν ενδιαφέρον ήταν να αποδειχθεί εάν αυτό ισχύει σε όλες τις χώρες, μιας και οι Tsapatsoulis et al. [81] σε μια προσπάθεια να παρέχουν κάποιες οδηγίες σχεδιασμού για τα SVAA παρατήρησαν ότι η απόδοση της σύστασης διαφέρει από σύνολο δεδομένων σε σύνολο δεδομένων. Πράγματι σε κάθε περίπτωση η αποτελεσματικότητα της σύστασης διέφερε από χώρα σε χώρα, με κάποιες τεχνικές να αποδίδουν καλύτερα από κάποιες άλλες. Παρόλα αυτά σε όλες τις περιπτώσεις, εκτός από μια χώρα, η μέθοδος SVM που εφαρμόστηκε στην κοινωνική σύσταση έδωσε τα καλύτερα αποτελέσματα. Η χώρα που αποτέλεσε εξαίρεση ήταν η Ολλανδία, η οποία έχει τα πιο λίγα δεδομένα εκπαίδευσης από όλες τις χώρες (μόλις 489). Υπό αυτές τις συνθήκες, η πιθανοτική μέθοδος Naïve Bayes είχε καλύτερη απόδοση από τη μέθοδο SVM, κάτι που είναι δικαιολογημένο αφού η SVM είναι πιο κατάλληλη για προβλήματα που έχουν πολλά δεδομένα [77], ενώ ο Naïve Bayes αποδίδει εξίσου καλά και με λίγα δεδομένα [98, 99].

Γενικά σε όλες τις περιπτώσεις, εκτός της Ολλανδίας, η μέθοδος SVM κατάφερε την καλύτερη επίδοση, κάτι που βρίσκεται σε συμφωνία με την μελέτη των Katakis et al. [57]. Στις πλείστες από αυτές τα νευρωνικά δίκτυα (που είναι επίσης τεχνική που εφαρμόστηκε για την κοινωνική σύσταση) είχαν τη δεύτερη μεγαλύτερη απόδοση, ενώ σε κάποιες χώρες τα αποτελέσματα τους ήταν σχεδόν ίδια με αυτά της μεθόδου SVM. Η συγκεκριμένη παρατήρηση συμπίπτει με το συμπέρασμα των Tsapatsoulis και Mendez [97], πως οι μη-γραμμικές τεχνικές μοντελοποίησης, όπως αυτές που βασίζονται τα νευρωνικά δίκτυα, ξεπερνούν τις γραμμικές μεθόδους του Mahalanobis. Αυτό δείχνει πως ο τρόπος που απαντούν στις ερωτήσεις οι χρήστες του VAA είναι πολύπλοκος, καθώς οι τεχνικές αυτές μπορούν να καταγράψουν πιο σύνθετες σχέσεις μεταξύ των χαρακτηριστικών εισόδου [26, 77, 24].

Η παρατήρηση για την πολυπλοκότητα στον τρόπο που απαντούν οι χρήστες επιβεβαιώνεται και από την αύξηση στην απόδοση των νευρωνικών δικτύων όταν ο αριθμός των κρυφών νευρώνων αυξανόταν, αφού όσο περισσότεροι νευρώνες χρησιμοποιούνται τόσο πιο σύνθετο είναι το πρόβλημα που παρουσιάζεται [100]. Το τι διαφέρει από χώρα σε χώρα είναι ο βαθμός πολυπλοκότητας, αφού κάποιες χρειάστηκαν περισσότερους και κάποιες λιγότερους νευρώνες για να επιτευχθεί η βέλτιστη απόδοση.

Όσον αφορά τις υπόλοιπες τεχνικές μηχανικής μάθησης που υιοθετήθηκαν για την κοινωνική σύσταση, υπήρχαν περιπτώσεις που η παραδοσιακή προσέγγιση ήταν καλύτερη ή ισοδύναμη με κάποιες από αυτές. Σχετικά με τις μεθόδους HMM και Mahalanobis, έρευνες έχουν δείξει πως επηρεάζονται από το μικρό αριθμό δεδομένων και από την πολυμορφία του δείγματος [96, 101, 70, 81]. Συγκεκριμένα, η απόδοση των HMMs εξαρτάται σε μεγάλο βαθμό από τη διαθεσιμότητα επαρκούς ποσότητας αντιπροσωπευτικών δεδομένων εκπαίδευσης [70] καθώς επίσης, η συμπεριφορά τους μοιάζει πολύ με αυτή του Mahalanobis ταξινομητή [96] και μειώνεται όταν το δείγμα χαρακτηρίζεται από πολυμορφία (οι χρήστες δηλαδή που υποστηρίζουν το ίδιο κόμμα απαντούν με διαφορετικό τρόπο τις δηλώσεις πολιτικής του VAA επειδή πιθανόν να προέρχονται από διαφορετικό πολιτικό και κοινωνικό υπόβαθρο [96, 101].

Στην Ενότητα 3.1 εξετάστηκε η ιδεολογική ταύτιση των χρηστών με το κόμμα που δήλωσαν πως θα ψηφίσουν στις επερχόμενες εκλογές και συνάμα παρατηρήθηκε κατά πόσο υπάρχει ποικιλομορφία στο δείγμα. Για να συμβεί αυτό οι χρήστες του κάθε κόμματος χωρίστηκαν σε επιμέρους ομάδες, σύμφωνα με την ομοιότητα στον τρόπο που απάντησαν στις δηλώσεις πολιτικής του ερωτηματολογίου. Στη συνέχεια εξετάστηκαν τα ιδεολογικά χαρακτηριστικά των κομμάτων βάσει των χρηστών που τα υποστηρίζουν.

Η πειραματική διαδικασία που εφαρμόστηκε απέδειξε πως όντως υπάρχει πολυμορφία στα κόμματα της κάθε χώρας. Επιπρόσθετα, διαφορετικά κόμματα μπορούν να παρουσιάσουν παρόμοια ιδεολογική ταύτιση. Σε αυτές τις περιπτώσεις, φανερώνεται είτε πως οι χρήστες είχαν παρόμοια χαρακτηριστικά γνωρίσματα και για αυτό ανταποκρίθηκαν όμοια στις δηλώσεις πολιτικής, είτε πως παρά τη διαφορετικότητα τους έδωσαν όμοιες απαντήσεις. Τα “όμοια” κόμματα μπερδεύουν το σύστημα, το οποίο δυσκολεύεται να αναγνωρίσει που πραγματικά ανήκει ο κάθε χρήστης και στην περίπτωση που εφαρμόζεται κυρίως η μέθοδος Mahalanobis, τείνει να τους ταξινομήσει στο κόμμα με το μεγαλύτερο μέγεθος. Ωστόσο όταν υπάρχει συσπείρωση μεταξύ των χρηστών, ακόμα και στα μικρά κόμματα το σύστημα μπορεί να δώσει πολύ ικανοποιητικά αποτελέσματα. Ενώ στην αντίθετη περίπτωση που οι χρήστες ενός μεγάλου σε μέγεθος κόμματος δεν παρουσιάζουν ιδεολογική ταύτιση, η επίδοση του συστήματος αναμένεται να μειωθεί σημαντικά.

Γενικά όταν υπάρχουν πολλά κόμματα, ο διαχωρισμός μεταξύ των θέσεων των κομμάτων γίνεται πιο δύσκολος. Έτσι στις χώρες με τα πολλά κόμματα η επίδοση του συστήματος τείνει να μειώνεται. Αυτό είναι ένα από τα προβλήματα που απασχολεί τους ερευνητές που ασχολούνται με τις μεθόδους υπολογισμού της ομοιότητας μεταξύ χρηστών και κομμάτων [60, 43, 46] και είναι ένας από τους λόγους που η κοινωνική σύσταση αποδίδει καλύτερα από την παραδοσιακή.

Στα δεδομένα της Σκωτίας οι μέθοδοι HMM και Mahalanobis της κοινωνικής σύστασης είχαν επίσης χαμηλότερο ποσοστό Recall από την παραδοσιακή σύσταση, καθώς το ποσοστό Precision της μεθόδου SVM, που είχε την καλύτερη απόδοση, δεν ήταν ιδιαίτερα πιο ψηλό από όλες τις μεθόδους που εφαρμόστηκαν (βλέπε Ενότητα 1.7 για το πώς υπολογίζεται το Recall και το Precision). Αυτό αποδεικνύει πως ο τρόπος που απάντησαν οι χρήστες της Σκωτίας δεν ήταν εύκολα διαχωρίσιμος από κόμμα σε κόμμα.

Οι μέθοδοι HMM και Mahalanobis παρουσίασαν χειρότερα αποτελέσματα από την παραδοσιακή προσέγγιση και στα κόμματα της Λιθουανίας και Πολωνίας. Κατάφεραν να ταξινομήσουν πιο λίγους χρήστες από αυτούς που ανήκαν πραγματικά στα κόμματα, παρουσιάζοντας πολύ χαμηλό ποσοστό Recall σε σχέση με τις άλλες μεθόδους. Αυτό δικαιολογείται, αφού τα αποτελέσματα στην ενότητα 3.1 δείχνουν πως τα δεδομένα αυτά χαρακτηρίζονται από ποικιλομορφία, κάτι που επηρεάζει αρνητικά τις συγκεκριμένες μεθόδους σύμφωνα με τις έρευνες που ήδη αναφέρθηκαν.

Χαμηλό Recall σε σχέση με την παραδοσιακή προσέγγιση παρουσιάστηκε στα κόμματα 2 και 3 της Λιθουανίας, τα οποία όπως φαίνεται υποστηρίζονται από χρήστες με όμοιες δηλώσεις πολιτικής (βλέπε αποτελέσματα Ενότητας 3.1). Έτσι, οι τεχνικές που χρησιμοποιούνται δεν μπορούν να διαχωρίσουν τους χρήστες που ανήκουν στα συγκεκριμένα κόμματα, έχοντας ως δεδομένο μόνο τις δηλώσεις πολιτικής. Το

ίδιο παρατηρείται σχεδόν σε όλα τα κόμματα της Πολωνίας, με το Mahalanobis ταξινομητή να επηρεάζεται ιδιαίτερα αρνητικά από τα τρία πιο μικρά κόμματα (3, 7 και 8), κάτι που συνέβη και στην έρευνα των Tsaratsoulis et al. [81].

Η έρευνα των Tsaratsoulis et al. [81] έδειξε πως ο ταξινομητής Mahalanobis χρειάζεται γύρω στους 600 χρήστες (δηλαδή ένα δείγμα 20 φορές μεγαλύτερο από τον αριθμό των 30 ερωτήσεων του VAA) για τη δημιουργία αποτελεσματικών μοντέλων κόμματος. Επιπρόσθετα δεν υπάρχει σημαντική διαφορά στην πρόβλεψη ψήφου εάν πρόκειται για τους πρώτους 600 χρήστες ή οι τους τελευταίους 600 του VAA για τη δημιουργία των μοντέλων. Έτσι όπως ήταν προσδοκώμενο, στο σύνολο δεδομένων της Ολλανδίας η συμπεριφορά του Mahalanobis ταξινομητή ήταν η χειρότερη από όλων των άλλων τεχνικών που εφαρμόστηκαν, ακόμα και από την παραδοσιακή προσέγγιση, μιας και οι χρήστες στο σύνολο εκπαίδευσης της Ολλανδίας είναι λιγότεροι συνολικά από 600 (είναι μόλις 489 για όλα τα κόμματα μαζί).

Το μικρό μέγεθος του δείγματος της Λετονίας, η οποία είναι η δεύτερη πιο μικρή χώρα, φαίνεται να μην επηρεάζει την απόδοση των HMMs και Mahalanobis στα δύο πρώτα κόμματα, σύμφωνα με τα αποτελέσματα της Ενότητας 3.1 και των Παραρτημάτων. Αυτό συμβαίνει μιας και οι χρήστες του πρώτου κόμματος είναι πολύ συσπειρωμένοι, πάνω από το 80% δίνει παρόμοιες απαντήσεις στις δηλώσεις πολιτικής σύμφωνα με το Διάγραμμα VII.11. Επίσης η καλή απόδοση στο κόμμα 2 σε σχέση με το κόμμα 3 (βλέπε II.21) οφείλεται στο γεγονός πως τα δύο αυτά κόμματα μοιράζονται χρήστες με την ίδια ιδεολογία, όμως λόγω του μεγέθους του κόμματος 2, που είναι πολύ πιο μεγάλο από το κόμμα 3 (βλέπε Παράρτημα III.7), οι ταξινομητές τοποθέτησαν τους περισσότερους χρήστες στο κόμμα 2 (βλέπε II.21). Επιπρόσθετα η χώρα αυτή είχε και τον πιο μικρό αριθμό κομμάτων (μόλις 3).

Έτσι επιβεβαιώνεται η παρατήρηση προηγούμενων ερευνών πως τα HMMs μπορούν να εκπαιδευτούν αποτελεσματικά ακόμα και με ένα πολύ μικρό δείγμα όταν δεν υπάρχει ποικιλομορφία στον τρόπο που απαντούν οι χρήστες που υποστηρίζουν το ίδιο κόμμα [96]. Επιπλέον ο Mahalanobis ταξινομητής αποδίδει καλύτερα όταν ο αριθμός των κομμάτων είναι μικρός [81] και όταν το δείγμα είναι εύκολα διαχωρίσιμο.

Σε αρκετές χώρες η παραδοσιακή σύσταση είχε παρόμοια απόδοση με τη μέθοδο των k -κοντινότερων γειτόνων (KNN). Αυτό συμφωνεί και με άλλες έρευνες που έχουν πραγματοποιηθεί [86, 96] και φανερώνει πως στις περιπτώσεις αυτές δεν υπάρχει μεγάλη διαφορά μεταξύ των χαρακτηριστικών γνωρισμάτων των χρηστών σε κάθε κόμμα, με αποτέλεσμα να απαντούν όλοι ή η μεγάλη πλειοψηφία με παρόμοιο τρόπο. Έτσι είτε χρησιμοποιείται ο KNN ταξινομητής, που λαμβάνει υπόψη του ένα μικρό αριθμό από το δείγμα, είτε εφαρμόζεται η παραδοσιακή προσέγγιση, που λαμβάνει υπόψη της όλο το δείγμα, τα αποτελέσματα είναι παρόμοια, μιας και το μικρό δείγμα αντιπροσωπεύει το μεγάλο.

Στην Ισπανία όλες οι τεχνικές απέδωσαν λιγότερο, με την παραδοσιακή προσέγγιση να εμφανίζει τη χειρότερη επίδοση. Αυτό οφείλεται στον πολύ μεγάλο βαθμό ανομοιογένειας του δείγματος εξαιτίας των διαφορετικών περιφερειών που έχει αλλά και εξαιτίας του μεγάλου αριθμού των κομμάτων της, αφού είχε τα περισσότερα κόμματα από όλες τις χώρες.

Αφού εξετάστηκαν οι διάφορες τεχνικές μηχανικής μάθησης και βρέθηκε πως καλύτερα αποδίδει η μέθοδος SVM για την πρόβλεψη της πρόθεσης ψήφου του χρήστη, τα πειράματα συνεχίστηκαν, έτσι ώστε να βρεθούν τρόποι που βελτιώνουν ακόμη περισσότερο την αποτελεσματικότητα της μεθόδου. Έρευνες έχουν δείξει πως ο τρόπος που ένας άνθρωπος καταλήγει στις διάφορες αποφάσεις του εξαρτάται από πολλούς παράγοντες, όπως την εμπειρία του παρελθόντος [102], τις γνωστικές προκαταλήψεις [103], καθώς και την ηλικία και άλλες δημογραφικές διαφορές [104]. Ερευνητές απέδειξαν πως οι προτάσεις που γίνονται μέσω των VAAs, επηρεάζουν κυρίως τους αναποφάσιστους, τις γυναίκες, τα άτομα κάτω των 34 χρόνων και τα άτομα που ψηφίζουν για πρώτη φορά [42, 39].

Έτσι αποφασίστηκε οι απαντήσεις του χρήστη στις δηλώσεις πολιτικής να συνδυαστούν με τις απαντήσεις του στις συμπληρωματικές ερωτήσεις που αφορούσαν το φύλο, την ηλικία, το μορφωτικό επίπεδο, το λόγο για τον οποίο ψηφίζει ένα κόμμα, το πόσο πολύ ενδιαφέρεται για την πολιτική και την αυτοτοποθέτηση του στον πολιτικό χάρτη και να επαναληφθεί η διαδικασία της κοινωνικής σύστασης με τη μέθοδο SVM (βλέπε Ενότητα 2.2). Τα αποτελέσματα, όπως ήταν αναμενόμενο βελτιώθηκαν με την επιπλέον πληροφορία που δόθηκε στο σύστημα. Όπως ήταν προσδοκώμενο ο συνδυασμός μεταξύ του ενδιαφέροντος για την πολιτική, των κριτηρίων ψήφου και των δημογραφικών χαρακτηριστικών, ήταν αυτά που βοήθησαν περισσότερο σε όλες τις χώρες.

Η αυτοτοποθέτηση του χρήστη στον πολιτικό χάρτη ενίσχυσε λιγότερο την απόδοση της πρόβλεψης ψήφου, σχεδόν σε όλες τις περιπτώσεις. Εξαιρεση αποτέλεσε ο συνδυασμός xy , ο οποίος έδειχνε την αυτοτοποθέτηση του χρήστη στην Οικονομική και Φιλελεύθερη-Συντηρητική κλίμακα και σε αρκετές περιπτώσεις απέδωσε καλύτερα από τους άλλους συνδυασμούς που έγιναν. Ωστόσο γενικά σε όλες οι συμπληρωματικές ερωτήσεις βοήθησαν πολύ στην πρόβλεψη ψήφου, αυξάνοντας κατά μέσο όρο πέντε μονάδες την απόδοση που είχε το VAA όταν λάμβανε υπόψη μόνο τις δηλώσεις πολιτικής του χρήστη.

Οι χρήστες, όμως, που απάντησαν στις συγκεκριμένες συμπληρωματικές ερωτήσεις ήταν λιγότεροι από το δείγμα που χρησιμοποιήθηκε. Έτσι υπολογίστηκαν οι απαντήσεις στις συμπληρωματικές ερωτήσεις για όσους δεν τις απάντησαν με τη μέθοδο παραγοντοποίησης πινάκων SVD, έτσι ώστε η διαδικασία να εφαρμοστεί σε ολοκληρωμένο δείγμα (βλέπε Ενότητα 2.3). Για την εκτίμηση των απαντήσεων, χρησιμοποιήθηκε ως είσοδο η ανταπόκριση του χρήστη στις δηλώσεις πολιτικής. Αφού η αυτοτοποθέτηση του χρήστη είναι κάτι υποκειμενικό για τον καθένα, έχει πολλές πιθανές τιμές και είναι πιο δύσκολο να εκτιμηθεί, αποφασίστηκε να γίνει πρόβλεψη μόνο για τις υπόλοιπες συμπληρωματικές ερωτήσεις. Η μέθοδος SVD ανταποκρίθηκε πολύ καλά στην εκτίμηση των αναπάντητων ερωτημάτων.

Αφού εκτιμήθηκαν οι απαντήσεις στις συμπληρωματικές ερωτήσεις του χρήστη, συμπληρώθηκε το δείγμα που έλειπε και η μέθοδος SVM εφαρμόστηκε τώρα σε ολοκληρωμένο δείγμα, παίρνοντας ως είσοδο τις απαντήσεις στις δηλώσεις πολιτικής και τις απαντήσεις στις συμπληρωματικές ερωτήσεις (μαζί με αυτές που εκτιμήθηκαν). Τα αποτελέσματα ήταν παρόμοια με αυτά της Ενότητας 2.2. Αυτό ήταν αναμενόμενο μιας και οι Tsaratsoulis et al. [81] έδειξαν πως γύρω στους 600 χρήστες είναι αρκετοί για τη δημιουργία αποτελεσματικών μοντέλων κόμματος. Έτσι, ακόμα και με μειωμένο δείγμα, όπως συνέβηκε στην περίπτωση που αφαιρέθηκαν όσοι χρήστες δεν απάντησαν στις συμπληρωματικές ερωτήσεις (βλέπε Ενότητα 2.2), τα αποτελέσματα δεν έπρεπε να είχαν μεγάλη διαφορά από αυτά με το ολοκληρωμένο δείγμα.

Ως εκ τούτου αποδείχτηκε πως με το να λαμβάνονται υπόψη οι συμπληρωματικές ερωτήσεις που απάντησε ο χρήστης μαζί με τις απαντήσεις του στις δηλώσεις πολιτικής, η απόδοση της κοινωνικής σύστασης αυξάνεται ακόμη περισσότερο. Όμως διαφέρουν από χώρα σε χώρα οι συμπληρωματικές ερωτήσεις που βοηθούν περισσότερο στην απόδοση του VAA. Σε κάθε περίπτωση όμως η συμπερίληψη του ενδιαφέροντος του χρήστη για την πολιτική καθώς και τα κριτήρια ψήφου του είναι επιλογές που βελτιώνουν την απόδοση της πρόβλεψης ψήφου και είναι καλό να λαμβάνονται υπόψη.

Ακόμη μια καινοτομία της παρούσας διατριβής, ήταν να προσπαθήσει να λύσει το πρόβλημα σύστασης του VAA ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων (βλέπε Ενότητα 3.3). Έτσι οι απαντήσεις στις δηλώσεις πολιτικής του ερωτηματολογίου, αθροίστηκαν κάτω από τρεις κατηγορίες (1- Θέματα Ευρωπαϊκής Ένωσης, 2- Θέματα Οικονομίας και 3- Θέματα Πολιτικής-Κουλτούρας) και έδωσαν τρεις συνολικές βαθμολογίες, τις οποίες το σύστημα λάμβανε ως είσοδο τις βαθμολογίες σε τρία κριτήρια. Ωστόσο, αν και οι διαστάσεις του προβλήματος μειώνονται με το να λαμβάνει ως είσοδο 3 τιμές αντί 30, η ομαδοποίηση των ερωτήσεων φαίνεται να μην λειτουργεί στην περίπτωση του VAA. Ίσως μάλιστα να πρέπει να αυξηθούν οι ερωτήσεις που χρησιμοποιούνται, μιας και η επιπλέον γνώση φαίνεται

πως βοηθά το σύστημα να γνωρίσει καλύτερα το χρήστη. Βέβαια αυτό πρέπει να γίνει με μέτρο [46].

4.1.2 Η ταυτότητα και συμπεριφορά του Ευρωπαίου χρήστη στο EUnox

Οι Ενότητες στις οποίες διερευνήσαμε άμεσα τη συμπεριφορά του Ευρωπαίου χρήστη στο περιβάλλον του EUnox ήταν οι Ενότητες 3.1 και 3.2. Στις άλλες Ενότητες φάνηκε μέσα από τα αποτελέσματα κατά την προσπάθεια για βελτιστοποίηση της πρόθεσης ψήφου του χρήστη, κατά πόσο επηρεάζεται από διάφορους παράγοντες και ιδιαίτερα από τα κριτήρια ψήφου και το ενδιαφέρον του για την πολιτική.

Σύμφωνα με έρευνες που έχουν γίνει, ο τυπικός χρήστης του VAA είναι άντρας, νέος, με υψηλή μόρφωση και ενδιαφέρεται ιδιαίτερα για την πολιτική (Garzia & Marschall, [56], pp. 99-101). Κάτι που επιβεβαιώνεται και στους χρήστες όλων των χωρών που χρησιμοποιήθηκαν για την παρούσα μελέτη. Τα ποσοστά όμως διαφέρουν από χώρα σε χώρα. Όσον αφορά το φύλο, στη Λετονία και Λιθουανία το 40% των χρηστών είναι γυναίκες και το υπόλοιπο 60% άντρες. Οι άντρες χρήστες καταλαμβάνουν γύρω στο 80% του δείγματος της Κύπρου, Τσεχίας, Γερμανίας, Ελλάδας, Ιρλανδίας, Ολλανδίας, Πολωνίας και Σλοβακίας, ενώ στις υπόλοιπες χώρες τα δείγματα περιλαμβάνουν περίπου 30% γυναίκες και 70% άντρες. Τα ποσοστά που καταλαμβάνουν στο δείγμα οι ηλικιακές ομάδες, επίσης ποικίλουν από χώρα σε χώρα, με τους χρήστες κάτω των 34 χρονών να αποτελούν γύρω στο 50% με 60% των πλειστων χωρών. Στη Δανία και Ρουμανία το ποσοστό αυτών των ηλικιών είναι γύρω στο 40%, στην Πολωνία και Σουηδία γύρω στο 70% και στη Λιθουανία γύρω στο 80%.

Το μορφωτικό επίπεδο της μεγάλης πλειοψηφίας των χρηστών για όλες τις χώρες, με εξαίρεση τη Λετονία, συγκεντρώνεται στους απόφοιτους λυκείου/τεχνικού λυκείου/ επαγγελματικής σχολής και στα άτομα με πτυχίο, μεταπτυχιακό ή και διδακτορικό. Στη Λετονία, το μορφωτικό επίπεδο του 69% των χρηστών περιορίζεται στους απόφοιτους λυκείου/τεχνικού λυκείου/επαγγελματικής σχολής, ενώ σχεδόν το υπόλοιπο ποσοστό (το 23%) απευθύνεται στους απόφοιτους γυμνασίου. Η Λετονία ήταν επίσης η μόνη χώρα που πάνω από το 20% των χρηστών της (συγκεκριμένα το 24%) δήλωσαν πως ενδιαφέρονται ελάχιστα για την πολιτική, ενώ συνολικά σε όλες τις χώρες οι χρήστες που εξέφρασαν πως δεν ενδιαφέρονται καθόλου για την πολιτική ήταν κάτω από το 2% του δείγματος. Έτσι, γενικά η συντριπτική πλειοψηφία των χρηστών σε όλες τις περιπτώσεις ενδιαφέρονται από πολύ μέχρι αρκετά για την πολιτική.

Η πλειοψηφία των χρηστών (κατά μέσο όρο πάνω από το 60%) σε όλες τις χώρες δήλωσε πως ψηφίζει το κόμμα που ταιριάζει περισσότερο με την ιδεολογία τους. Εξαίρεση για ακόμη μια φορά αποτέλεσε η Λετονία, όπου το 42% υποστήριξε πως ψηφίζει για ένα κόμμα βάσει της ιδεολογίας του και το 36% επειδή το κόμμα είναι ικανότερο από τα υπόλοιπα. Κάτι αντίστοιχο συμβαίνει και με τη Βουλγαρία με τα ποσοστά της να διαμορφώνονται 49% για την ιδεολογία, 29% για την ικανότητα του κόμματος και 12% για τη βοήθεια που προσφέρει το κόμμα σε ανθρώπους όπως το χρήστη.

Στην Ενότητα 3.2 εξετάσαμε κατά πόσο το ενδιαφέρον για την πολιτική του Ευρωπαίου χρήστη του VAA επηρεάζεται από τα δημογραφικά χαρακτηριστικά του και εάν μπορούν να εξαχθούν συμπεράσματα για το πόσο πολύ ενδιαφέρεται για την πολιτική ανάλογα με το λόγο για τον οποίο επέλεξε να ψηφίσει ένα κόμμα. Τα αποτελέσματα έδειξαν πως οι άντρες χρήστες που είναι μορφωμένοι τείνουν να ενδιαφέρονται περισσότερο για την πολιτική. Κάτι που η ηλικία φαίνεται να μην επηρεάζει σε μεγάλο βαθμό, αφού τις πιο πολλές φορές η αύξηση που παρατηρείται στο ενδιαφέρον του χρήστη καθώς αυξάνεται η ηλικιακή του ομάδα, είναι πολύ μικρή. Με τα ευρήματα αυτά συμφωνεί η έρευνα των van de Pol et al. ([105]), οι οποίοι έδειξαν ότι οι άνδρες και οι άνθρωποι με ανώτερο μορφωτικό επίπεδο είναι πιο πιθανό να ενδιαφέρονται πολύ για την πολιτική, ενώ οι νεαροί και οι γυναίκες είναι πιο αβέβαιοι όσον αφορά την ψήφο τους και έχουν μικρό πολιτικό ενδιαφέρον.

Επίσης όσοι χρήστες δήλωσαν προτίμηση για το κόμμα που υποστηρίζουν οι φίλοι ή η οικογένεια

τους, έδειξαν να ενδιαφέρονται από ελάχιστα μέχρι αρκετά για την πολιτική, όσοι απάντησαν πως υποστηρίζουν το κόμμα που βοηθάει ανθρώπους σαν αυτούς τείνουν να ενδιαφέρονται αρκετά για την πολιτική στις πλείστες περιπτώσεις, ενώ όσοι εξέφρασαν κάτι άλλο από αυτές τις δύο δηλώσεις φαίνεται να ενδιαφέρονται αρκετά έως πολύ για την πολιτική. Άρα, εάν ο πολίτης επιλέγει κάποιο κόμμα εξαιτίας ενός εξωτερικού παράγοντα και όχι των ιδεών, της ικανότητας του και της ηγεσίας του κόμματος, φαίνεται να ενδιαφέρεται από ελάχιστα έως αρκετά για την πολιτική.

4.2 Συμπεράσματα

Μέσα από την παρούσα διατριβή ερευνήθηκαν διάφορες τεχνικές από τα Συστήματα Συστάσεων για τη βελτιστοποίηση της πρόβλεψης ψήφου που γίνεται από τα VAAs. Ταυτόχρονα ερευνήθηκε η συμπεριφορά του Ευρωπαίου χρήστη και πώς διάφοροι παράγοντες επηρεάζουν τον τρόπο που ψηφίζει αλλά και το ενδιαφέρον του για την πολιτική.

Γενικά ο Ευρωπαίος χρήστης του VAA για την πλειοψηφία των χωρών φαίνεται να είναι άντρας, νέος, με υψηλή μόρφωση και ενδιαφέρεται πολύ για την πολιτική. Μάλιστα οι άντρες χρήστες που είναι μορφωμένοι τείνουν να ενδιαφέρονται περισσότερο για την πολιτική από ότι οι γυναίκες ή αυτοί με χαμηλό μορφωτικό επίπεδο. Η ηλικιακή ομάδα, έδειξε να μην αποτελεί ιδιαίτερα σημαντικό παράγοντα για το ενδιαφέρον του χρήστη, αφού τις πιο πολλές φορές η αύξηση που παρατηρήθηκε στο ενδιαφέρον του για την πολιτική καθώς αυξανόταν η ηλικιακή του ομάδα, ήταν πολύ μικρή. Επίσης όσοι δήλωσαν πως επιλέγουν το κόμμα που θα ψηφίσουν εξαιτίας των φίλων ή της οικογένειας τους ή λόγω της βοήθειας που προσφέρει στους συνανθρώπους τους, τείνουν να ενδιαφέρονται για την πολιτική από ελάχιστα ως αρκετά. Ενώ όσοι δήλωσαν πως ψηφίζουν για την ιδεολογία του, την ικανότητα και την ηγεσία του, ενδιαφέρονται για την πολιτική από αρκετά έως πολύ.

Τα πειράματα που διεξήχθησαν απέδειξαν για ακόμα μια φορά πως η κοινωνική προσέγγιση των VAAs αποδίδει καλύτερα από την παραδοσιακή, κάτι που υποστηρίζεται μέχρι τώρα από τις υπάρχουσες έρευνες. Επίσης σε όλες σχεδόν τις περιπτώσεις, οι Μηχανές Διανυσμάτων Υποστήριξης (SVMs) και τα Νευρωνικά Δίκτυα παρείχαν τα καλύτερα αποτελέσματα, δηλώνοντας πως υπάρχει πολυπλοκότητα στις δηλώσεις πολιτικής των χρηστών. Ο βαθμός πολυπλοκότητας διέφερε από χώρα σε χώρα.

Η αποτελεσματικότητα των μεθόδων που χρησιμοποιήθηκαν, ειδικά στις περιπτώσεις των Mahalanobis και HMMs, εξαρτήθηκε από το μέγεθος των δεδομένων της κάθε χώρας, τον αριθμό των κομμάτων που μοντελοποιήθηκαν και την πολυμορφία του δείγματος σε κάθε κόμμα. Τα χειρότερα αποτελέσματα παρουσιάστηκαν στην Ισπανία, που είχε με μεγάλη διαφορά τα πιο πολλά δεδομένα από όλες τις χώρες και τα πιο πολλά κόμματα σε αριθμό (12 κόμματα). Επιπρόσθετα στην Ισπανία τα κόμματα χαρακτηρίζονταν από ανομοιογένεια, καθώς επίσης πολλοί χρήστες έδιναν όμοιες απαντήσεις με χρήστες που υποστήριζαν άλλα κόμματα (βλέπε Διάγραμμα VI.11). Έτσι τα όρια μεταξύ των κομμάτων δεν ήταν ευδιάκριτα και η κατηγοριοποίηση των χρηστών δεν έγινε με μεγάλη επιτυχία.

Τα πειράματα που διεξήχθησαν έδειξαν πως οι χρήστες που υποστηρίζουν το ίδιο κόμμα, αν και θα αναμενόταν να απαντούν με τον ίδιο τρόπο στις δηλώσεις πολιτικής του ερωτηματολογίου, αυτό δε συμβαίνει πάντα προκαλώντας πολυμορφία στο δείγμα των κομμάτων. Παρόλα αυτά παρατηρήθηκε πως στις περισσότερες περιπτώσεις, η μεγάλη πλειοψηφία των χρηστών που υποστηρίζουν ένα κόμμα παρουσιάζουν ιδεολογική ταύτιση μαζί του. Επιπλέον εντοπίστηκαν πολλοί χρήστες που ανήκουν σε διαφορετικά κόμματα αλλά μοιράζονται την ίδια ιδεολογία. Κάτι τέτοιο οδηγεί το σύστημα σε λανθασμένες εκτιμήσεις, αφού δεν μπορεί να διαχωρίσει τους χρήστες βάσει των πολιτικών τους θέσεων. Σε αυτές τις περιπτώσεις, όταν το σύστημα έχει να επιλέξει μεταξύ δύο ή περισσότερων κομμάτων, τείνει να ταξινο-

μει τους χρήστες σε αυτό με το μεγαλύτερο μέγεθος, ειδικά εάν στο συγκεκριμένο υπάρχει συσπείρωση μεταξύ των χρηστών. Η συσπείρωση των χρηστών είναι σημαντική και στα μικρά κόμματα, αφού το σύστημα μπορεί να επιφέρει καλά αποτελέσματα σε τέτοιες συνθήκες, ειδικά εάν χρησιμοποιείται η μέθοδος Mahalanobis.

Ακόμα ένα από τα ευρήματα της παρούσας έρευνας είναι πως οι συμπληρωματικές ερωτήσεις βοηθούν το σύστημα να «γνωρίσει» καλύτερα το χρήστη και να προβεί σε πιο αποδοτικές προβλέψεις ψήφου. Έτσι, ακόμα και όταν ο χρήστης παραλείπει να απαντήσει στις συμπληρωματικές ερωτήσεις, τότε το σύστημα μπορεί να χρησιμοποιεί τις απαντήσεις του στις δηλώσεις πολιτικής για να προβλέψει την ανταπόκριση του στις συμπληρωματικές. Κάτι που όπως φάνηκε, μπορεί να κάνει πολύ καλά η μέθοδος SVD. Βέβαια, δε βοηθούν όλες οι συμπληρωματικές ερωτήσεις στον ίδιο βαθμό, αλλά κάποιες ενισχύουν περισσότερο τη γνώση για το χρήστη και κάποιες λιγότερο, ανάλογα από την κάθε χώρα ξεχωριστά.

Επιπρόσθετα εξετάστηκε κατά πόσο οι ερωτήσεις που καλείται ο χρήστης να απαντήσει στο VAA, μπορούν να ομαδοποιηθούν σύμφωνα με κάποια κριτήρια για να μειωθούν οι διαστάσεις του προβλήματος σύστασης. Με αυτόν τον τρόπο για πρώτη φορά το πρόβλημα σύστασης στα VAAs αντιμετωπίστηκε ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων που είναι ένα καλά εδραιωμένο πεδίο στον τομέα της Επιστήμης Αποφάσεων. Ωστόσο, η ομαδοποίηση των ερωτήσεων φαίνεται να μην λειτουργεί στην περίπτωση του VAA. Ίσως μάλιστα να πρέπει να αυξηθούν οι ερωτήσεις που χρησιμοποιούνται, μιας και η επιπλέον γνώση φαίνεται πως βοηθά το σύστημα να γνωρίσει καλύτερα το χρήστη.

4.3 Μελλοντική Έρευνα

Στην παρούσα μελέτη τα VAAs αντιμετωπίστηκαν υπό το πρίσμα των Συστήματα Συστάσεων. Δύο από τα μεγαλύτερα προβλήματα που αντιμετωπίζουν τα εν λόγω συστήματα είναι το πρόβλημα της ψυχρής εκκίνησης (cold-start problem) και τα αραιά δεδομένα (data sparsity) [4, 30]. Τα αραιά δεδομένα προκύπτουν όταν οι χρήστες βαθμολογούν ένα μικρό ποσοστό των διαθέσιμων στοιχείων δημιουργώντας μια ανισότητα μεταξύ των γνωστών και των άγνωστων τιμών, με τις μηδενικές τιμές να υπερτερούν. Το πρόβλημα της ψυχρής εκκίνησης παρουσιάζεται όταν οι χρήστες ή τα στοιχεία που προστέθηκαν πρόσφατα στο σύστημα δεν έχουν καθόλου αξιολογήσεις, με αποτέλεσμα το σύστημα να μην μπορεί να στηριχθεί στην υπάρχουσα πληροφορία για να προβεί σε συστάσεις.

Οι ελλείπουσες τιμές που παρουσιάζονται στα VAAs είναι πολύ λίγες [86] και έχει αποδειχθεί πως δεν επηρεάζουν σε αξιολογο βαθμό την απόδοση τους [95]. Έτσι δεν χρειάζεται να παρθούν οποιαδήποτε μέτρα. Στην περίπτωση που πιθανόν οι ελλείπουσες τιμές να αποτελούσαν κάποιο πρόβλημα, είναι όταν δεν απαντώνται οι συμπληρωματικές ερωτήσεις από τους χρήστες και είναι επιθυμητό το σύστημα να τις λάβει υπόψη για τις προβλέψεις του. Υπό αυτές τις συνθήκες, όμως, μπορεί να χρησιμοποιηθεί η μέθοδος SVD για την εκτίμηση των απαντήσεων των χρηστών στις συμπληρωματικές ερωτήσεις, όπως ήδη έχει γίνει με επιτυχία στην παρούσα έρευνα.

Το πρόβλημα της ψυχρής εκκίνησης στα VAAs, όσον αφορά την κοινωνική σύσταση ερμηνεύεται ως ο αριθμός των χρηστών που χρειάζονται για την καλύτερη εκπαίδευση των μοντέλων κόμματος [57]. Σύμφωνα με τους Tsapatsoulis et al., ([81]), για την εκπαίδευση του προβλήματος μάθησης, το σύνολο εκπαίδευσης μπορεί να περιέχει τους 600-700 πρώτους χρήστες από το σύνολο δεδομένων του EUvox για κάθε χώρα, μιας και έχει αποδειχθεί ότι μπορεί να γίνεται χωρίς να αλλοιώνει την απόδοση της σύστασης. Μέχρι, όμως, να μαζευτούν αυτοί οι πρώτοι χρήστες, το σύστημα μπορεί να κάνει συστάσεις σύμφωνα με την παραδοσιακή προσέγγιση, αφού το μόνο που χρειάζεται είναι οι απαντήσεις των κομμάτων και δε στηρίζεται στις απαντήσεις των άλλων χρηστών (όπως κάνει η κοινωνική σύσταση).

Ένας άλλος τρόπος, όμως, με τον οποίο θα μπορούσε να αντιμετωπιστεί το πρόβλημα της ψυχρής εκκίνησης στα VAAs, είναι συγκρίνοντας τις απαντήσεις των νέων χρηστών με τα μοντέλα κόμματος που δημιουργήθηκαν σε πιο παλιά VAAs για τις ίδιες ερωτήσεις. Έτσι, πρόκειται να χρησιμοποιηθούν τα μοντέλα κόμματος που δημιουργήθηκαν σε αυτή την εργασία σε μελλοντικά VAAs.

Στην παρούσα μελέτη έγινε διερεύνηση της ιδεολογικής ταύτισης του Ευρωπαίου χρήστη με το κόμμα που υποστηρίζει, χρησιμοποιώντας τον αλγόριθμο k -means. Ως μελλοντική εργασία θα χρησιμοποιηθούν και άλλες τεχνικές ομαδοποίησης (clustering) [80, 24] για περαιτέρω εξέταση του θέματος.

Επιπλέον έχει παρατηρηθεί πως οι ερωτήσεις που εμφανίζονται στα VAAs σχετίζονται μεταξύ τους [101, 96]. Σε μεταγενέστερο στάδιο θα γίνει εκμετάλλευση της συσχέτισης των ερωτήσεων για ενδεχόμενη εύρεση χρηστών που δίνουν ψευδείς ή τυχαίες απαντήσεις. Ακόμη, η διασύνδεση αυτή μεταξύ του περιεχομένου των ερωτήσεων πρόκειται να εκμεταλλευτεί για τη βελτίωση της διαδραστικότητας του VAA, βγάζοντας προειδοποιητικά μηνύματα στην οθόνη του χρήστη στην περίπτωση που ανταποκρίνεται διαφορετικά σε ερωτήσεις με παρόμοιο περιεχόμενο,

4.4 Επίλογος

Η παρούσα διατριβή διαπραγματεύτηκε την εφαρμογή των ΣΣ, τα οποία έχουν εφαρμοστεί ευρέως σε τομείς όπως το ηλεκτρονικό εμπόριο, την ψυχαγωγία και τις διαδικτυακές υπηρεσίες, στους Ηλεκτρονικούς Συμβούλους Ψήφου (VAAs). Τα VAAs προτείνουν στους χρήστες έναν υποψήφιο ή ένα υποψήφιο κόμμα από επερχόμενες εκλογές, σύμφωνα με τις απαντήσεις τους σε ένα online ερωτηματολόγιο. Όταν το πρόβλημα σύστασης του VAA προσεγγίζεται παραδοσιακά, γίνεται απλά σύγκριση μεταξύ των απαντήσεων του χρήστη και των απαντήσεων των κομμάτων, για να καταλήξει το σύστημα στο κόμμα που ταιριάζει καλύτερα με το χρήστη. Στην περίπτωση που ακολουθείται η κοινωνική προσέγγιση, τεχνικές συνεργατικού φιλτραρίσματος από τα ΣΣ χρησιμοποιούνται για να συγκρίνουν τις απαντήσεις μεταξύ των χρηστών και να προτείνουν σε έναν χρήστη, το πολιτικό κόμμα/υποψήφιο που δήλωσε πως προτιμάται να ψηφίσει η πλειοψηφία των χρηστών με κοντινές απαντήσεις.

Στόχος της παρούσας διατριβής ήταν να μελετήσει τη συμπεριφορά του Ευρωπαίου χρήστη στο VAA και να βρει τρόπους για τη βελτιστοποίηση της απόδοσης του συστήματος. Όπως αποδείχθηκε από τα πειράματα, όταν ο Ευρωπαίος χρήστης είναι άντρας με υψηλή μόρφωση το ενδιαφέρον του για την πολιτική αυξάνεται περισσότερο, παρά όταν είναι γυναίκα και χωρίς υψηλό επίπεδο μόρφωσης. Όσο αυξάνεται η ηλικιακή του ομάδα, μεγαλώνει και στις περισσότερες χώρες το ενδιαφέρον του για την πολιτική, αλλά όχι σε σημαντικό επίπεδο. Καθώς επίσης, όταν ψηφίζει ένα κόμμα επειδή το υποστηρίζει η κοινότητα του ή επειδή βοηθά τους συνανθρώπους του, δείχνει το ενδιαφέρον του να μειώνεται.

Παρατηρήθηκε ανομοιογένεια στα κόμματα των χωρών, σε διαφορετικό βαθμό από χώρα σε χώρα και από κόμμα σε κόμμα. Επίσης πολλοί χρήστες έδωσαν όμοιες απαντήσεις με χρήστες που δήλωσαν προτίμηση για κάποιο άλλο κόμμα. Όταν συμβαίνει αυτό είναι δύσκολο από το σύστημα να αναγνωρίσει σε ποιο κόμμα ανήκει ο κάθε χρήστης, ειδικά στην περίπτωση του Mahalanobis ταξινομητή. Τα κόμματα που παρουσιάζουν μεγαλύτερη ιδεολογική ταύτιση με τους χρήστες που τα υποστηρίζουν, είναι πιο εύκολο να μοντελοποιηθούν με επιτυχία.

Στις περιπτώσεις που το σύστημα για τη δημιουργία της σύστασης λαμβάνει υπόψη μαζί με τις δηλώσεις πολιτικής και κάποιες συμπληρωματικές ερωτήσεις, παρατηρείται αύξηση στην απόδοση του συστήματος. Βέβαια, τα αποτελέσματα από χώρα σε χώρα ποικίλουν και στην Ενότητα 2.4 παρουσιάζονται οι συμπληρωματικές ερωτήσεις που βοήθησαν περισσότερο τη σύσταση σε κάθε χώρα ξεχωριστά.

Όταν ο χρήστης δεν απαντά σε όλες ή σε κάποιες από τις συμπληρωματικές ερωτήσεις, η τεχνική

παραγοντοποίησης πινάκων SVD μπορεί να υπολογίζει με επιτυχία τις απαντήσεις του. Κάτι που μπορεί να γίνεται γενικά στα VAAs, αφού η επιπλέον γνώση για το χρήστη συντελεί σε αποδοτικότερες συστάσεις ψήφου. Επιπρόσθετα, οι ερωτήσεις που λαμβάνονται υπόψη από το σύστημα είναι καλό να μένουν ως έχουν και όχι να ομαδοποιούνται, αφού όταν το πρόβλημα σύστασης στα VAAs αντιμετωπίστηκε ως πρόβλημα λήψης αποφάσεων με τη χρήση πολλαπλών κριτηρίων και οι ερωτήσεις ομαδοποιήθηκαν κάτω από τρία κύρια κριτήρια, κάτι που συμβαίνει για πρώτη φορά, τα αποτελέσματα δεν ήταν ιδιαίτερα ενθαρρυντικά.

Η μέθοδος που είχε την καλύτερη συμπεριφορά σε όλες τις χώρες ήταν αυτή των Μηχανών Διανυσμάτων Υποστήριξης (SVMs) που εφαρμόστηκε για την κοινωνική σύσταση, αποδεικνύοντας πως οι Ευρωπαίοι χρήστες απαντούν με πολύπλοκο τρόπο στο VAA ερωτηματολόγιο και πως η κοινωνική σύσταση υπερτερεί της παραδοσιακής. Αυτό που διέφερε από χώρα σε χώρα ήταν ο βαθμός πολυπλοκότητας. Επιπλέον, η απόδοση όλων των μεθόδων που χρησιμοποιήθηκαν, εξαρτήθηκε σε μεγάλο βαθμό από τον αριθμό των κομμάτων της κάθε χώρας και του βαθμού ομοιότητας μεταξύ των απαντήσεων που έδωσαν οι χρήστες που υποστήριξαν το ίδιο κόμμα. Βέβαια, εδώ πρέπει να αναφερθεί πως για να μετρηθεί η απόδοση της κοινωνικής σύστασης ψήφου, γίνεται σύγκριση μεταξύ των αποτελεσμάτων της τεχνικής που χρησιμοποιήθηκε και της πρόθεσης ψήφου του χρήστη. Ορισμένοι, όμως, από τους χρήστες που εκφράζουν συγκεκριμένη πρόθεση ψήφου δεν το κάνουν ειλικρινά και για αυτό κάποια από τα αποτελέσματα της κοινωνικής προσέγγισης μπορεί να θεωρούνται λανθασμένα, αλλά στην πραγματικότητα να είναι σωστά. Έτσι, η εκτίμηση της πρόθεσης ψήφου από το πολιτικό τους προφίλ μπορεί στην ουσία να είναι πιο αποδοτική από ότι φαίνεται. Ως εκ τούτου πολλές φορές, μπορεί να βοηθήσει για τις προβλέψεις που γίνονται στις εκλογές πριν την ανακοίνωση του αποτελέσματος [106].

Γενικά οι τεχνικές που χρησιμοποιούνται στα ΣΣ λειτούργησαν πολύ καλά στα VAAs. Ως μελλοντική δουλειά, πρόκειται να χρησιμοποιηθούν η παραδοσιακή προσέγγιση και τα μοντέλα κόμματος από προηγούμενα VAAs για την αντιμετώπιση ενός κοινού προβλήματος στα ΣΣ, του λεγόμενου προβλήματος ψυχρής εκκίνησης. Αυτό προκαλείται όταν δεν υπάρχει αρκετή πληροφορία για τη δημιουργία της σύστασης και παρατηρείται στους πρώτους χρήστες του VAA, όπου τα δείγματα δεν είναι αρκετά για τη σωστή εκπαίδευση του συστήματος. Επιπρόσθετα πρόκειται να διερευνηθεί περαιτέρω η ιδεολογική ταύτιση των χρηστών με τα κόμματα που υποστηρίζουν, χρησιμοποιώντας και άλλες τεχνικές ομαδοποίησης (clustering). Ακόμη θα γίνει εκμετάλλευση της συσχέτισης των ερωτήσεων για ενδεχόμενη εύρεση τυχαίων και ψευδών απαντήσεων, καθώς επίσης και για τη βελτίωση της διαδραστικότητας των VAAs.

Λίστα Δημοσιεύσεων

- M.Agathokleous and N.Tsapatsoulis, “Voting advice applications: Missing value estimation using matrix factorization and collaborative filtering,” in IFIP International Conference on Artificial Intelligence Applications and Innovations. Springer,2013,pp.20–29.
- M. Agathokleous, N. Tsapatsoulis, and I. Katakis, “On the quantification of missing value impact on voting advice applications,” in International Conference on Engineering Applications of Neural Networks. Springer,2013,pp.496–505.
- M. Agathokleous and N. Tsapatsoulis, “Learning user models in multi-criteria recommender systems,” in International Conference on Engineering Applications of Neural Networks. Springer,2014,pp.205–216.
- N. Tsapatsoulis, M. Agathokleous, C. Djouvas, and F. Mendez, “On the design of social voting recommendation applications,” International Journal on Artificial Intelligence Tools, vol. 24, no.3,2015.
- M. Agathokleous and N. Tsapatsoulis, “Applying hidden markov models to voting advice applications,” EPJDataScience,vol.5,pp.1–34,2016.
- M. Agathokleous, N. Tsapatsoulis, and C. Djouvas, “Estimating party-user similarity in voting advice applications using hidden markov models,” inIntRS@RecSys,2016,pp.6–13.
- N. Tsapatsoulis and M. Agathokleous, “Forecasting elections from VAA data: What the undecided would vote?” in Semantic and Social Media Adaptation and Personalization (SMAP), 2017 12th International Workshop on. IEEE,2017,pp.46–52.
- N. Tsapatsoulis and M. Agathokleous, “Voting Intention: “I Prefer not to Say” versus “I am Undecided” in ECPR general conference.

Βιβλιογραφία

- [1] F. Mendez, “Matching voters with political parties and candidates: An empirical test of four algorithms,” *International Journal of Electronic Governance*, vol. 5, no. 3-4, pp. 264–278, 2012.
- [2] G. Rabinowitz and S. E. Macdonald, “A directional theory of issue voting,” *American Political Science Review*, vol. 83, no. 1, pp. 93–121, 1989.
- [3] T. Berners-Lee, R. Cailliau, J.-F. Groff, and B. Pollermann, “World-wide web: The information universe,” *Internet Research*, vol. 20, no. 4, pp. 461–471, 2010.
- [4] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender Systems: An Introduction*, 1st ed. New York, NY, USA: Cambridge University Press, 2010.
- [5] F. Ricci, L. Rokach, and B. Shapira, “Recommender systems: introduction and challenges,” in *Recommender systems handbook*, F. Ricci, L. Rokach, and B. Shapira, Eds. Springer, 2015, ch. 1, pp. 1–36.
- [6] M. McLuhan, W. T. Gordon, E. Lamberti, and D. Scheffel-Dunand, *The Gutenberg galaxy: the making of typographic man*. University of Toronto Press, 2014. [Online]. Available: <http://www.utppublishing.com/The-Gutenberg-Galaxy.html>
- [7] N. Negroponte, *Being Digital*, 1st ed., M. Asher, Ed. New York, NY, USA: Random House Inc., 1995.
- [8] Μ. Δερτούζος, *Τι μέλλει γενέσθαι: Πώς ο νέος κόσμος της πληροφορίας θ’ αλλάξει τη ζωή μας*, 1st ed. Λιβάνης - Νέα Σύνορα, 1998.
- [9] T. O’Reilly, *What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software*. O’Reilly Media, Inc., 2005.
- [10] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, “Recommender systems,” *Physics Reports*, vol. 519, no. 1, pp. 1–49, 2012.
- [11] S. Perugini, M. A. Gonçalves, and E. A. Fox, “Recommender systems research: A connection-centric survey,” *Journal of Intelligent Information Systems*, vol. 23, no. 2, pp. 107–143, 2004.
- [12] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, Jun. 2005.
- [13] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, “GroupLens: an open architecture for collaborative filtering of netnews,” in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. ACM, 1994, pp. 175–186.

- [14] P. Resnick and H. R. Varian, “Recommender systems,” *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.
- [15] R. Burke, A. Felfernig, and M. H. Göker, “Recommender systems: An overview,” *Ai Magazine*, vol. 32, no. 3, pp. 13–18, 2011.
- [16] J. A. Konstan and J. Riedl, “Recommender systems: from algorithms to user experience,” *User Modeling and User-Adapted Interaction*, vol. 22, no. 1, pp. 101–123, 2012.
- [17] G. Adomavicius and A. Tuzhilin, “Context-aware recommender systems,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, and B. Shapira, Eds. Springer, 2015, ch. 6, pp. 191–226.
- [18] P. Lops, M. De Gemmis, and G. Semeraro, “Content-based recommender systems: State of the art and trends,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Springer, 2011, ch. 3, pp. 73–105.
- [19] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, “Recommender systems with social regularization,” in *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 2011, pp. 287–296.
- [20] J. S. Breese, D. Heckerman, and C. Kadie, “Empirical analysis of predictive algorithms for collaborative filtering,” in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1998, pp. 43–52.
- [21] M. D. Ekstrand, J. T. Riedl, and J. A. Konstan, “Collaborative filtering recommender systems,” *Foundations and Trends® in Human–Computer Interaction*, vol. 4, no. 2, pp. 81–173, 2011.
- [22] W. Hill, L. Stead, M. Rosenstein, and G. Furnas, “Recommending and evaluating choices in a virtual community of use,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 1995, pp. 194–201.
- [23] X. Amatriain and J. M. Pujol, “Data mining methods for recommender systems,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, and B. Shapira, Eds. Springer, 2015, ch. 7, pp. 227–262.
- [24] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Morgan Kaufmann, 2017.
- [25] S. Marsland, *Machine learning: an algorithmic perspective*. CRC press, 2015.
- [26] E. Alpaydin, *Introduction to machine learning*. MIT press, 2014.
- [27] W. B. Croft, D. Metzler, and T. Strohman, *Search engines: Information retrieval in practice*. Pearson Education, 2010.
- [28] D. Asanov, “Algorithms and methods in recommender systems,” *Berlin Institute of Technology, Berlin, Germany*, 2011.
- [29] D. Cosley, S. K. Lam, I. Albert, J. A. Konstan, and J. Riedl, “Is seeing believing?: How recommender system interfaces affect users’ opinions,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2003, pp. 585–592.
- [30] F. Ricci, L. Rokach, and B. Shapira, *Recommender systems handbook*. Springer, 2015.

-
- [31] J. B. Schafer, J. Konstan, and J. Riedl, “Recommender systems in e-commerce,” in *Proceedings of the 1st ACM conference on Electronic commerce*. ACM, 1999, pp. 158–166.
- [32] J. Reason, *Human error*. Cambridge university press, 1990.
- [33] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, “Analysis of recommendation algorithms for e-commerce,” in *Proceedings of the 2nd ACM conference on Electronic commerce*. ACM, 2000, pp. 158–167.
- [34] A. Gunawardana and G. Shani, “Evaluating recommender systems,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, and B. Shapira, Eds. Springer, 2015, ch. 8, pp. 265–308.
- [35] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, “Recommender system application developments: a survey,” *Decision Support Systems*, vol. 74, pp. 12–32, 2015.
- [36] G. Linden, B. Smith, and J. York, “Amazon. com recommendations: Item-to-item collaborative filtering,” *IEEE Internet computing*, vol. 7, no. 1, pp. 76–80, 2003.
- [37] A. Ladner and J. Pianzola, “Do voting advice applications have an effect on electoral participation and voter turnout? evidence from the 2007 swiss federal elections,” *Electronic participation*, pp. 211–224, 2010.
- [38] K. Gemenis and M. Rosema, “Voting advice applications and electoral turnout,” *Electoral studies*, vol. 36, pp. 281–289, 2014.
- [39] O. Ruusuvirta and M. Rosema, “Do online vote selectors influence electoral participation and the direction of the vote,” in *ECPR general conference*, 2009, pp. 13–12.
- [40] J. Fivaz and D. Schwarz, “Nailing the pudding to the wall—e-democracy as catalyst for transparency and accountability,” in *International conference on direct democracy in Latin America*, 2007, pp. 14–15.
- [41] M. E. Milakovich, “The internet and increased citizen participation in government,” *JeDEM—eJournal of eDemocracy and Open Government*, vol. 2, no. 1, pp. 1–9, 2010.
- [42] J. Fivaz, J. Pianzola, and A. Ladner, “More than toys: a first assessment of voting advice applications’ impact on the electoral decision of voters,” *National Centre of Competence in Research (NCCR): Challenges to Democracy in the 21st Century, Working Paper No. 48*, 2010.
- [43] T. Louwerse and M. Rosema, “The design effects of voting advice applications: Comparing methods of calculating matches,” *Acta politica*, vol. 49, no. 3, pp. 286–312, 2014.
- [44] A. Marzuca, U. Serdült, and Y. Welp, “Questão pública: first voting advice application in latin america,” *Electronic Participation*, pp. 216–227, 2011.
- [45] M. Rosema, J. Anderson, and S. Walgrave, “The design, purpose, and effects of voting advice applications,” *Electoral studies*, vol. 36, pp. 240–243, 2014.
- [46] M. Wagner and O. Ruusuvirta, “Matching voters to parties: Voting advice applications and models of party choice,” *Acta politica*, vol. 47, no. 4, pp. 400–422, 2012.
- [47] A. Downs, “An economic theory of political action in a democracy,” *Journal of Political Economy*, vol. 65, no. 2, pp. 135–150, 1957.
-

- [48] J. De Graaf, “The irresistible rise of stemwijzer,” in *Voting Advice Applications in Europe: The State of the Art*, L. Cedroni and D. Garzia, Eds. Scriptaweb, Napoli, 2010, ch. 2, pp. 35–46.
- [49] S. Walgrave, P. Van Aelst, and M. Nuytemans, “‘do the vote test’: The electoral effects of a popular vote advice application at the 2004 belgian elections,” *Acta Politica*, vol. 43, no. 1, pp. 50–70, 2008.
- [50] A. Baka, L. Figgou, and V. Triga, “‘neither agree, nor disagree’: a critical analysis of the middle answer category in voting advice applications,” *International Journal of Electronic Governance*, vol. 5, no. 3-4, pp. 244–263, 2012.
- [51] D. Garzia and S. Marschall, “Voting advice applications under review: the state of research,” *International Journal of Electronic Governance*, vol. 5, no. 3, pp. 203–222, 2012.
- [52] M. L. Sudulich, D. Garzia, A. H. Trechsel, and K. Vassil, “Matching voters with parties in supranational elections: the case of the eu profiler,” in *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*, D. Garzia and S. Marschall, Eds. ECPR Press, Colchester, 2014, ch. 13, pp. 175–182.
- [53] F. Breuer, “The eu profiler: a new way for voters to meet parties and to understand european elections,” in *The 2009 Elections to the European Parliament – Country Reports*, W. Gagatsek, Ed. European Union Democracy Observatory, 2010, ch. 4, pp. 27–31.
- [54] K. Vassil, “Role of self selection in estimating the effects of voting advice applications empirical evidence on the basis of swiss smartvote data,” in *Paper Presented at the 6th ECPR General Conference, 25–27 August, Reykjavik, Iceland*. ECPR, 2011.
- [55] L. Cedroni and D. Garzia, *Voting Advice Applications in Europe: The State of the Art*. Scriptaweb, Napoli, 2010.
- [56] D. Garzia and S. Marschall, *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*. Ecpr Press, Colchester, 2014.
- [57] I. Katakis, N. Tsapatsoulis, F. Mendez, V. Triga, and C. Djouvas, “Social voting advice applications—definitions, challenges, datasets and evaluation,” *IEEE Transactions on cybernetics*, vol. 44, no. 7, pp. 1039–1052, 2014.
- [58] J. Lefevere and S. Walgrave, “A perfect match? the impact of statement selection on voting advice applications’ ability to match voters and parties,” *Electoral Studies*, vol. 36, pp. 252–262, 2014.
- [59] K. Van Camp, J. Lefevere, and S. Walgrave, “The content and formulation of statements in voting advice applications: A comparative analysis of 26 vaas,” in *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*, D. Garzia and S. Marschall, Eds. ECPR Press, Colchester, 2014, ch. 2, pp. 11–31.
- [60] K. Gemenis, “Estimating parties’ policy positions through voting advice applications: Some methodological considerations,” *Acta politica*, vol. 48, no. 3, pp. 268–295, 2013.
- [61] J. A. Krosnick and S. Presser, “Question and questionnaire design,” in *Handbook of survey research*, P. V. Marsden and J. D. Wright, Eds. Emerald Group Publishing, 2010, ch. 9, pp. 263–314.

-
- [62] K. Gemenis, I. Konstantinidis, and V. Triga, “Οι ηλεκτρονικοί σύμβουλοι ψήφου στην Ελλάδα - voting advice applications in greece,” *Επιστήμη και κοινωνία: επιθεώρηση πολιτικής και ηθικής θεωρίας - Science and society: journal of political and moral theory*, vol. 33, pp. 177–210, 2015.
- [63] I. Katakis, N. Tsapatsoulis, V. Triga, C. Tziouvas, and F. Mendez, “Clustering online poll data: Towards a voting assistance system,” in *Semantic and Social Media Adaptation and Personalization (SMAP), 2012 Seventh International Workshop on*. IEEE, 2012, pp. 54–59.
- [64] K. Gemenis and C. Van Ham, “Comparing methods for estimating parties’ positions in voting advice applications,” in *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*, D. Garzia and S. Marschall, Eds. ECPR Press, Colchester, 2014, ch. 3, pp. 33–47.
- [65] F. Mendez, “What’s behind a matching algorithm: A critical assessment of how vaas produce voting recommendations,” in *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*, D. Garzia and S. Marschall, Eds. ECPR Press, Colchester, 2014, ch. 4, pp. 49–66.
- [66] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, “On the surprising behavior of distance metrics in high dimensional spaces,” in *Proceedings of the 8th International Conference on Database Theory*, ser. ICDT ’01. Springer-Verlag, 2001, pp. 420–434.
- [67] I. Andreadis, “Data quality and data cleaning,” in *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*, D. Garzia and S. Marschall, Eds. ECPR Press, Colchester, 2014, ch. 6, pp. 79–91.
- [68] C. Djouvas, K. Gemenis, and F. Mendez, “Weeding out the rogues: how to identify them and why it matters for vaa-generated datasets,” in *Proceedings of the 2014 European consortium for political research general conference*. ECPR, Colchester, 2014, pp. 1–7.
- [69] C. Djouvas, F. Mendez, and N. Tsapatsoulis, “Mining online political opinion surveys for suspect entries: An interdisciplinary comparison,” *Journal of Innovation in Digital Ecosystems*, vol. 3, no. 2, pp. 172–182, 2016.
- [70] M. Agathokleous and N. Tsapatsoulis, “Applying hidden markov models to voting advice applications,” *EPJ Data Science*, vol. 5, pp. 1–34, 2016.
- [71] N. Tsapatsoulis, M. Agathokleous, C. Djouvas, and F. Mendez, “On the design of social voting recommendation applications,” *International Journal on Artificial Intelligence Tools*, vol. 24, no. 3, 2015.
- [72] W. Khreich, E. Granger, A. Miri, and R. Sabourin, “A survey of techniques for incremental learning of hmm parameters,” *Information Sciences*, vol. 197, pp. 105–130, 2012.
- [73] L. Rabiner and B. Juang, “An introduction to hidden markov models,” *ieee assp magazine*, vol. 3, no. 1, pp. 4–16, 1986.
- [74] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the em algorithm,” *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.
- [75] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, “A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains,” *The annals of mathematical statistics*, vol. 41, no. 1, pp. 164–171, 1970.
-

- [76] D. Bremner, E. Demaine, J. Erickson, J. Iacono, S. Langerman, P. Morin, and G. Toussaint, “Output-sensitive algorithms for computing nearest-neighbour decision boundaries,” *Discrete & Computational Geometry*, vol. 33, no. 4, pp. 593–604, 2005.
- [77] G. McLachlan, *Discriminant analysis and statistical pattern recognition*. John Wiley & Sons, 2004, vol. 544.
- [78] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–444, 2015.
- [79] C.-C. Chang and C.-J. Lin, “Libsvm: A library for support vector machines,” *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, 2011.
- [80] T. Segaran, *Programming collective intelligence: building smart web 2.0 applications*. O’Reilly Media, Inc., 2007.
- [81] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.
- [82] C. Manning, P. Raghavan, and H. Schütze, *Introduction to information retrieval*. Cambridge University Press, Cambridge, England, 2008.
- [83] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, “Recommender systems survey,” *Knowledge-based systems*, vol. 46, pp. 109–132, 2013.
- [84] L. Terán, A. Ladner, J. Fivaz, and S. Gerber, “Using a fuzzy-based cluster algorithm for recommending candidates in e-elections,” in *Digital Democracy: Concepts, Methodologies, Tools and Applications*, M. Khosrow-Pour, S. Clarke, M. E. Jennex, A. Becker, and A.-V. Anttiroiko, Eds. Information Science Reference-Imprint of: IGI Publishing, 2012, ch. 34, pp. 684–705.
- [85] R. Bambini, P. Cremonesi, and R. Turrin, “A recommender system for an iptv service provider: a real large-scale production environment,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Springer, 2011, ch. 7, pp. 299–331.
- [86] X. Ning, C. Desrosiers, and G. Karypis, “A comprehensive survey of neighborhood-based recommendation methods,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, and B. Shapira, Eds. Springer, 2015, ch. 2, pp. 37–76.
- [87] M. Agathokleous and N. Tsapatsoulis, “Voting advice applications: Missing value estimation using matrix factorization and collaborative filtering,” in *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 2013, pp. 20–29.
- [88] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, “Singular value decomposition and principal component analysis,” in *A practical approach to microarray data analysis*, D. P. Berrar, W. Dubitzky, and M. Granzow, Eds. Springer, 2003, ch. 5, pp. 91–109.
- [89] R. Salakhutdinov and A. Mnih, “Probabilistic matrix factorization,” *Advances in neural information processing systems*, vol. 20, no. 1, pp. 1257–1264, 2008.
- [90] T. Zhou, H. Shan, A. Banerjee, and G. Sapiro, “Kernelized probabilistic matrix factorization: Exploiting graphs and side information,” in *Proceedings of the 2012 SIAM international Conference on Data mining*. SIAM, 2012, pp. 403–414.
- [91] A. K. Jain, “Data clustering: 50 years beyond k-means,” *Pattern recognition letters*, vol. 31, no. 8, pp. 651–666, 2010.

-
- [92] G. Lutz, “Low turnout in direct democracy,” *Electoral Studies*, vol. 26, no. 3, pp. 624–632, 2007.
- [93] A. Ladner, J. Fivaz, and J. Pianzola, “Voting advice applications and party choice: evidence from smartvote users in switzerland,” *International Journal of Electronic Governance*, vol. 5, no. 3-4, pp. 367–387, 2012.
- [94] J. S. Dodgson, M. Spackman, A. Pearman, and L. D. Phillips, “Multi-criteria analysis: a manual,” Department for Communities and Local Government: London, Tech. Rep., 2009.
- [95] M. Agathokleous and N. Tsapatsoulis, “Learning user models in multi-criteria recommender systems,” in *International Conference on Engineering Applications of Neural Networks*. Springer, 2014, pp. 205–216.
- [96] M. Agathokleous, N. Tsapatsoulis, and I. Katakis, “On the quantification of missing value impact on voting advice applications,” in *International Conference on Engineering Applications of Neural Networks*. Springer, 2013, pp. 496–505.
- [97] N. Tsapatsoulis and F. Mendez, “Social vote recommendation: Building party models using the probability to vote feedback of vaa users,” in *Semantic and Social Media Adaptation and Personalization (SMAP), 2014 9th international Workshop on*. IEEE, 2014, pp. 124–129.
- [98] G. Forman and I. Cohen, “Learning from little: Comparison of classifiers given little training,” *Knowledge Discovery in Databases: PKDD 2004*, pp. 161–172, 2004.
- [99] M. Stern, J. Beck, and B. Woolf, “Naive bayes classifiers for user modeling,” *Center for Knowledge Communication, Computer Science Department, University of Massachusetts*, 1999.
- [100] S. Karsoliya, “Approximating number of hidden layer neurons in multiple hidden layer bpnn architecture,” *International Journal of Engineering Trends and Technology*, vol. 3, no. 6, pp. 714–717, 2012.
- [101] M. Agathokleous, N. Tsapatsoulis, and C. Djouvas, “Estimating party-user similarity in voting advice applications using hidden markov models,” in *IntRS@ RecSys*, 2016, pp. 6–13.
- [102] E. Á. Juliusson, N. Karlsson, and T. Gärling, “Weighing the past and the future in decision making,” *European Journal of Cognitive Psychology*, vol. 17, no. 4, pp. 561–575, 2005.
- [103] R. F. West, R. J. Meserve, and K. E. Stanovich, “Cognitive sophistication does not attenuate the bias blind spot,” *Journal of personality and social psychology*, vol. 103, no. 3, pp. 506–519, 2012.
- [104] W. Bruine de Bruin, A. M. Parker, and B. Fischhoff, “Individual differences in adult decision-making competence,” *Journal of personality and social psychology*, vol. 92, no. 5, pp. 938–956, 2007.
- [105] J. Van de Pol, B. Holleman, N. Kamoen, A. Krouwel, and C. De Vreese, “Beyond young, highly educated males: a typology of vaa users,” *Journal of Information Technology & Politics*, vol. 11, no. 4, pp. 397–411, 2014.
- [106] N. Tsapatsoulis and M. Agathokleous, “Forecasting elections from vaa data: What the undecided would vote?” in *Semantic and Social Media Adaptation and Personalization (SMAP), 2017 12th International Workshop on*. IEEE, 2017, pp. 46–52.
-

Παράρτημα Ι

Τα ονόματα των κομμάτων σε κάθε χώρα

Πίνακας Ι.1: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Α

Κόμμα	Αυστρία	Βουλγαρία	Κύπρος	Τσεχία	Γερμανία
1	SPÖ	ГЕРБ	ΔΗΣΥ	ČSSD	CDU/CSU
2	ÖVP	БСП	ΑΚΕΛ	ANO 2011	SPD
3	FPÖ	Атака	ΔΗΚΟ	KSČM	Bündnis 90/Die Grünen
4	Grüne	Реформаторски блок	ΕΔΕΚ	TOP 09	Die Linke
5	NEOS	България без цензура	Συμμαχία Πολιτών	ODS	FDP
6	Europa Anders		ΕΛΑΜ	ÚSVIT	Alternative für Deutschland
7	REKOS			KDU-ČSL	Piratenpartei
8	EU-STOP			Strana Zelených	
9				Česká Pirátská Strana	
10				SVOBODNÍ	

Πίνακας Ι.2: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Β

Κόμμα	Δανία	Εσθονία	Αγγλία	Βόρεια Ιρλανδία
1	Socialdemokraterne	Isamaa ja Res Publica Liit	Conservative	Democratic Unionist Party
2	Det Radikale Venstre	Keskerakond	Labour	Sinn Féin
3	De Konservative	Konservatiivne	Liberal Democrats	Social Democratic and Labour Party
4	Socialistisk Folkeparti	Reformierakond	Green	Ulster Unionist party
5	Liberal Alliance	Sotsiaaldemokraadid	UKIP	Alliance
6	Dansk Folkeparti	Hääletan üksikkandidaadi poolt		Green
7	Folkebevægelsen mod EU			

Πίνακας Ι.3: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Γ

Κόμμα	ΗΒ-Σκωτία	ΗΒ-Ουαλία	Ισπανία	Φινλανδία
1	Labour	Labour	PP	Kokoomus
2	SNP	Conservative	PSOE	SDP
3	Conservative	Plaid Cymru	UPyD	Perussuomalaiset
4	Liberal Democrats	Liberal Democrats	Ciudadanos	Keskusta
5	Green	Green	VOX	Vasemmistoliitto
6	UKIP	UKIP	Podemos	Vihreä liitto
7			RED	RKP
8			IU-ICV-Anova	Kristillisdemokraatit
9			CIU-PNV-CC-CxG	
10			ERC-Necat-CatSi	
11			Bildu-BNG-AA-ANC-UP	
12			Equo-Compromis-Chunta	

Πίνακας Ι.4: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Δ

Κόμμα	Γαλλία	Ελλάδα	Κροατία	Ουγγαρία	Ιρλανδία
1	Front de Gauche	Νέα Δημοκρατία	SDP-HNS-IDS-HSU	Fidesz	Fine Gael
2	Parti Socialiste	ΣΥΡΙΖΑ	HDZ-HSS-HSP AS-BUZ-ZDS-HDS	Jobbik	Labour Party
3	Les Verts Europe Ecologie	Ελιά	HDSSB-ABH-A HSS-HRAST-HSP-HZ-OS-ZZH	LMP	Fianna Fáil
4	UDI-MoDem	ΑΝΕΛ	Hrvatski laburisti – stranka rada	Együtt PM	Sinn Féin
5	UMP	ΚΚΕ	OraH	Demokratikus Koalíció	Socialist Party
6	Debout la République	Χρυσή Αυγή			Green Party
7	Front National	ΔΗΜΑΡ			Independent
8		Το Ποτάμι			
9		Δράση			

Πίνακας Ι.5: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Ε

Κόμμα	Ιταλία	Λιθουανία	Λετονία
1	Fratelli d'Italia - Alleanza Nazionale	Lietuvos socialdemokratų partija	Saskaņa sociāldemokrātiskā partija
2	Forza Italia	Tēvynės sąjunga - Lietuvos krikščionis demokratus	Vienotība
3	Lega Nord	Lietuvos Respublikos liberalų sąjūdį	Nacionālā apvienība Visu Latvijai! -Tēvzemei un Brīvībai /LNNK
4	Movimento 5 Stelle	Partija Tvarka ir teisingumas	
5	Nuovo Centro Destra - Unione di Centro		
6	Partito Democratico		
7	Scelta Europea		
8	L'altra Europa con Tsipras		

Πίνακας Ι.6: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Στ

Κόμμα	Ολλανδία	Πολωνία	Πορτογαλία
1	VVD	Platforma Obywatelska	Aliança Portugal: PSD/CDS-PP
2	PVV	Prawo i Sprawiedliwość	CDU: PCP/PEV
3	SP Socialistische Partij	SLD-UP	BE
4	D66 Democraten 66	Europa + Twój Ruch	PS
5	CU-SGP	Nowa Prawica	Livre
6	GroenLinks	Ruch Narodowy	MPT
7		Polska Razem	
8		Zieloni	

Πίνακας Ι.7: Τα ονόματα των κομμάτων για κάθε χώρα - Μέρος Η

Κόμμα	Ρουμανία	Σουηδία	Σλοβακία
1	PSD-UNPR-PC USD	Socialdemokraterna	KDH
2	PNL	Moderaterna	ESNS
3	PDL	Kristdemokraterna	Most-Híd
4	PMP	Miljöpartiet	NOVA
5	UDMR RMDSZ	Feministiskt initiativ	OĽaNO
6		Centerpartiet	SaS
7		Vänsterpartiet	SDKÚ-DS
8		Folkpartiet	SMER-SD
9		Piratpartiet	
10		Sverigedemokraterna	

Παράρτημα II

Αναλυτικοί πίνακες αποτελεσμάτων

Σε αυτό το Παράρτημα βρίσκονται τα αποτελέσματα των ταξινομητών, αναλυτικά για κάθε κόμμα της κάθε χώρας.

Πίνακας II.1: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Αυστρία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.1732	0.2312	0.1980	0.1082	0.2381	0.1488	0.2771	0.1855	0.2222
2	0.4853	0.3625	0.4150	0.2092	0.5814	0.3077	0.3473	0.3255	0.3360
3	0.6903	0.5474	0.6106	0.7920	0.5812	0.6704	0.5708	0.6202	0.5945
4	0.5054	0.5442	0.5241	0.5039	0.5310	0.5171	0.5364	0.4799	0.5066
5	0.4070	0.5451	0.4660	0.7736	0.3421	0.4744	0.3854	0.4074	0.3961
6	0.4874	0.4265	0.4550	0.1218	0.5577	0.2000	0.1933	0.4299	0.2667
7	0.1333	0.0784	0.0988	0	0	0	0.1000	0.1500	0.1200
8	0.2812	0.2571	0.2687	0	0	0	0.0625	0.4000	0.1081
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.0260	0.3000	0.0478	0	0	0	0.0996	0.3594	0.1559
2	0.4351	0.3838	0.4078	0.4226	0.4262	0.4244	0.3891	0.4266	0.407
3	0.7832	0.5549	0.6495	0.7566	0.6357	0.6909	0.7434	0.6486	0.6928
4	0.8775	0.4177	0.5660	0.707	0.4861	0.5761	0.6806	0.5208	0.5901
5	0.0377	0.8750	0.0724	0.6065	0.3961	0.4792	0.5741	0.412	0.4797
6	0.0882	0.6774	0.1561	0	0	0	0.2731	0.5856	0.3725
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
Κόμμα	NN7			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.1169	0.3214	0.1714	0.1255	0.2929	0.1758	0.2944	0.1937	0.2337
2	0.364	0.5088	0.4244	0.3305	0.5683	0.4180	0.3305	0.3970	0.3607
3	0.7301	0.6322	0.6776	0.8009	0.6373	0.7098	0.4513	0.5368	0.4904
4	0.6822	0.5183	0.5890	0.6620	0.5226	0.5841	0.4109	0.5176	0.4581
5	0.6011	0.4069	0.4853	0.6065	0.3996	0.4818	0.5418	0.3865	0.4512
6	0.2521	0.6061	0.3561	0.2521	0.6186	0.3582	0.1218	0.3295	0.1779
7	0	0	0	0.0667	0.2222	0.1026	0.0333	0.0125	0.0182
8	0	0	0	0.0938	0.7500	0.1667	0.3125	0.1389	0.1923

Πίνακας ΙΙ.2: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Βουλγαρία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3426	0.5114	0.4103	0.8325	0.4426	0.5780	0.6396	0.4352	0.5180
2	0.2832	0.3478	0.3122	0.4425	0.3289	0.3774	0.2389	0.3699	0.2903
3	0.6000	0.4636	0.5231	0.1765	0.5172	0.2632	0.4353	0.4022	0.4181
4	0.5706	0.6884	0.6240	0.2673	0.7479	0.3938	0.5015	0.5623	0.5302
5	0.3409	0.1786	0.2344	0.1250	0.2821	0.1732	0.0909	0.2162	0.1280
Mahalanobis			NN1			NN2			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6117	0.4273	0.5031	0.7868	0.3813	0.5137	0.7690	0.4927	0.6006
2	0.3009	0.3696	0.3317	0	0	0	0.6460	0.3042	0.4136
3	0.1647	0.5385	0.2523	0	0	0	0	0	0
4	0.2883	0.6316	0.3959	0.5616	0.6471	0.6013	0.4955	0.6680	0.5690
5	0.0455	0.1818	0.0727	0	0	0	0	0	0
NN3			SVM			Party Coding			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7437	0.5034	0.6004	0.7538	0.4655	0.5756	0.1929	0.4606	0.2719
2	0.5752	0.3421	0.4290	0.5310	0.4412	0.4819	0.0531	0.1176	0.0732
3	0.2941	0.3788	0.3311	0.3882	0.5238	0.4459	0.5647	0.3179	0.4068
4	0.5255	0.6629	0.5863	0.4595	0.7051	0.5564	0.6456	0.5283	0.5811
5	0	0	0	0.0795	0.1667	0.1077	0.2159	0.0841	0.1210

Πίνακας ΙΙ.3: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Κύπρος

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7569	0.7842	0.7703	0.8924	0.6624	0.7604	0.8472	0.6472	0.7338
2	0.7198	0.9030	0.8011	0.6763	0.9150	0.7778	0.8551	0.7532	0.8009
3	0.2258	0.2029	0.2137	0.1774	0.1667	0.1719	0.1452	0.2000	0.1682
4	0.1538	0.2308	0.1846	0.1538	0.2069	0.1765	0.1538	0.2727	0.1967
5	0.2593	0.4468	0.3281	0.3086	0.3012	0.3049	0.1605	0.3095	0.2114
6	0.6780	0.3922	0.4969	0.3220	0.4872	0.3878	0.3220	0.5135	0.3958
Mahalanobis			NN3			NN6			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.8646	0.6484	0.7411	0.8264	0.6819	0.7473	0.8368	0.6770	0.7484
2	0.7150	0.8655	0.7831	0.7971	0.8250	0.8108	0.8164	0.8244	0.8204
3	0.2903	0.2535	0.2707	0	0	0	0.129	0.1667	0.1455
4	0.1026	0.8000	0.1818	0	0	0	0.2051	0.4211	0.2759
5	0.0617	0.3125	0.1031	0.5309	0.2829	0.3691	0.3210	0.3250	0.3230
6	0.5763	0.3036	0.3977	0.4746	0.4828	0.4786	0.4746	0.5490	0.5091
NN9			SVM			Party Coding			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.8229	0.6791	0.7441	0.8819	0.6828	0.7697	0.4896	0.8011	0.6078
2	0.8116	0.8317	0.8215	0.8792	0.8235	0.8505	0.6522	0.8824	0.7500
3	0.1774	0.1864	0.1818	0.1129	0.2917	0.1628	0.1935	0.1364	0.1600
4	0.1538	0.3750	0.2182	0.2051	0.5333	0.2963	0.3590	0.1750	0.2353
5	0.3086	0.3012	0.3049	0.358	0.3816	0.3694	0.4938	0.2410	0.3239
6	0.4237	0.5000	0.4587	0.4407	0.5098	0.4727	0.4576	0.4286	0.4426

Πίνακας Π.4: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Τσεχία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4778	0.4687	0.4732	0.3472	0.4682	0.3987	0.5444	0.2638	0.3554
2	0.2597	0.3262	0.2892	0.3323	0.2917	0.3107	0.3122	0.2217	0.2593
3	0.5411	0.2970	0.3835	0.4110	0.3681	0.3883	0.2740	0.2469	0.2597
4	0.6953	0.5115	0.5894	0.7650	0.4699	0.5822	0.4818	0.4958	0.4887
5	0.2192	0.2874	0.2487	0.0365	0.2857	0.0648	0.2900	0.2153	0.2471
6	0.4246	0.1836	0.2563	0.0670	0.4000	0.1148	0.0223	0.0816	0.0351
7	0.2889	0.7027	0.4094	0.1250	0.6000	0.2069	0.2306	0.4663	0.3086
8	0.4941	0.5476	0.5194	0.4927	0.5268	0.5092	0.3937	0.4649	0.4263
9	0.2281	0.4171	0.2949	0.1984	0.4123	0.2679	0.1719	0.2835	0.2140
10	0.6980	0.7183	0.7080	0.8016	0.6302	0.7056	0.6445	0.6995	0.6709
Κόμμα	Mahalanobis			NN10			NN11		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2833	0.5667	0.3778	0.4528	0.4441	0.4484	0.4333	0.4522	0.4426
2	0.2674	0.3040	0.2845	0.3972	0.3261	0.3582	0.3926	0.3286	0.3577
3	0.2260	0.4400	0.2986	0.363	0.4569	0.4046	0.3425	0.4098	0.3731
4	0.7368	0.4242	0.5385	0.6894	0.5038	0.5822	0.6627	0.5045	0.5729
5	0.3653	0.2309	0.2829	0.1256	0.3481	0.1846	0.1804	0.3264	0.2324
6	0.3240	0.2094	0.2544	0	0	0	0	0	0
7	0.1472	0.5354	0.2309	0.3528	0.4922	0.411	0.3583	0.5223	0.4250
8	0.1915	0.7005	0.3008	0.5244	0.4726	0.4972	0.5535	0.4844	0.5166
9	0.2516	0.3523	0.2935	0.1938	0.4106	0.2633	0.2094	0.4295	0.2815
10	0.6696	0.6829	0.6762	0.7798	0.6706	0.7211	0.7757	0.6685	0.7181
Κόμμα	NN26			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4611	0.4523	0.4567	0.4889	0.5116	0.5000	0.3083	0.2913	0.2996
2	0.4080	0.3346	0.3677	0.4219	0.3482	0.3816	0.0757	0.2707	0.1184
3	0.3562	0.4298	0.3895	0.4041	0.4214	0.4126	0.6301	0.2473	0.3552
4	0.6471	0.5029	0.5660	0.6583	0.5151	0.5779	0.6316	0.4988	0.5574
5	0.2146	0.3745	0.2729	0.1986	0.3655	0.2574	0.5502	0.1997	0.2930
6	0.0056	0.5000	0.0110	0.0838	0.3409	0.1345	0.1508	0.1200	0.1337
7	0.3750	0.4909	0.4252	0.3361	0.5105	0.4054	0.3028	0.2884	0.2954
8	0.5535	0.4549	0.4994	0.5482	0.4976	0.5217	0.3104	0.4728	0.3748
9	0.2250	0.4915	0.3087	0.2437	0.4771	0.3226	0.0844	0.4091	0.1399
10	0.7579	0.6903	0.7225	0.7814	0.6706	0.7218	0.5360	0.6427	0.5845

Πίνακας ΙΙ.5: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Γερμανία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7801	0.3412	0.4747	0.6722	0.3904	0.4939	0.4938	0.3510	0.4103
2	0.3079	0.5177	0.3861	0.6000	0.3579	0.4484	0.5421	0.3411	0.4187
3	0.4573	0.4438	0.4505	0.0701	0.4510	0.1214	0.2409	0.3405	0.2821
4	0.6227	0.6592	0.6404	0.6016	0.6064	0.6040	0.4776	0.5934	0.5292
5	0.2621	0.7500	0.3885	0.3398	0.7609	0.4698	0.3786	0.5493	0.4483
6	0.7807	0.8380	0.8083	0.8571	0.7933	0.8240	0.7767	0.7959	0.7862
7	0.1736	0.2500	0.2049	0.0208	0.3000	0.0390	0.0556	0.2222	0.0889
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7386	0.3242	0.4506	0.5892	0.4201	0.4905	0.5270	0.4847	0.5050
2	0.2974	0.4414	0.3553	0.7237	0.3429	0.4653	0.6237	0.3977	0.4857
3	0.2226	0.4506	0.2980	0	0	0	0.3659	0.4898	0.4188
4	0.5224	0.5910	0.5546	0.6227	0.6396	0.6310	0.5937	0.6777	0.6329
5	0.4466	0.6479	0.5287	0	0	0	0.4563	0.5	0.4772
6	0.8330	0.7901	0.8110	0.9034	0.7975	0.8472	0.8934	0.8177	0.8538
7	0.2500	0.2057	0.2257	0	0	0	0	0	0
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5311	0.4885	0.5089	0.5436	0.5574	0.5504	0.4108	0.4950	0.4490
2	0.5974	0.4374	0.5050	0.6500	0.4540	0.5346	0.3421	0.3439	0.3430
3	0.3780	0.5061	0.4328	0.3445	0.5330	0.4185	0.5000	0.3071	0.3805
4	0.6201	0.6088	0.6144	0.6359	0.6276	0.6317	0.4512	0.5552	0.4978
5	0.4854	0.6329	0.5495	0.4854	0.6494	0.5556	0.3689	0.3333	0.3502
6	0.9034	0.7905	0.8432	0.9074	0.7749	0.836	0.8350	0.7743	0.8035
7	0.0417	0.4615	0.0764	0.0694	0.2632	0.1099	0	0	0

Πίνακας II.6: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Δανία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2919	0.4925	0.3663	0.6291	0.3860	0.4785	0.4957	0.3671	0.4218
2	0.5103	0.4763	0.4927	0.4013	0.5460	0.4626	0.3787	0.3959	0.3871
3	0.7948	0.2014	0.3213	0.3333	0.3982	0.3629	0.1897	0.3304	0.2410
4	0.3955	0.5779	0.4696	0.3429	0.5051	0.4084	0.3554	0.4431	0.3945
5	0.1735	0.6101	0.2702	0.3189	0.4978	0.3888	0.3715	0.2817	0.3204
6	0.6132	0.7837	0.6880	0.8588	0.6616	0.7474	0.7047	0.7041	0.7044
7	0.5419	0.5483	0.5451	0.2323	0.7108	0.3502	0.4632	0.5778	0.5142
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3511	0.4676	0.4011	0.7251	0.3700	0.4899	0.3959	0.4981	0.4412
2	0.2538	0.5330	0.3438	0.1463	0.3628	0.2086	0.6246	0.4200	0.5022
3	0.7481	0.1867	0.2988	0	0	0	0	0	0
4	0.2848	0.5777	0.3815	0.02373	0.6288	0.0457	0.4598	0.5234	0.4896
5	0.1289	0.6370	0.2144	0.4269	0.3139	0.3618	0.3794	0.3509	0.3646
6	0.7912	0.6765	0.7294	0.9064	0.6295	0.7430	0.8858	0.6516	0.7508
7	0.5661	0.5696	0.5678	0.5708	0.5434	0.5568	0.4907	0.6999	0.5769
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4971	0.4768	0.4867	0.4950	0.4703	0.4823	0.4214	0.4128	0.4171
2	0.5353	0.4788	0.5055	0.5278	0.4869	0.5065	0.4309	0.4397	0.4352
3	0.2438	0.4617	0.3191	0.2823	0.4231	0.3387	0.2923	0.3119	0.3018
4	0.4392	0.5480	0.4876	0.4281	0.5651	0.4871	0.4681	0.3746	0.4162
5	0.3197	0.5092	0.3927	0.3139	0.5082	0.3881	0.3218	0.4042	0.3583
6	0.8773	0.6578	0.7519	0.8838	0.6574	0.7540	0.7298	0.6654	0.6961
7	0.5438	0.6310	0.5841	0.5424	0.6433	0.5885	0.1620	0.3039	0.2114

Πίνακας ΙΙ.7: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Εσθονία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2933	0.4133	0.3431	0.0715	0.3968	0.1212	0.4936	0.3018	0.3746
2	0.8446	0.2582	0.3955	0.5100	0.5792	0.5424	0.3825	0.6713	0.4873
3	0.3246	0.3024	0.3131	0.1152	0.5238	0.1888	0.1571	0.4688	0.2353
4	0.3578	0.5159	0.4226	0.7022	0.3777	0.4912	0.3395	0.4397	0.3831
5	0.6201	0.3342	0.4343	0.1234	0.5033	0.1982	0.4351	0.3363	0.3793
6	0.0476	0.3636	0.0841	0.4901	0.2914	0.3655	0.2285	0.3128	0.2641
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.1645	0.4440	0.2401	0.0014	0.1429	0.0028	0.2575	0.4369	0.3240
2	0.6693	0.4590	0.5446	0.5976	0.6148	0.6061	0.5458	0.6256	0.5830
3	0.4241	0.2382	0.3051	0	0	0	0.2042	0.4699	0.2847
4	0.0282	0.4792	0.0532	0.7341	0.4125	0.5282	0.6618	0.4423	0.5302
5	0.4545	0.3660	0.4055	0.4854	0.4769	0.4811	0.4805	0.5086	0.4942
6	0.5086	0.2481	0.3335	0.4465	0.3380	0.3847	0.3646	0.3395	0.3516
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3004	0.3763	0.3341	0.2461	0.4257	0.3119	0.3720	0.3338	0.3518
2	0.5857	0.6000	0.5927	0.6335	0.6023	0.6175	0.5538	0.2465	0.3411
3	0.2199	0.4468	0.2947	0.2356	0.4369	0.3061	0.4817	0.1874	0.2698
4	0.5699	0.4189	0.4829	0.6556	0.4396	0.5263	0.4583	0.4083	0.4319
5	0.4756	0.4819	0.4788	0.4692	0.4801	0.4745	0.1023	0.1816	0.1308
6	0.3382	0.3580	0.3478	0.3554	0.3635	0.3594	0.0621	0.2017	0.0950

Πίνακας ΙΙ.8: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Αγγλία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7719	0.4540	0.5718	0.7281	0.4688	0.5704	0.5797	0.4644	0.5157
2	0.2227	0.5166	0.3112	0.8176	0.3964	0.534	0.5565	0.4051	0.4689
3	0.3190	0.4200	0.3626	0.07707	0.6298	0.1373	0.1864	0.3381	0.2403
4	0.7702	0.5558	0.6457	0.1621	0.7565	0.267	0.4155	0.5813	0.4846
5	0.5610	0.8563	0.6779	0.8081	0.7887	0.7983	0.8276	0.7476	0.7856
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4921	0.5885	0.5360	0.6646	0.5326	0.5914	0.6563	0.5425	0.5940
2	0.2667	0.4841	0.3440	0.6509	0.4182	0.5092	0.6396	0.4525	0.5300
3	0.0757	0.6353	0.1352	0	0	0	0.2200	0.5228	0.3097
4	0.8391	0.4984	0.6253	0.4754	0.6478	0.5484	0.4744	0.6577	0.5512
5	0.7974	0.7418	0.7686	0.8831	0.7604	0.8171	0.8816	0.7635	0.8183
Κόμμα	NN7			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6401	0.5664	0.6010	0.6248	0.5846	0.6040	0.3590	0.6170	0.4539
2	0.6255	0.4556	0.5272	0.6638	0.4511	0.5372	0.2270	0.4360	0.2986
3	0.2429	0.4962	0.3261	0.2368	0.509	0.3232	0.5063	0.2350	0.3210
4	0.4911	0.6517	0.5601	0.4597	0.6711	0.5457	0.6017	0.5346	0.5662
5	0.8882	0.7586	0.8183	0.9022	0.7445	0.8158	0.7893	0.7512	0.7698

Πίνακας Π.9: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Βόρεια Ιρλανδία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3175	0.4651	0.3774	0.1429	0.3600	0.2045	0.5873	0.4157	0.4868
2	0.6000	0.7075	0.6493	0.6400	0.6202	0.6299	0.6400	0.5634	0.5993
3	0.3494	0.3152	0.3314	0.3976	0.2845	0.3317	0.2771	0.2644	0.2706
4	0.3333	0.3000	0.3158	0.5926	0.3200	0.4156	0.2593	0.3590	0.3011
5	0.5097	0.5064	0.5080	0.5097	0.5000	0.5048	0.4581	0.4410	0.4494
6	0.5200	0.4483	0.4815	0.3100	0.4921	0.3804	0.2300	0.3151	0.2659
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4603	0.4328	0.4462	0.8095	0.3893	0.5258	0.6032	0.4578	0.5205
2	0.776	0.6599	0.7132	0.8240	0.6821	0.7464	0.7200	0.6716	0.6950
3	0.0964	0.3200	0.1481	0.3494	0.3816	0.3648	0.2771	0.3151	0.2949
4	0.2963	0.3478	0.3200	0	0	0	0.3333	0.3396	0.3364
5	0.6258	0.4199	0.5026	0.6839	0.4274	0.5261	0.6194	0.4229	0.5026
6	0.3200	0.3556	0.3368	0	0	0	0.2000	0.5556	0.2941
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4921	0.4247	0.4559	0.5079	0.4156	0.4571	0.2698	0.4359	0.3333
2	0.7840	0.7000	0.7396	0.816	0.6846	0.7445	0.7760	0.4641	0.5808
3	0.3253	0.4030	0.3600	0.2892	0.4706	0.3582	0.4699	0.2349	0.3133
4	0.3148	0.3542	0.3333	0.2963	0.4000	0.3404	0.6667	0.3051	0.4186
5	0.6258	0.4575	0.5286	0.7290	0.4520	0.5580	0.1871	0.6591	0.2915
6	0.2000	0.4348	0.2740	0.2000	0.5714	0.2963	0	0	0

Πίνακας ΙΙ.10: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Σκωτία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4850	0.3687	0.4189	0.4352	0.4906	0.4613	0.3987	0.3509	0.3733
2	0.7146	0.4973	0.5865	0.9004	0.3599	0.5142	0.6054	0.5032	0.5496
3	0.4848	0.6316	0.5486	0.6768	0.4241	0.5214	0.4949	0.5269	0.5104
4	0.5923	0.1552	0.2460	0.0308	0.1905	0.0530	0.1538	0.4082	0.2235
5	0.1256	0.7823	0.2165	0.0492	0.6032	0.09102	0.5505	0.6037	0.5759
6	0.5316	0.6885	0.6000	0.6203	0.5568	0.5868	0.5696	0.5172	0.5422
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5748	0.3453	0.4314	0.5083	0.4048	0.4507	0.6179	0.4035	0.4882
2	0.7625	0.4854	0.5931	0.8257	0.5000	0.6228	0.7548	0.5786	0.6550
3	0.6162	0.3315	0.4311	0.3131	0.6739	0.4276	0.6263	0.5536	0.5877
4	0.0231	0.2000	0.0414	0	0	0	0	0	0
5	0.1852	0.7296	0.2955	0.4326	0.6588	0.5223	0.5039	0.6885	0.5819
6	0.8228	0.3476	0.4887	0.7089	0.5091	0.5926	0.6329	0.5952	0.6135
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5050	0.4457	0.4735	0.5382	0.4309	0.4786	0.2392	0.4417	0.3103
2	0.7644	0.5684	0.6520	0.7548	0.5645	0.6459	0.1782	0.6242	0.2772
3	0.6364	0.5526	0.5915	0.6061	0.5607	0.5825	0.4141	0.5467	0.4713
4	0.1154	0.3191	0.1695	0.1846	0.3636	0.2449	0.4615	0.2076	0.2864
5	0.5285	0.6678	0.5900	0.5026	0.6867	0.5804	0.7850	0.5293	0.6322
6	0.6835	0.6136	0.6467	0.6835	0.5934	0.6353	0.5443	0.5244	0.5342

Πίνακας Π.11: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ουαλία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4906	0.4860	0.4883	0.7170	0.3666	0.4851	0.6384	0.3684	0.4672
2	0.4239	0.3451	0.3805	0.5652	0.3059	0.3969	0.2717	0.3676	0.3125
3	0.3439	0.5588	0.4258	0.2670	0.5315	0.3554	0.2805	0.4921	0.3573
4	0.3671	0.1726	0.2348	0.0253	0.2222	0.0454	0.1013	0.1778	0.129
5	0.4397	0.4728	0.4556	0.0311	0.5000	0.0586	0.2568	0.4648	0.3308
6	0.6744	0.7016	0.6877	0.8140	0.6034	0.6931	0.7829	0.6012	0.6801
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3522	0.4195	0.3829	0.7987	0.3769	0.5121	0.5943	0.4479	0.5108
2	0.4348	0.2837	0.3433	0.2826	0.5532	0.3741	0.4674	0.2986	0.3644
3	0.3756	0.5804	0.4560	0.4299	0.6169	0.5067	0.4253	0.6225	0.5054
4	0.0380	0.2500	0.0659	0	0	0	0	0	0
5	0.4358	0.4534	0.4444	0	0	0	0.3696	0.4974	0.4241
6	0.9380	0.4144	0.5748	0.9147	0.5198	0.6629	0.8760	0.5825	0.6997
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5472	0.400	0.4622	0.6415	0.4331	0.5171	0.1887	0.4225	0.2609
2	0.4674	0.4057	0.4343	0.4457	0.4659	0.4556	0.4022	0.4205	0.4111
3	0.3439	0.608	0.4393	0.4932	0.6193	0.5491	0.3937	0.3734	0.3833
4	0.1392	0.2500	0.1789	0.0886	0.2059	0.1239	0.2152	0.1278	0.1604
5	0.3619	0.4769	0.4115	0.2685	0.5897	0.369	0.5253	0.3629	0.4293
6	0.8682	0.5685	0.6871	0.9147	0.5463	0.6841	0.6434	0.6484	0.6459

Πίνακας ΙΙ.12: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ισπανία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.8701	0.3351	0.4839	0.7073	0.4841	0.5748	0.6181	0.2624	0.3684
2	0.1295	0.3345	0.1867	0.2905	0.2379	0.2616	0.3874	0.1901	0.2550
3	0.0960	0.1043	0.1000	0.6100	0.2741	0.3783	0.2654	0.2815	0.2732
4	0.2272	0.1356	0.1698	0.0010	0.0882	0.0020	0.0957	0.1661	0.1214
5	0.0391	0.5454	0.0729	0.0374	0.5000	0.0696	0.2864	0.3559	0.3174
6	0.1740	0.4376	0.2490	0.2321	0.3967	0.2929	0.2756	0.3527	0.3094
7	0.3560	0.0193	0.0366	0.0428	0.0163	0.0236	0.0098	0.0184	0.0127
8	0.2343	0.3710	0.2872	0.4736	0.3154	0.3787	0.2874	0.3030	0.2950
9	0.2744	0.2280	0.2490	0.0122	0.7143	0.0240	0.1573	0.3086	0.2084
10	0.0281	0.7121	0.0541	0.0006	0.0303	0.0012	0.1677	0.4000	0.2363
11	0.3511	0.1690	0.2281	0.0739	0.0250	0.0373	0.0883	0.1842	0.1194
12	0.1144	0.2839	0.1631	0.0005	0.1818	0.0009	0.0882	0.2136	0.1249
Κόμμα	Mahalanobis			NN5			NN10		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4602	0.5954	0.5191	0.7202	0.4987	0.5893	0.6641	0.5258	0.5869
2	0.0560	0.4497	0.1057	0.3708	0.2630	0.3078	0.3946	0.2591	0.3128
3	0.3705	0.2901	0.3254	0.5778	0.2501	0.3491	0.5475	0.2526	0.3457
4	0.4092	0.1319	0.1995	0	0	0	0	0	0
5	0.2791	0.3035	0.2908	0	0	0	0.3385	0.4522	0.3872
6	0.1982	0.4206	0.2694	0.2524	0.4076	0.3117	0.2489	0.4328	0.3160
7	0.1712	0.0242	0.0424	0	0	0	0	0	0
8	0.0336	0.4982	0.0630	0.5095	0.313	0.3877	0.4774	0.3183	0.3820
9	0.0963	0.3970	0.1551	0	0	0	0.0488	0.3604	0.0859
10	0.5695	0.215	0.3121	0.3862	0.3455	0.3647	0.4084	0.3694	0.3879
11	0.2361	0.1566	0.1883	0	0	0	0.0996	0.2304	0.1391
12	0.3310	0.1631	0.2185	0	0	0	0.0619	0.3271	0.1042
Κόμμα	NN15			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6638	0.5198	0.5830	0.7366	0.5121	0.6041	0.2825	0.5921	0.3825
2	0.3720	0.2852	0.3229	0.3782	0.3133	0.3427	0.1692	0.1314	0.1479
3	0.5519	0.2502	0.3443	0.4459	0.3706	0.4048	0.1098	0.3056	0.1615
4	0.0060	0.2466	0.0118	0.0597	0.3274	0.1009	0	0	0
5	0.3596	0.4115	0.3838	0.3302	0.4993	0.3975	0.1094	0.1445	0.1245
6	0.2692	0.4210	0.3284	0.3065	0.4237	0.3557	0.0155	0.0039	0.0062
7	0	0	0	0.0022	0.3333	0.0043	0	0	0
8	0.4692	0.3219	0.3819	0.561	0.3155	0.4039	0.0545	0.3238	0.0933
9	0.2500	0.3400	0.2881	0.2953	0.3583	0.3238	0.0078	0.0119	0.0094
10	0.3539	0.3873	0.3698	0.4498	0.4293	0.4393	0.3383	0.2572	0.2923
11	0.1037	0.2457	0.1458	0.1481	0.228	0.1796	0.1354	0.0181	0.0319
12	0.0650	0.3157	0.1078	0.1025	0.3619	0.1598	0.2257	0.0485	0.0798

Πίνακας II.13: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Φινλανδία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5788	0.7348	0.6475	0.8185	0.5085	0.6273	0.726	0.5222	0.6074
2	0.2883	0.2623	0.2747	0.2162	0.3038	0.2526	0.1712	0.2111	0.1891
3	0.6979	0.8243	0.7559	0.762	0.7691	0.7655	0.5675	0.8267	0.6730
4	0.3153	0.1824	0.2311	0.1196	0.1964	0.1486	0.1304	0.2222	0.1644
5	0.4685	0.6624	0.5488	0.5360	0.5064	0.5208	0.5270	0.4718	0.4979
6	0.6026	0.5391	0.5691	0.3799	0.5370	0.4450	0.6026	0.5149	0.5553
7	0.2500	0.1053	0.1481	0	0	0	0.2000	0.1270	0.1553
8	0.3333	0.1273	0.1842	0.0476	0.1111	0.0667	0.0476	0.0667	0.0556
Κόμμα	Mahalanobis			NN5			NN10		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4623	0.6750	0.5488	0.8082	0.5856	0.6791	0.7774	0.5896	0.6706
2	0.0090	0.2500	0.0174	0.1261	0.2593	0.1697	0.2162	0.2609	0.2365
3	0.9794	0.5232	0.6821	0.7826	0.8162	0.7991	0.7414	0.8265	0.7817
4	0.0109	0.3333	0.0211	0.1522	0.2917	0.2	0.1630	0.2308	0.1911
5	0.5991	0.5076	0.5496	0.5135	0.6162	0.5602	0.4189	0.6739	0.5167
6	0.3493	0.5926	0.4396	0.6769	0.4627	0.5496	0.7293	0.4758	0.5759
7	0.1000	0.1818	0.1290	0	0	0	0	0	0
8	0	0	0	0	0	0	0.1905	0.1905	0.1905
Κόμμα	NN15			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7979	0.5508	0.6517	0.8048	0.5980	0.6861	0.3048	0.6846	0.4218
2	0.1171	0.3171	0.1711	0.1802	0.3774	0.2439	0.4595	0.2152	0.2931
3	0.746	0.8295	0.7855	0.8146	0.8036	0.8091	0.4943	0.7687	0.6017
4	0.1522	0.1918	0.1697	0.2391	0.4000	0.2993	0.0870	0.0635	0.0734
5	0.5000	0.5842	0.5388	0.455	0.6392	0.5316	0.6081	0.4397	0.5104
6	0.6725	0.4768	0.5580	0.7074	0.5000	0.5859	0.4148	0.4774	0.4439
7	0	0	0	0.0750	0.7500	0.1364	0.2000	0.0870	0.1212
8	0	0	0	0.2381	0.3571	0.2857	0.2381	0.06944	0.1075

Πίνακας ΙΙ.14: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Γαλλία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4792	0.6715	0.5593	0.4115	0.6583	0.5064	0.4948	0.4822	0.4884
2	0.2654	0.5177	0.3510	0.5236	0.4068	0.4579	0.3055	0.3889	0.3422
3	0.3961	0.5153	0.4479	0.1569	0.5263	0.2417	0.3725	0.4095	0.3901
4	0.8473	0.3807	0.5254	0.6873	0.4355	0.5331	0.4909	0.3879	0.4334
5	0.2140	0.5733	0.3116	0.5224	0.4688	0.4941	0.4677	0.3900	0.4253
6	0.4167	0.1923	0.2632	0.1250	0.2727	0.1714	0.0833	0.2222	0.1212
7	0.7473	0.8015	0.7735	0.8114	0.7677	0.7889	0.7758	0.7927	0.7842
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5417	0.6624	0.5960	0.6771	0.4943	0.5714	0.5104	0.6806	0.5833
2	0.1164	0.5333	0.1910	0.5964	0.3933	0.474	0.4909	0.5056	0.4982
3	0.4784	0.4586	0.4683	0.1216	0.4769	0.1938	0.4667	0.5434	0.5021
4	0.0509	0.5385	0.0930	0.0727	0.3922	0.1227	0.4836	0.4981	0.4908
5	0.7811	0.2679	0.3990	0.7015	0.3606	0.4764	0.6517	0.4338	0.5209
6	0.0208	0.5000	0.0400	0	0	0	0	0	0
7	0.9324	0.6093	0.7370	0.9039	0.7471	0.8180	0.8897	0.7622	0.8210
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5573	0.6221	0.5879	0.5625	0.6171	0.5886	0.2969	0.7703	0.4286
2	0.4218	0.5000	0.4576	0.4509	0.4921	0.4706	0.3564	0.3203	0.3373
3	0.4667	0.5064	0.4857	0.5059	0.542	0.5233	0.6745	0.3532	0.4636
4	0.4509	0.4921	0.4706	0.4582	0.5551	0.5020	0.1600	0.5432	0.2472
5	0.6418	0.4243	0.5109	0.6418	0.451	0.5298	0.5771	0.4280	0.4915
6	0	0	0	0.0833	0.3636	0.1356	0.1042	0.0526	0.0699
7	0.8826	0.7492	0.8105	0.8932	0.7426	0.8110	0.5623	0.7418	0.6397

Πίνακας II.15: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ελλάδα

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3807	0.6176	0.4711	0.6448	0.5159	0.5732	0.6504	0.4282	0.5164
2	0.5150	0.7434	0.6085	0.8547	0.5862	0.6955	0.7423	0.5921	0.6587
3	0.0799	0.2252	0.1179	0	0	0	0.1246	0.1653	0.1421
4	0.4593	0.1390	0.2134	0.0052	0.1176	0.0100	0.1129	0.1937	0.1426
5	0.4451	0.4057	0.4245	0.06113	0.6724	0.1121	0.3009	0.4885	0.3725
6	0.4935	0.5971	0.5404	0.5844	0.6034	0.5938	0.4731	0.6281	0.5397
7	0.6341	0.0880	0.1546	0	0	0	0.0731	0.1511	0.0986
8	0.0691	0.3542	0.1156	0.4948	0.3831	0.4318	0.2439	0.3429	0.2850
9	0.6612	0.2539	0.3670	0.0049	1	0.0098	0.1678	0.3400	0.2247
	Mahalanobis			NN5			NN10		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4759	0.4932	0.4844	0.7167	0.5167	0.6005	0.7122	0.5112	0.5952
2	0.2848	0.8037	0.4205	0.8045	0.6339	0.7091	0.8066	0.6396	0.7134
3	0.1581	0.2920	0.2052	0	0	0	0.0415	0.3714	0.0747
4	0.1024	0.0736	0.0856	0	0	0	0	0	0
5	0.4953	0.3261	0.3933	0.2900	0.5929	0.3895	0.2947	0.6551	0.4065
6	0.9378	0.2383	0.3800	0.6020	0.6199	0.6108	0.6410	0.6040	0.6220
7	0.0453	0.2653	0.0774	0	0	0	0	0	0
8	0.3258	0.4183	0.3663	0.4988	0.3842	0.4341	0.4733	0.4043	0.4361
9	0.1201	0.6033	0.2003	0.1447	0.5906	0.2325	0.2072	0.5207	0.2965
	NN15			SVM			Party Coding		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6992	0.5216	0.5974	0.6924	0.5274	0.5987	0.4130	0.5432	0.4693
2	0.8004	0.6444	0.7140	0.8019	0.6439	0.7143	0.4390	0.7197	0.5453
3	0.0479	0.3448	0.0841	0.0831	0.3444	0.1338	0.2364	0.1503	0.1837
4	0	0	0	0.0551	0.4667	0.0986	0.4331	0.1375	0.2087
5	0.2962	0.6655	0.4100	0.2994	0.7375	0.4259	0.4890	0.2581	0.3378
6	0.6494	0.5993	0.6233	0.6744	0.5774	0.6222	0.3553	0.5511	0.4320
7	0.0070	0.1333	0.0132	0.0488	0.2979	0.0838	0.3345	0.0629	0.1059
8	0.4768	0.3968	0.4331	0.4669	0.4028	0.4325	0.0842	0.3040	0.1319
9	0.2368	0.4898	0.3193	0.1760	0.5377	0.2652	0.4688	0.2867	0.3558

Πίνακας ΙΙ.16: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Κροατία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5187	0.5045	0.5115	0.5467	0.4699	0.5054	0.514	0.3406	0.4097
2	0.3692	0.6000	0.4571	0.6359	0.5849	0.6093	0.5692	0.5578	0.5635
3	0.7083	0.3228	0.4435	0.2917	0.4773	0.3621	0.25	0.3529	0.2927
4	0.5514	0.1947	0.2878	0.2150	0.1756	0.1933	0.0935	0.2222	0.1316
5	0.3405	0.6961	0.4573	0.5324	0.6016	0.5649	0.5492	0.5917	0.5697
Mahalanobis			NN3			NN5			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.1916	0.6212	0.2929	0.5467	0.4875	0.5154	0.5607	0.4633	0.5074
2	0.6256	0.5351	0.5768	0.7128	0.5673	0.6318	0.6821	0.5684	0.6200
3	0.1667	0.3636	0.2286	0.1528	0.3929	0.2200	0.2500	0.4186	0.3130
4	0.0280	0.1875	0.0488	0	0	0	0.1121	0.2264	0.1500
5	0.8753	0.5514	0.6766	0.7218	0.6118	0.6623	0.6523	0.6538	0.6531
NN10			SVM			Party Coding			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5187	0.4066	0.4559	0.6262	0.4201	0.5028	0.4953	0.4398	0.4659
2	0.6821	0.6244	0.652	0.6462	0.5972	0.6207	0.2923	0.6951	0.4116
3	0.3056	0.5000	0.3793	0.1667	0.3636	0.2286	0.6944	0.3401	0.4566
4	0.0467	0.2083	0.0763	0.0467	0.2000	0.0758	0.5327	0.1851	0.2747
5	0.6643	0.6142	0.6382	0.6259	0.6259	0.6259	0.3381	0.6211	0.4379

Πίνακας ΙΙ.17: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ουγγαρία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7686	0.5057	0.6100	0.7162	0.6029	0.6547	0.6943	0.5300	0.6011
2	0.7672	0.8240	0.7946	0.7957	0.846	0.8201	0.7078	0.8347	0.7661
3	0.3380	0.5748	0.4257	0.3657	0.4489	0.4031	0.3796	0.408	0.3933
4	0.2287	0.5728	0.3269	0.7636	0.4592	0.5735	0.4186	0.4635	0.4399
5	0.7461	0.3170	0.4450	0.1308	0.4146	0.1988	0.2769	0.3711	0.3172
Mahalanobis			NN3			NN5			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3100	0.6339	0.4164	0.6550	0.6410	0.6479	0.6201	0.6425	0.6311
2	0.9691	0.5965	0.7385	0.829	0.7968	0.8126	0.8171	0.8075	0.8123
3	0.1296	0.5185	0.2074	0.4722	0.5075	0.4892	0.5463	0.4155	0.4720
4	0.7287	0.4688	0.5706	0.7093	0.4085	0.5184	0.6783	0.4487	0.5401
5	0.2077	0.3971	0.2727	0	0	0	0	0	0
NN10			SVM			Party Coding			
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6812	0.6047	0.6407	0.6638	0.7103	0.6862	0.5066	0.464	0.4843
2	0.7886	0.8579	0.8218	0.8646	0.8180	0.8406	0.5416	0.8201	0.6524
3	0.5231	0.3553	0.4232	0.5231	0.4708	0.4956	0.6528	0.3169	0.4266
4	0.4380	0.3831	0.4087	0.7054	0.4703	0.5643	0.0271	0.3889	0.0507
5	0.1462	0.3016	0.1969	0.1000	0.3714	0.1576	0.5538	0.2384	0.3333

Πίνακας Π.18: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ιρλανδία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4630	0.4643	0.4636	0.7389	0.3646	0.4882	0.6083	0.3203	0.4197
2	0.5444	0.3310	0.4117	0.3668	0.3626	0.3647	0.4595	0.2439	0.3186
3	0.4136	0.2762	0.3312	0.1257	0.2474	0.1667	0.1937	0.2176	0.2050
4	0.1980	0.4552	0.2760	0.4870	0.3513	0.4082	0.3377	0.3824	0.3586
5	0.3469	0.2193	0.2688	0.0102	0.1111	0.0187	0.1224	0.1765	0.1446
6	0.2669	0.4375	0.3315	0.1525	0.5000	0.2337	0.1965	0.3490	0.2514
7	0.2411	0.3266	0.2774	0.2857	0.2759	0.2807	0.0744	0.2841	0.1179
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2611	0.4706	0.3359	0.7359	0.3798	0.5010	0.6202	0.4130	0.4958
2	0.0039	0.5000	0.0077	0.5174	0.2894	0.3712	0.39	0.3870	0.3885
3	0.6230	0.2117	0.3161	0	0	0	0.3089	0.3010	0.3049
4	0.2727	0.4330	0.3347	0.6753	0.3432	0.4551	0.5779	0.3973	0.4709
5	0.2347	0.3239	0.2722	0	0	0	0	0	0
6	0.4751	0.3885	0.4274	0.2874	0.4623	0.3544	0.4194	0.3865	0.4023
7	0.3780	0.2555	0.3049	0	0	0	0.1429	0.3137	0.1963
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6855	0.4118	0.5145	0.6528	0.4215	0.5122	0.3561	0.4027	0.3780
2	0.5019	0.3421	0.4069	0.4131	0.3603	0.3849	0.4788	0.2157	0.2974
3	0.2408	0.3172	0.2738	0.2618	0.3704	0.3067	0.3770	0.1989	0.2604
4	0.461	0.3890	0.4220	0.5714	0.3801	0.4565	0.4221	0.3125	0.3591
5	0.1122	0.2619	0.1571	0.1531	0.3333	0.2098	0.3061	0.1775	0.2247
6	0.2874	0.4667	0.3557	0.3460	0.4453	0.3894	0.0645	0.5116	0.1146
7	0.2262	0.3290	0.2681	0.2143	0.3495	0.2657	0.0060	0.2857	0.0117

Πίνακας ΙΙ.19: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ιταλία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2857	0.1854	0.2249	0.1429	0.3173	0.1970	0.2641	0.1728	0.2089
2	0.2935	0.1947	0.2341	0.1493	0.2703	0.1923	0.1642	0.1823	0.1728
3	0.2176	0.3881	0.2788	0.0711	0.4359	0.1223	0.2008	0.2892	0.2370
4	0.5577	0.6532	0.6016	0.7471	0.5379	0.6255	0.6412	0.5291	0.5798
5	0.4779	0.0982	0.1629	0.0885	0.2439	0.1299	0.0797	0.1800	0.1104
6	0.5527	0.6389	0.5927	0.8607	0.5084	0.6393	0.6164	0.5743	0.5946
7	0.3270	0.3856	0.3539	0.0304	0.6154	0.0580	0.1635	0.4057	0.2331
8	0.6305	0.6340	0.6322	0.1478	0.8938	0.2536	0.4942	0.5909	0.5383
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3680	0.1197	0.1807	0	0	0	0	0	0
2	0.1144	0.2674	0.1603	0	0	0	0.0846	0.2394	0.1250
3	0.5063	0.2516	0.3361	0.3808	0.3655	0.3730	0.4059	0.3593	0.3811
4	0.5504	0.6431	0.5932	0.7442	0.5849	0.6550	0.7313	0.5783	0.6459
5	0.0708	0.3810	0.1194	0	0	0	0	0	0
6	0.7342	0.5853	0.6513	0.7601	0.5845	0.6608	0.761	0.6153	0.6804
7	0.1939	0.4636	0.2735	0	0	0	0.1293	0.4474	0.2006
8	0.4088	0.7409	0.5269	0.5591	0.6655	0.6076	0.5724	0.6878	0.6248
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.0909	0.2838	0.1377	0.1342	0.2925	0.1840	0.1775	0.1524	0.1640
2	0.2090	0.2561	0.2301	0.2090	0.2745	0.2373	0.1244	0.2525	0.1667
3	0.318	0.4294	0.3654	0.3473	0.4716	0.4000	0.3473	0.2814	0.3109
4	0.7262	0.6109	0.6636	0.7341	0.6008	0.6608	0.5634	0.6020	0.5821
5	0.0619	0.2121	0.0959	0.1770	0.3704	0.2395	0.3009	0.1318	0.1833
6	0.7548	0.6160	0.6784	0.7557	0.5993	0.6685	0.3436	0.6436	0.4480
7	0.1825	0.5161	0.2697	0.1901	0.5952	0.2882	0.2776	0.2303	0.2517
8	0.5784	0.6910	0.6297	0.5245	0.7081	0.6026	0.7820	0.4946	0.6060

Πίνακας II.20: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Λιθουανία

		HMM			Naïve Bayes			KNN		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1	
1	0.7279	0.1147	0.19820	0.1029	0.2642	0.1481	0.2574	0.2333	0.2448	
2	0.1555	0.5373	0.2412	0.5983	0.4301	0.5005	0.5032	0.4989	0.5011	
3	0.2898	0.7370	0.4160	0.6190	0.6701	0.6436	0.6639	0.6313	0.6472	
4	0.5000	0.3064	0.3800	0.2500	0.5588	0.3455	0.1579	0.6000	0.2500	
		Mahalanobis			NN3			NN5		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1	
1	0.6250	0.1512	0.2436	0.1397	0.3800	0.2043	0.1838	0.3425	0.2392	
2	0.2073	0.5680	0.3038	0.5443	0.5780	0.5606	0.5335	0.5222	0.5278	
3	0.3714	0.7339	0.4932	0.7850	0.6564	0.7150	0.7293	0.6466	0.6854	
4	0.4211	0.1042	0.1671	0.3158	0.5333	0.3967	0.2632	0.5714	0.3604	
		NN10			SVM			Party Coding		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1	
1	0.1397	0.3800	0.2043	0.1838	0.4032	0.2525	0.3088	0.2545	0.2791	
2	0.5054	0.4825	0.4937	0.5184	0.5647	0.5405	0.6156	0.4974	0.5502	
3	0.7075	0.6303	0.6667	0.7864	0.6576	0.7162	0.5483	0.7120	0.6195	
4	0.3421	0.5200	0.4127	0.3026	0.5227	0.3833	0.4474	0.3208	0.3736	

Πίνακας II.21: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Λετονία

		HMM			Naïve Bayes			KNN		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1	
1	0.7679	0.8600	0.8113	0.7500	0.7778	0.7636	0.7500	0.8077	0.7778	
2	0.7113	0.7914	0.7492	0.9077	0.7403	0.8155	0.8274	0.7297	0.7755	
3	0.5662	0.4375	0.4936	0.2794	0.6129	0.3838	0.2941	0.4211	0.3463	
		Mahalanobis			NN3			NN5		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1	
1	0.6429	0.7660	0.6990	0.8571	0.8421	0.8496	0.8036	0.7500	0.7759	
2	0.8185	0.7161	0.7639	0.8512	0.7627	0.8045	0.6964	0.7524	0.7233	
3	0.2868	0.4021	0.3348	0.3603	0.5104	0.4224	0.4265	0.3694	0.3959	
		NN10			SVM			Party Coding		
Κόμμα	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1	
1	0.6607	0.8810	0.7551	0.9286	0.8814	0.9043	0.6250	0.7447	0.6796	
2	0.6905	0.6905	0.6905	0.8958	0.7488	0.8157	0.6607	0.7708	0.7115	
3	0.3162	0.2867	0.3007	0.2647	0.5373	0.3547	0.5588	0.3938	0.4620	

Πίνακας ΙΙ.22: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ολλανδία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6000	0.3793	0.4648	0.6909	0.6786	0.6847	0.5273	0.5370	0.5321
2	0.7292	0.7447	0.7368	0.8542	0.8200	0.8367	0.7292	0.7955	0.7609
3	0.2667	0.6666	0.3809	0.4000	0.6316	0.4898	0.3667	0.3793	0.3729
4	0.6142	0.6903	0.6500	0.8425	0.7181	0.7754	0.6457	0.6406	0.6431
5	0.2500	0.5000	0.3333	0.3333	0.5714	0.4211	0.1667	0.3333	0.2222
6	0.6296	0.5574	0.5913	0.5556	0.6667	0.6061	0.5741	0.4769	0.5210
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2909	0.4324	0.3478	0.3636	0.4348	0.396	0.2727	0.4167	0.3297
2	0.6042	0.7838	0.6824	0.7708	0.6271	0.6916	0.625	0.6522	0.6383
3	0.0333	0.5000	0.0625	0	0	0	0.2667	0.4000	0.3200
4	0.5906	0.5474	0.5682	0.7953	0.6558	0.7189	0.7953	0.6433	0.7113
5	0.1667	0.0278	0.0476	0	0	0	0.4167	0.3846	0.4000
6	0.3519	0.4634	0.4000	0.7222	0.5821	0.6446	0.6296	0.6296	0.6296
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2364	0.3514	0.2826	0.4545	0.7143	0.5556	0.4000	0.6286	0.4889
2	0.3958	0.6129	0.4810	0.8750	0.8235	0.8485	0.7708	0.7551	0.7629
3	0.0667	0.1818	0.0976	0.3000	0.6923	0.4186	0.4000	0.4286	0.4138
4	0.685	0.5148	0.5878	0.8268	0.6481	0.7266	0.5039	0.6598	0.5714
5	0.3333	0.4444	0.3810	0.5000	0.6000	0.5455	0.2500	0.3333	0.2857
6	0.5185	0.4058	0.4553	0.6481	0.6364	0.6422	0.8333	0.4167	0.5556

Πίνακας Π.23: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Πολωνία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.1571	0.7671	0.2608	0.9166	0.5443	0.6830	0.7055	0.6184	0.6591
2	0.2321	0.6667	0.3443	0.4420	0.5711	0.4983	0.4131	0.5016	0.4531
3	0.1305	0.3941	0.1960	0.0279	0.5968	0.05331	0.2677	0.2769	0.2722
4	0.2761	0.4280	0.3356	0.0429	0.4537	0.0785	0.2656	0.3764	0.3114
5	0.4033	0.9553	0.5671	0.8796	0.7192	0.7914	0.7942	0.7193	0.7549
6	0.5997	0.2221	0.3242	0.0598	0.3429	0.1018	0.3887	0.3000	0.3386
7	0.0280	0.2253	0.0498	0.0052	0.3750	0.01036	0.0648	0.1869	0.0962
8	0.2913	0.2868	0.2891	0.1181	0.4412	0.1863	0.1732	0.5116	0.2588

Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.3232	0.7425	0.4504	0.8972	0.5611	0.6904	0.8556	0.6118	0.7134
2	0.4581	0.5063	0.4810	0.4574	0.5522	0.5004	0.4614	0.5729	0.5111
3	0.0468	0.5487	0.0862	0	0	0	0	0	0
4	0.7187	0.255	0.3764	0	0	0	0.3269	0.3914	0.3563
5	0.6907	0.8333	0.7553	0.8805	0.7251	0.7952	0.8782	0.7263	0.7950
6	0.3422	0.362	0.3518	0.3804	0.3491	0.3641	0.4037	0.3744	0.3885
7	0.2207	0.178	0.1970	0	0	0	0	0	0
8	0.1102	0.4667	0.1783	0	0	0	0	0	0

Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.8542	0.6231	0.7206	0.8208	0.6446	0.7221	0.4582	0.7157	0.5587
2	0.499	0.5826	0.5376	0.5023	0.5657	0.5321	0.4400	0.4659	0.4526
3	0.1116	0.567	0.1865	0.1961	0.5451	0.2884	0.1259	0.4207	0.1938
4	0.3751	0.4529	0.4104	0.3313	0.4489	0.3812	0.7064	0.263	0.3833
5	0.8602	0.7528	0.8029	0.8805	0.7297	0.7980	0.7408	0.8443	0.7892
6	0.4302	0.35	0.386	0.3555	0.3970	0.3751	0.1661	0.2703	0.2058
7	0.0088	0.1667	0.01664	0.06305	0.2553	0.1011	0.2504	0.0955	0.1384
8	0	0	0	0.1811	0.5111	0.2674	0	0	0

Πίνακας ΙΙ.24: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Πορτογαλία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7177	0.7658	0.7409	0.9196	0.5274	0.6703	0.7601	0.6834	0.7197
2	0.1600	0.5980	0.2525	0.4923	0.4138	0.4496	0.5050	0.3271	0.3970
3	0.2040	0.2645	0.2303	0.0105	0.2034	0.0200	0.1891	0.2248	0.2054
4	0.3119	0.3242	0.3179	0.4710	0.2553	0.3311	0.3167	0.2866	0.3009
5	0.2089	0.3694	0.2668	0.0568	0.3000	0.0954	0.2001	0.2812	0.2338
6	0.3308	0.1502	0.2066	0	0	0	0.0119	0.1481	0.0221
Κόμμα	Mahalanobis			NN5			NN10		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.6442	0.7805	0.7059	0.8519	0.6523	0.7389	0.8548	0.6393	0.7315
2	0.1735	0.6515	0.2740	0.4183	0.4796	0.4468	0.4358	0.4696	0.4520
3	0.0210	0.3934	0.0399	0.1506	0.2915	0.1986	0.0946	0.2872	0.1423
4	0.2337	0.3377	0.2763	0.5678	0.283	0.3777	0.5740	0.3070	0.4000
5	0.6579	0.2425	0.3544	0.2831	0.3666	0.3195	0.3268	0.3422	0.3343
6	0.1252	0.1922	0.1516	0	0	0	0	0	0
Κόμμα	NN15			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.8482	0.6479	0.7346	0.8841	0.6263	0.7332	0.4855	0.8543	0.6191
2	0.4385	0.4657	0.4517	0.4734	0.4493	0.4610	0.1735	0.4161	0.2449
3	0.1156	0.2966	0.1664	0.1103	0.3316	0.1656	0.6401	0.1980	0.3024
4	0.5567	0.3059	0.3948	0.5387	0.3061	0.3904	0.4094	0.2334	0.2973
5	0.3486	0.3512	0.3499	0.2860	0.3606	0.3190	0.0422	0.1530	0.0661
6	0	0	0	0.0030	0.2500	0.0059	0.0298	0.0930	0.0451

Πίνακας ΙΙ.25: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Ρουμανία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.8086	0.3851	0.5217	0.3666	0.6296	0.4634	0.6658	0.4901	0.5646
2	0.0255	0.7500	0.0493	0.8514	0.3655	0.5115	0.5732	0.4072	0.4762
3	0.2690	0.3348	0.2983	0.006897	0.4	0.01356	0.1103	0.2857	0.1592
4	0.4328	0.5088	0.4677	0.3313	0.5286	0.4073	0.3791	0.4922	0.4283
5	0.2791	0.6316	0.3871	0.2558	0.449	0.3259	0.3837	0.6111	0.4714
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.7520	0.5365	0.6263	0.6092	0.6402	0.6243	0.6146	0.6514	0.6325
2	0.1507	0.6455	0.2444	0.6773	0.4733	0.5572	0.6730	0.4641	0.5494
3	0.6069	0.2754	0.3789	0	0	0	0	0	0
4	0.0328	0.5500	0.06197	0.6299	0.4710	0.5390	0.6328	0.4753	0.5429
5	0.9419	0.2746	0.4252	0.7326	0.5431	0.6238	0.7442	0.5714	0.6465
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.5903	0.6385	0.6134	0.6442	0.662	0.653	0.4825	0.4891	0.4858
2	0.6518	0.4617	0.5405	0.6603	0.4607	0.5428	0.2654	0.5459	0.3571
3	0.1621	0.3534	0.2222	0.0172	0.2778	0.0325	0.2517	0.2704	0.2607
4	0.4896	0.4726	0.4809	0.6179	0.4894	0.5462	0.1552	0.2989	0.2043
5	0.6744	0.5631	0.6138	0.7326	0.5727	0.6429	0.7791	0.2597	0.3895

Πίνακας II.26: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Σουηδία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.0918	0.2308	0.1314	0.1939	0.3654	0.2533	0.4286	0.2373	0.3055
2	0.0471	0.3636	0.0833	0.5176	0.4037	0.4536	0.4000	0.4146	0.4072
3	0.6667	0.2687	0.3830	0.0741	0.1818	0.1053	0.1111	0.2727	0.1579
4	0.6048	0.4570	0.5206	0.8024	0.4073	0.5403	0.1737	0.3973	0.2417
5	0.3043	0.2500	0.2745	0.0145	0.2000	0.0270	0.2029	0.2121	0.2074
6	0.5424	0.3478	0.4238	0.2712	0.3721	0.3137	0.4068	0.2759	0.3288
7	0.5571	0.4432	0.4937	0.2429	0.4595	0.3178	0.5286	0.3217	0.4000
8	0.2676	0.4634	0.3393	0.3803	0.3600	0.3699	0.2394	0.3696	0.2906
9	0.3733	0.4118	0.3916	0.2267	0.4857	0.3091	0.2400	0.4091	0.3025
10	0.8143	0.7125	0.7600	0.8857	0.6526	0.7515	0.8571	0.6667	0.7500
Κόμμα	Mahalanobis			NN5			NN10		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4184	0.2715	0.3293	0.2857	0.3373	0.3094	0.3367	0.3511	0.3438
2	0.2706	0.4182	0.3286	0.6118	0.3077	0.4094	0.4588	0.3362	0.3881
3	0.2593	0.3684	0.3043	0	0	0	0.1852	0.2500	0.2128
4	0.5150	0.4279	0.4674	0.6587	0.4297	0.5201	0.4192	0.4762	0.4459
5	0.1304	0.3103	0.1837	0	0	0	0.1449	0.2326	0.1786
6	0.1186	0.5833	0.1972	0	0	0	0.2373	0.5600	0.3333
7	0.3429	0.4615	0.3934	0.4714	0.3667	0.4125	0.5286	0.3558	0.4253
8	0.6338	0.3782	0.4737	0.0845	0.2400	0.1250	0.3944	0.2979	0.3394
9	0.2267	0.4474	0.3009	0.3733	0.3415	0.3567	0.4	0.4615	0.4286
10	0.9143	0.5565	0.6919	0.8143	0.6628	0.7308	0.8143	0.6867	0.7451
Κόμμα	NN15			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2347	0.2875	0.2584	0.3163	0.3647	0.3388	0.3673	0.2500	0.2975
2	0.4235	0.3130	0.3600	0.4588	0.3939	0.4239	0.1647	0.2692	0.2044
3	0.2963	0.3478	0.3200	0.3333	0.4091	0.3673	0.2222	0.3158	0.2609
4	0.4371	0.4650	0.4506	0.4850	0.5062	0.4954	0.1317	0.5238	0.2105
5	0.2029	0.3182	0.2478	0.2174	0.2586	0.2362	0.1884	0.1429	0.1625
6	0.2542	0.3750	0.3030	0.2881	0.4857	0.3617	0.2881	0.2208	0.2500
7	0.4286	0.3659	0.3947	0.5143	0.4675	0.4898	0.6857	0.2623	0.3794
8	0.3803	0.2903	0.3293	0.4507	0.4444	0.4476	0.3521	0.4237	0.3846
9	0.4667	0.4605	0.4636	0.5333	0.4938	0.5128	0.1467	0.2895	0.1947
10	0.8286	0.7160	0.7682	0.9000	0.6176	0.7326	0.8143	0.6628	0.7308

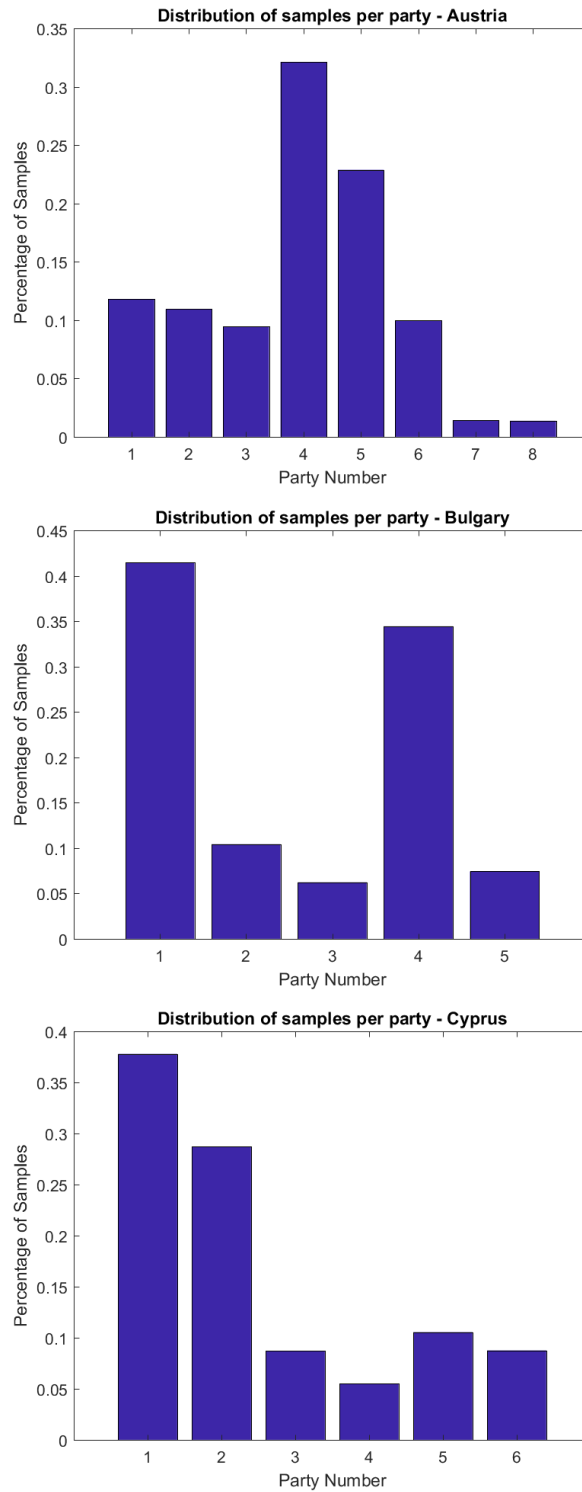
Πίνακας ΙΙ.27: Τα αποτελέσματα των ταξινομητών για κάθε κόμμα – Σλοβακία

Κόμμα	HMM			Naïve Bayes			KNN		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.2366	0.4231	0.3034	0.2258	0.3281	0.2675	0.5376	0.2941	0.3802
2	0.5833	0.1944	0.2917	0.2222	0.4706	0.3019	0.2500	0.4737	0.3273
3	0.6379	0.1045	0.1796	0.01724	0.2000	0.03175	0.2241	0.1275	0.1625
4	0.0054	0.3333	0.0106	0.03243	0.2609	0.05769	0.1784	0.2115	0.1935
5	0.1842	0.1931	0.1885	0.1118	0.2297	0.1504	0.1579	0.2353	0.1890
6	0.4587	0.6370	0.5333	0.7600	0.4831	0.5907	0.5973	0.5173	0.5545
7	0.0311	0.2381	0.0549	0.4845	0.3047	0.3741	0.2298	0.2846	0.2543
8	0.3051	0.2279	0.2609	0.339	0.2247	0.2703	0.2034	0.4444	0.2791
Κόμμα	Mahalanobis			NN3			NN5		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.1183	0.4583	0.1880	0.5914	0.3837	0.4490	0.3548	0.3837	0.3687
2	0.0556	0.3333	0.0952	0	0.3800	0	0.5278	0.38	0.4419
3	0.0345	0.2222	0.0597	0	0	0	0	0	0
4	0.0757	0.2979	0.1207	0	0.3077	0	0.0216	0.3077	0.0404
5	0.7763	0.1735	0.2837	0.0855	0.2832	0.1307	0.2105	0.2832	0.2415
6	0.416	0.6215	0.4984	0.8453	0.5165	0.6204	0.7920	0.5165	0.6253
7	0.1366	0.3333	0.1938	0.4348	0.3105	0.3599	0.4224	0.3105	0.3579
8	0.1525	0.2500	0.1895	0.2542	0.2151	0.2256	0.3390	0.2151	0.2632
Κόμμα	NN10			SVM			Party Coding		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
1	0.4624	0.5059	0.4831	0.4731	0.4074	0.4378	0.2581	0.3871	0.3097
2	0.3333	0.3429	0.338	0.3889	0.3256	0.3544	0.4444	0.3019	0.3596
3	0	0	0	0	0	0	0.2414	0.1197	0.1600
4	0.0973	0.2903	0.1457	0.0919	0.4250	0.1511	0.3189	0.2757	0.2957
5	0.1513	0.2674	0.1933	0.2039	0.2768	0.2348	0.0263	0.2222	0.0471
6	0.7387	0.4973	0.5944	0.7813	0.5327	0.6335	0.4933	0.6250	0.5514
7	0.3913	0.289	0.3325	0.3851	0.3196	0.3493	0.0683	0.2341	0.1058
8	0.3390	0.2500	0.2878	0.3559	0.2211	0.2727	0.4068	0.1739	0.2437

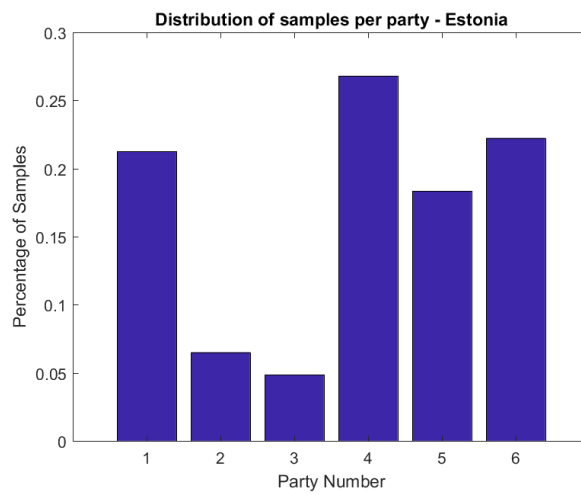
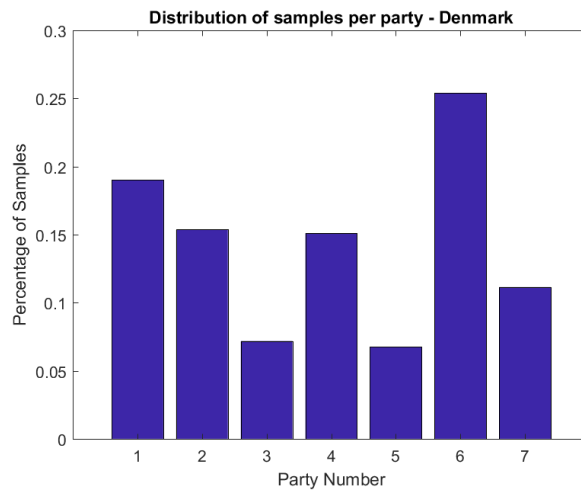
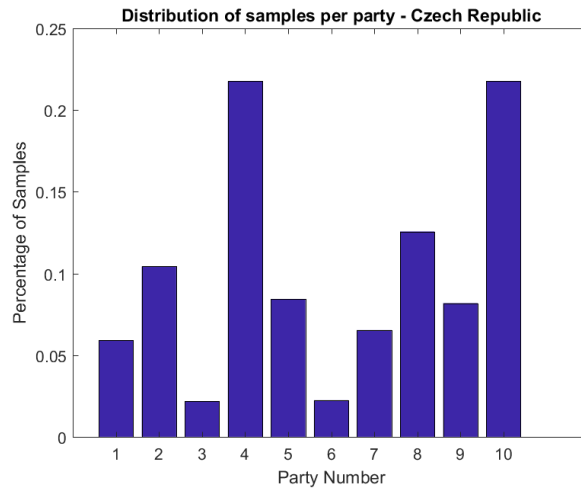
Παράρτημα ΙΙΙ

Η κατανομή του δείγματος

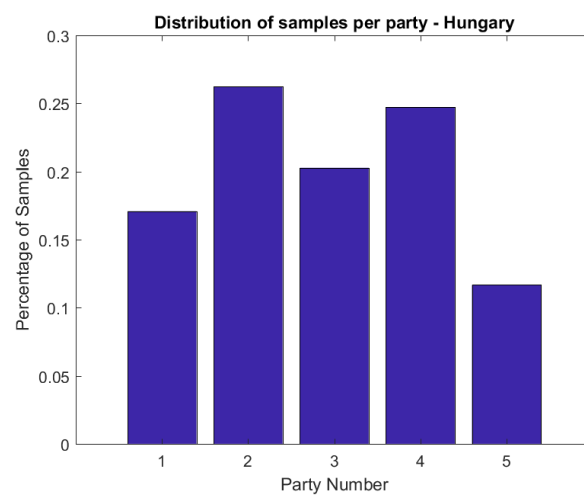
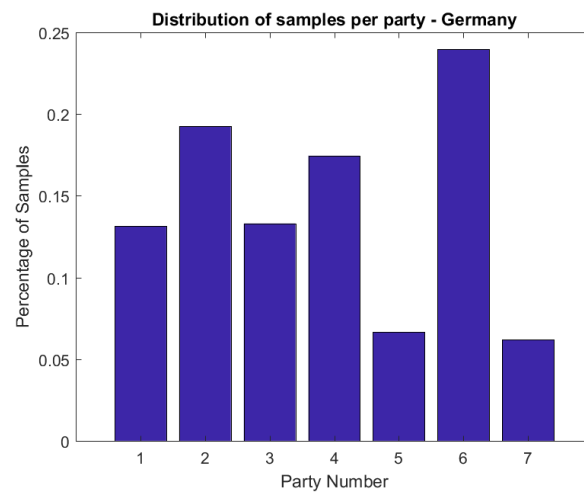
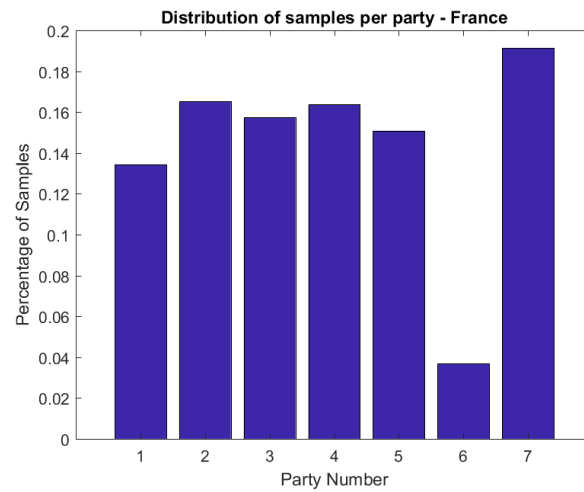
Το Παράρτημα περιλαμβάνει το ποσοστό των χρηστών που βρίσκεται σε κάθε κόμμα της κάθε χώρας.



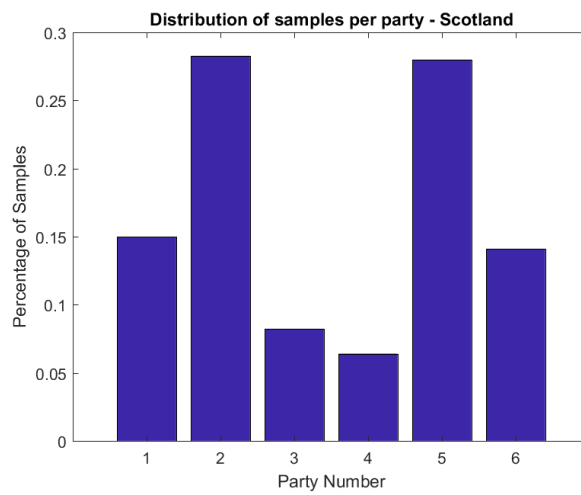
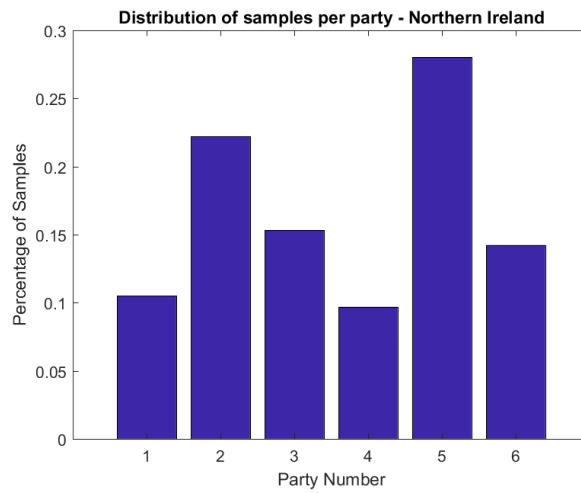
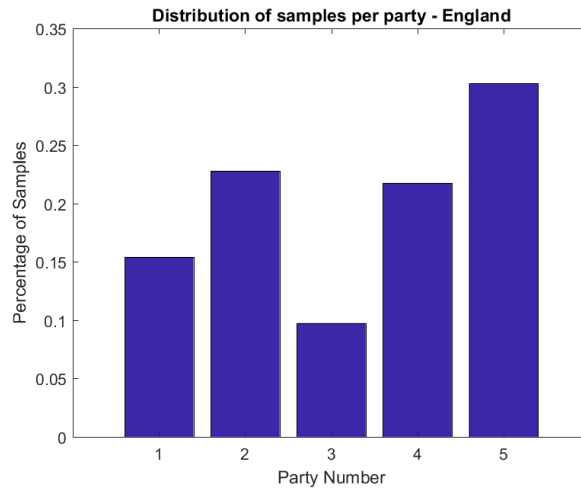
Σχήμα ΙΙΙ.1: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Αυστρία (β) Βουλγαρία (γ) Κύπρος



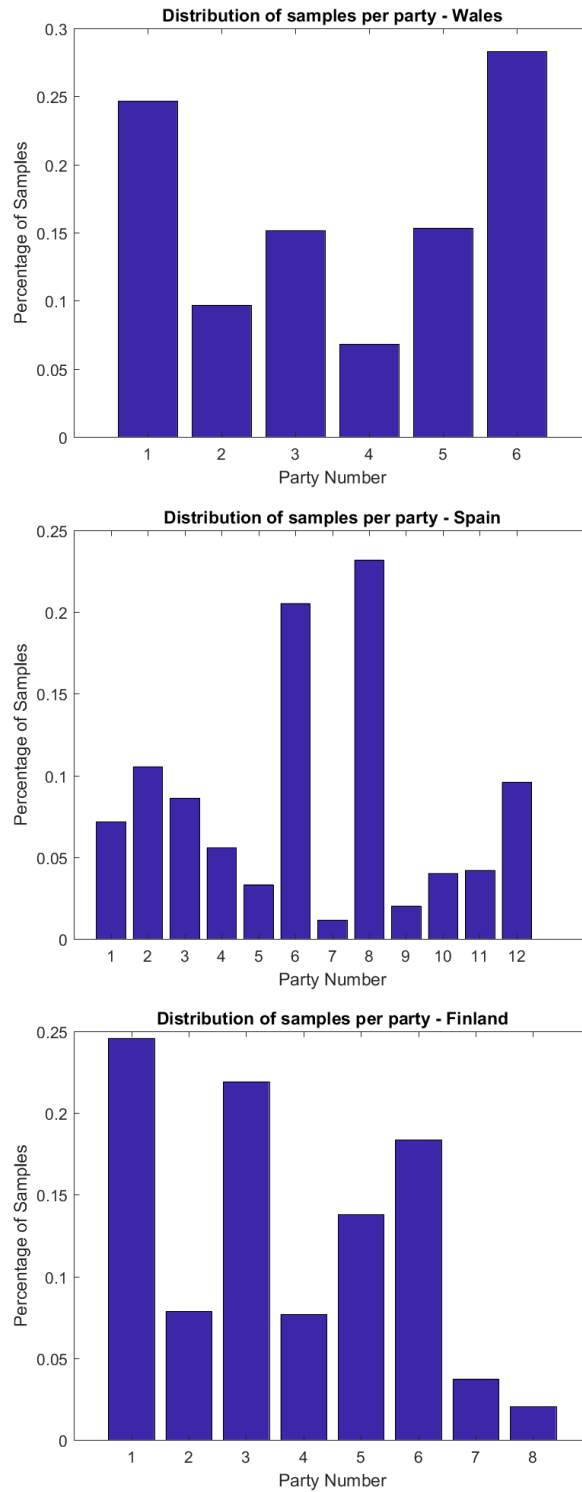
Σχήμα ΙΙΙ.2: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Τσεχία (β) Δανία (γ) Εσθονία



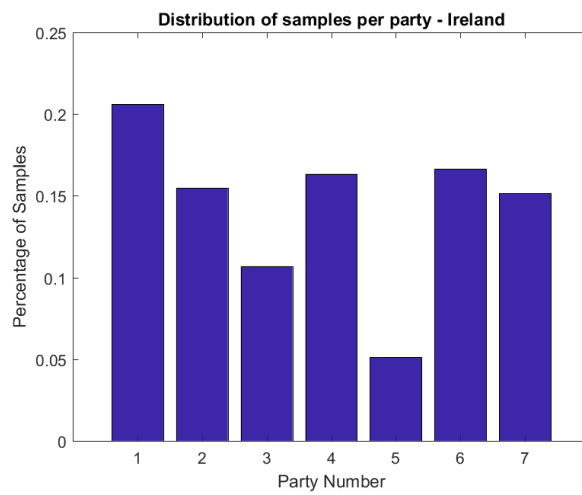
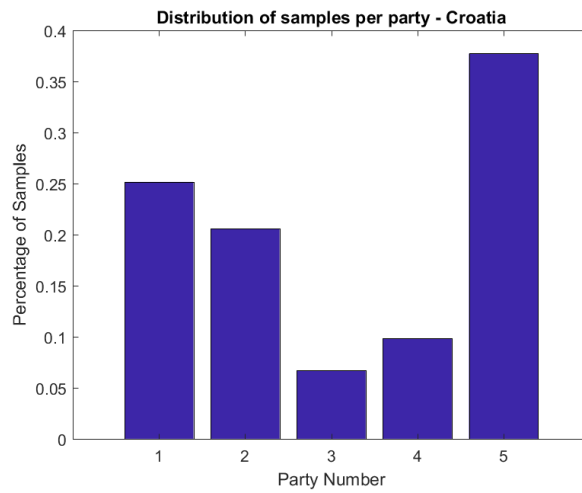
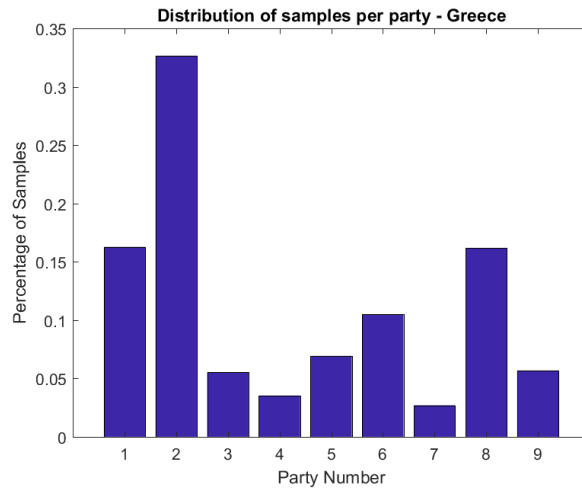
Σχήμα ΙΙΙ.3: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Γαλλία (β) Γερμανία (γ) Ουγγαρία



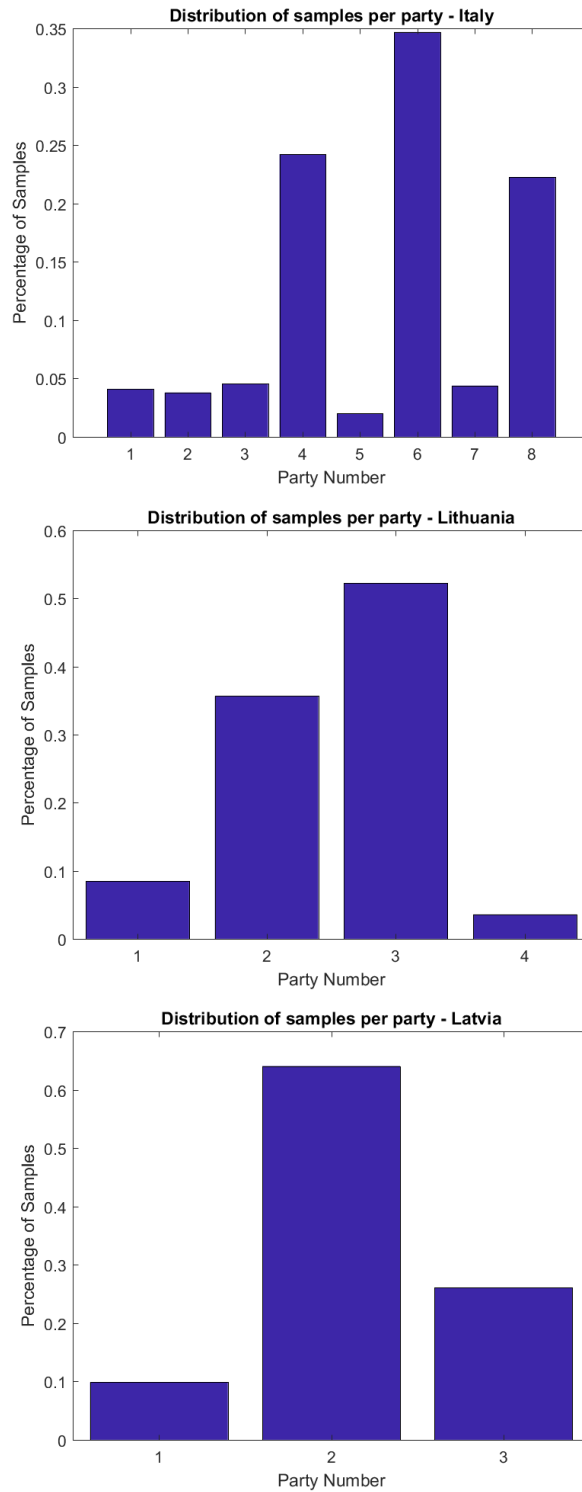
Σχήμα III.4: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Αγγλία (β) Β.Ιρλανδία (γ) Σκωτία



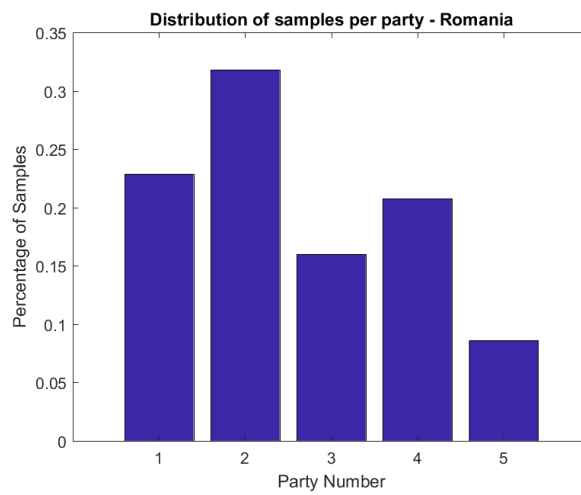
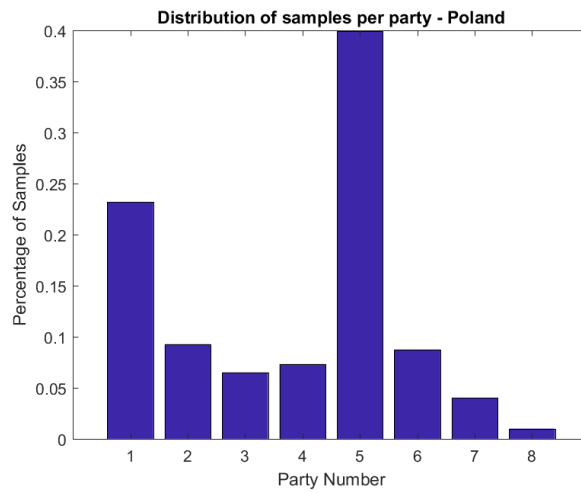
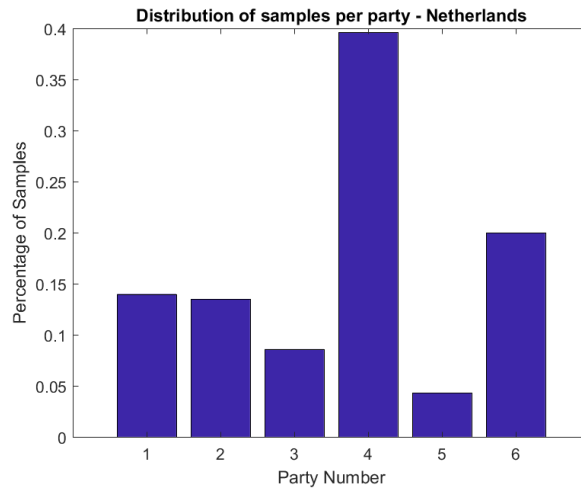
Σχήμα ΙΙΙ.5: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ουαλία (β) Ισπανία (γ) Φινλανδία



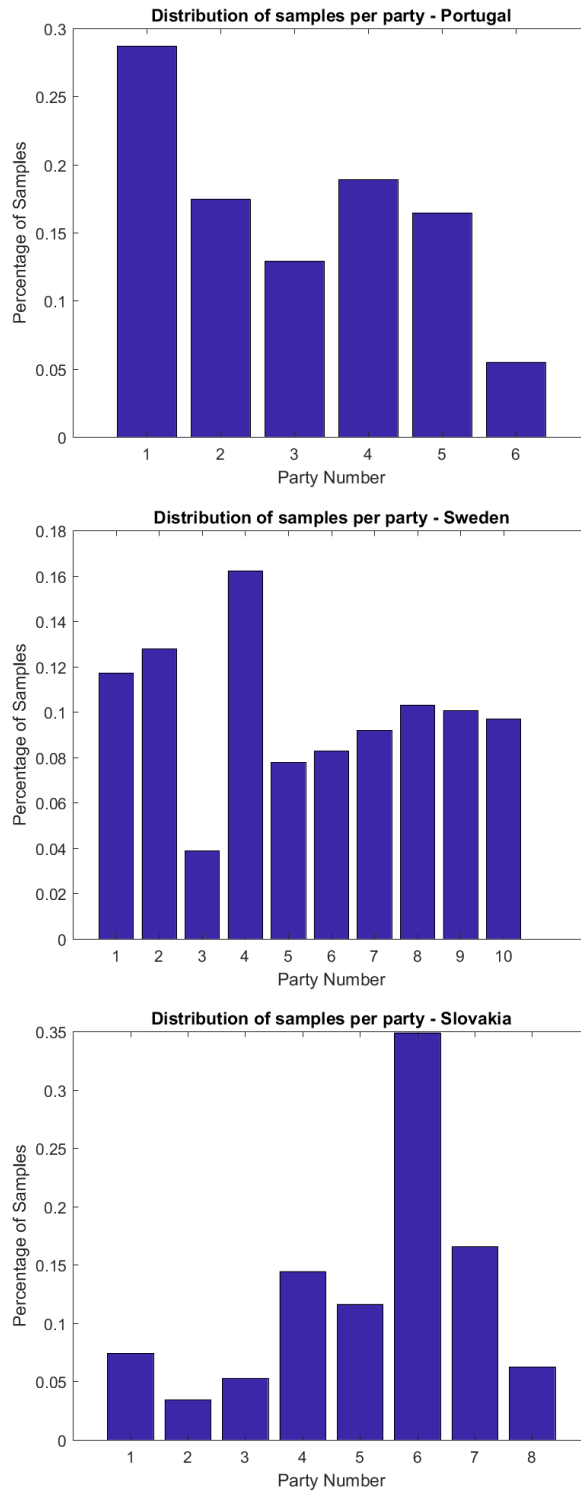
Σχήμα III.6: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ελλάδα (β) Κροατία (γ) Ιρλανδία



Σχήμα ΙΙΙ.7: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ιταλία (β) Λιθουανία (γ) Λετονία



Σχήμα III.8: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Ολλανδία (β) Πολωνία (γ) Ρουμανία



Σχήμα ΙΙΙ.9: Το ποσοστό των χρηστών ανά κόμμα για κάθε χώρα (α) Πορτογαλία (β) Σουηδία (γ) Σλοβακία

Παράρτημα IV

Αποτελέσματα των συνδυασμών μεταξύ των δηλώσεων πολιτικής και των συμπληρωματικών ερωτήσεων

Πίνακας IV.1: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Αυστρία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Αυστρία		0.5000	0.5024	0.5012	0.6899
	Q_3	0.5729	0.5513	0.5619	0.7417
	Q_3Q_8	0.5825	0.5625	0.5723	0.7477
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5773	0.5543	0.5656	0.7453
	Q_3Q_9	0.5825	0.5614	0.5718	0.7482
	Q_3Q_{10}	0.5773	0.5482	0.5623	0.7461
	Q_3Q_{11}	0.5825	0.5637	0.573	0.7478
	Q_8	0.5773	0.5568	0.5668	0.7458
	Q_8Q_9	0.5878	0.5715	0.5795	0.7514
	$Q_8Q_9Q_{10}$	0.5860	0.5676	0.5767	0.7505
	Q_8Q_{10}	0.5808	0.5564	0.5684	0.7507
	Q_9	0.5852	0.5652	0.5750	0.7515
	Q_9Q_{10}	0.5790	0.5572	0.5679	0.7457
	Q_{10}	0.5764	0.5527	0.5643	0.7444
	Q_{11}	0.5755	0.5421	0.5583	0.7444
	$Q_{11}Q_8$	0.5712	0.5449	0.5577	0.7448
	$Q_{11}Q_9$	0.5738	0.5422	0.5576	0.7426
	$Q_{11}Q_{10}$	0.5825	0.5614	0.5718	0.7482
	x	0.5432	0.5553	0.5491	0.7229
	xy	0.5471	0.5598	0.5534	0.7273
	xyz	0.5376	0.5346	0.5361	0.7235
	xz	0.5495	0.5631	0.5562	0.7235
	y	0.5455	0.5561	0.5508	0.7235
	yz	0.5313	0.5311	0.5312	0.7235
	z	0.5400	0.5435	0.5417	0.7235

Πίνακας IV.2: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Βουλγαρία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Βουλγαρία		0.4991	0.4784	0.4735	0.6917
	Q_3	0.5847	0.5633	0.5738	0.7513
	Q_3Q_8	0.5897	0.5684	0.5789	0.7548
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5897	0.5685	0.5789	0.7538
	Q_3Q_9	0.5897	0.5674	0.5784	0.7550
	Q_3Q_{10}	0.5897	0.5684	0.5789	0.7548
	Q_3Q_{11}	0.5759	0.5514	0.5634	0.7480
	Q_8	0.5897	0.5734	0.5814	0.7537
	Q_8Q_9	0.5897	0.5733	0.5814	0.7559
	$Q_8Q_9Q_{10}$	0.5897	0.5741	0.5818	0.7555
	Q_8Q_{10}	0.5897	0.5737	0.5816	0.7548
	Q_9	0.5922	0.5720	0.5819	0.7568
	Q_9Q_{10}	0.5809	0.5591	0.5698	0.7509
	Q_{10}	0.5872	0.5663	0.5766	0.7527
	Q_{11}	0.5885	0.5675	0.5778	0.7542
	$Q_{11}Q_8$	0.5809	0.5544	0.5674	0.7508
	$Q_{11}Q_9$	0.5759	0.5518	0.5636	0.7489
	$Q_{11}Q_{10}$	0.5772	0.5511	0.5638	0.7479
	x	0.5897	0.5684	0.5789	0.7548
	xy	0.5878	0.5746	0.5812	0.7560
	xyz	0.5829	0.5334	0.5571	0.7542
	xz	0.5993	0.538	0.5670	0.7624
	y	0.5780	0.5208	0.5479	0.7501
	yz	0.5796	0.5341	0.5559	0.7484
	z	0.5846	0.539	0.5609	0.7529

Πίνακας IV.3: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Κύπρος

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Κύπρος		0.6667	0.6153	0.6399	0.7766
	Q_3	0.6834	0.6155	0.6477	0.7944
	Q_3Q_8	0.6890	0.6248	0.6554	0.7993
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6909	0.6299	0.659	0.7980
	Q_3Q_9	0.6927	0.6361	0.6632	0.8014
	Q_3Q_{10}	0.6890	0.6385	0.6628	0.7988
	Q_3Q_{11}	0.6834	0.6300	0.6557	0.7951
	Q_8	0.6834	0.6342	0.6579	0.7909
	Q_8Q_9	0.6946	0.6552	0.6743	0.8004
	$Q_8Q_9Q_{10}$	0.6853	0.6395	0.6616	0.7948
	Q_8Q_{10}	0.6909	0.6429	0.666	0.7982
	Q_9	0.6816	0.6383	0.6592	0.7921
	Q_9Q_{10}	0.6834	0.6005	0.6393	0.7946
	Q_{10}	0.6816	0.6142	0.6461	0.7949
	Q_{11}	0.6872	0.6294	0.657	0.7986
	$Q_{11}Q_8$	0.6816	0.6299	0.6547	0.7947
	$Q_{11}Q_9$	0.6853	0.6274	0.6551	0.7967
	$Q_{11}Q_{10}$	0.6927	0.6361	0.6632	0.8014
	x	0.6749	0.6195	0.646	0.7933
	xy	0.6847	0.6740	0.6793	0.7916
	xyz	0.6847	0.6694	0.677	0.7932
	xz	0.6847	0.6308	0.6566	0.7963
	y	0.6921	0.6377	0.6638	0.7969
yz	0.6995	0.6688	0.6838	0.8014	
z	0.6921	0.6415	0.6659	0.8001	

Πίνακας IV.4: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Τσεχία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Τσεχία		0.5163	0.5042	0.5102	0.6881
	Q_3	0.5783	0.5663	0.5722	0.7348
	Q_3Q_8	0.5808	0.5680	0.5743	0.7361
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5830	0.5725	0.5777	0.7379
	Q_3Q_9	0.5812	0.5693	0.5752	0.7366
	Q_3Q_{10}	0.5860	0.5783	0.5821	0.7403
	Q_3Q_{11}	0.5899	0.5788	0.5843	0.7418
	Q_8	0.5871	0.5761	0.5816	0.7400
	Q_8Q_9	0.5885	0.5770	0.5827	0.7405
	$Q_8Q_9Q_{10}$	0.5894	0.5792	0.5843	0.7418
	Q_8Q_{10}	0.5981	0.5906	0.5943	0.7471
	Q_9	0.5922	0.5837	0.5879	0.7442
	Q_9Q_{10}	0.5785	0.5647	0.5715	0.7347
	Q_{10}	0.5803	0.5693	0.5748	0.7360
	Q_{11}	0.5819	0.5714	0.5766	0.7378
	$Q_{11}Q_8$	0.5853	0.5766	0.5809	0.7392
	$Q_{11}Q_9$	0.5805	0.5696	0.5750	0.7368
	$Q_{11}Q_{10}$	0.5812	0.5693	0.5752	0.7366
	x	0.5636	0.5497	0.5566	0.7224
	xy	0.5659	0.5502	0.5579	0.7261
	xyz	0.5696	0.5553	0.5624	0.7275
	xz	0.5591	0.5413	0.5501	0.7208
	y	0.5606	0.5465	0.5535	0.7211
	yz	0.5611	0.5460	0.5534	0.7211
	z	0.5566	0.5394	0.5479	0.7179

Πίνακας IV.5: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Γερμανία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Γερμανία		0.5999	0.5837	0.5917	0.7527
	Q_3	0.5987	0.5918	0.5952	0.7539
	Q_3Q_8	0.6038	0.5989	0.6013	0.7563
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6001	0.5911	0.5956	0.7547
	Q_3Q_9	0.6045	0.5991	0.6018	0.7568
	Q_3Q_{10}	0.5929	0.5837	0.5883	0.7494
	Q_3Q_{11}	0.6088	0.5997	0.6042	0.7592
	Q_8	0.6023	0.5936	0.5979	0.7563
	Q_8Q_9	0.6146	0.6083	0.6115	0.7622
	$Q_8Q_9Q_{10}$	0.6132	0.6061	0.6096	0.7618
	Q_8Q_{10}	0.6088	0.6022	0.6055	0.7601
	Q_9	0.6045	0.5951	0.5997	0.7569
	Q_9Q_{10}	0.6009	0.5939	0.5974	0.7534
	Q_{10}	0.6009	0.5991	0.6000	0.7557
	Q_{11}	0.5922	0.5801	0.5861	0.7504
	$Q_{11}Q_8$	0.5987	0.5891	0.5939	0.7553
	$Q_{11}Q_9$	0.5929	0.5811	0.5869	0.7492
	$Q_{11}Q_{10}$	0.6045	0.5991	0.6018	0.7568
	x	0.6158	0.5989	0.6072	0.7655
	xy	0.6158	0.5966	0.6060	0.7673
	xyz	0.6180	0.6000	0.6089	0.7686
	xz	0.6203	0.6017	0.6108	0.7686
	y	0.6113	0.5931	0.602	0.7647
	yz	0.6075	0.5870	0.5971	0.7621
	z	0.6045	0.5865	0.5954	0.7597

Πίνακας IV.6: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Δανία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Δανία		0.5586	0.5493	0.5540	0.7323
	Q_3	0.5856	0.5768	0.5812	0.7516
	Q_3Q_8	0.5827	0.5737	0.5782	0.7502
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5852	0.5761	0.5806	0.7518
	Q_3Q_9	0.5850	0.5761	0.5805	0.7518
	Q_3Q_{10}	0.5863	0.5775	0.5819	0.7527
	Q_3Q_{11}	0.5868	0.5778	0.5822	0.7527
	Q_8	0.5897	0.5805	0.5851	0.7544
	Q_8Q_9	0.5874	0.5782	0.5827	0.7533
	$Q_8Q_9Q_{10}$	0.5876	0.5787	0.5831	0.7535
	Q_8Q_{10}	0.5934	0.5843	0.5888	0.7576
	Q_9	0.5909	0.5822	0.5865	0.7552
	Q_9Q_{10}	0.5840	0.5755	0.5797	0.7510
	Q_{10}	0.5863	0.5774	0.5818	0.7525
	Q_{11}	0.5876	0.5791	0.5834	0.7535
	$Q_{11}Q_8$	0.5900	0.581	0.5855	0.7549
	$Q_{11}Q_9$	0.5869	0.5781	0.5824	0.7528
	$Q_{11}Q_{10}$	0.5850	0.5761	0.5805	0.7518
	x	0.5991	0.5939	0.5965	0.7641
	xy	0.6090	0.6051	0.6070	0.7701
	xyz	0.6173	0.6127	0.6150	0.7761
	xz	0.6077	0.6026	0.6051	0.7698
	y	0.5958	0.5919	0.5938	0.7606
	yz	0.6045	0.6003	0.6024	0.7671
	z	0.5950	0.5900	0.5925	0.7599

Πίνακας IV.7: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Εσθονία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Εσθονία		0.4411	0.4390	0.4401	0.6532
	Q_3	0.4566	0.4489	0.4527	0.6671
	Q_3Q_8	0.4599	0.4538	0.4568	0.6693
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.4628	0.4552	0.459	0.6711
	Q_3Q_9	0.4583	0.4518	0.455	0.6679
	Q_3Q_{10}	0.4603	0.4546	0.4575	0.6691
	Q_3Q_{11}	0.4707	0.4671	0.4689	0.6752
	Q_8	0.4702	0.4654	0.4678	0.6746
	Q_8Q_9	0.4744	0.4717	0.473	0.6774
	$Q_8Q_9Q_{10}$	0.4719	0.4694	0.4706	0.6764
	Q_8Q_{10}	0.4769	0.4721	0.4745	0.6795
	Q_9	0.4711	0.4672	0.4691	0.6752
	Q_9Q_{10}	0.4587	0.4526	0.4556	0.6666
	Q_{10}	0.4579	0.4499	0.4539	0.6679
	Q_{11}	0.4603	0.4544	0.4573	0.6698
	$Q_{11}Q_8$	0.4607	0.4532	0.4569	0.6700
	$Q_{11}Q_9$	0.4579	0.4514	0.4546	0.6679
	$Q_{11}Q_{10}$	0.4583	0.4518	0.455	0.6679
	x	0.4565	0.4566	0.4565	0.6666
	xy	0.4648	0.4669	0.4658	0.6730
	xyz	0.4589	0.4616	0.4602	0.6687
	xz	0.4604	0.461	0.4607	0.6673
	y	0.4526	0.4543	0.4534	0.6658
	yz	0.4545	0.4572	0.4559	0.6662
	z	0.4526	0.4555	0.454	0.6644

Πίνακας IV.8: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - ΗΒ-Αγγλία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
ΗΒ-Αγγλία		0.5748	0.5896	0.5821	0.7584
	Q_3	0.6683	0.6585	0.6633	0.8134
	Q_3Q_8	0.6675	0.6584	0.6629	0.8130
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6694	0.6599	0.6646	0.8141
	Q_3Q_9	0.6686	0.6597	0.6641	0.8136
	Q_3Q_{10}	0.6692	0.6598	0.6644	0.8141
	Q_3Q_{11}	0.6846	0.6765	0.6806	0.8225
	Q_8	0.6832	0.6743	0.6787	0.8220
	Q_8Q_9	0.6849	0.6769	0.6809	0.8227
	$Q_8Q_9Q_{10}$	0.6840	0.6756	0.6798	0.8222
	Q_8Q_{10}	0.6873	0.6785	0.6829	0.8242
	Q_9	0.6871	0.6783	0.6827	0.8239
	Q_9Q_{10}	0.6680	0.6587	0.6633	0.8130
	Q_{10}	0.6680	0.6586	0.6633	0.8132
	Q_{11}	0.6717	0.6623	0.667	0.8151
	$Q_{11}Q_8$	0.6710	0.6618	0.6664	0.8147
	$Q_{11}Q_9$	0.6708	0.6611	0.6659	0.8150
	$Q_{11}Q_{10}$	0.6686	0.6597	0.6641	0.8136
	x	0.6409	0.6286	0.6347	0.7988
	xy	0.6443	0.6358	0.64	0.8007
	xyz	0.6438	0.635	0.6394	0.7997
	xz	0.6431	0.6345	0.6388	0.7998
	y	0.6393	0.6258	0.6325	0.7971
	yz	0.6395	0.6252	0.6322	0.7972
	z	0.6368	0.6238	0.6302	0.7954

Πίνακας IV.9: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - ΗΒ-Βόρεια Ιρλανδία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
ΗΒ-Βόρεια Ιρλανδία		0.5099	0.5159	0.5129	0.7065
	Q ₃	0.5550	0.5297	0.5421	0.7341
	Q ₃ Q ₈	0.5478	0.5276	0.5375	0.7294
	Q ₃ Q ₈ Q ₉ Q ₁₀ Q ₁₁	0.5574	0.5391	0.5481	0.7361
	Q ₃ Q ₉	0.5478	0.5313	0.5394	0.7304
	Q ₃ Q ₁₀	0.5502	0.5327	0.5413	0.7335
	Q ₃ Q ₁₁	0.5502	0.5301	0.5400	0.7326
	Q ₈	0.5526	0.5317	0.5420	0.7339
	Q ₈ Q ₉	0.5550	0.5354	0.5450	0.7340
	Q ₈ Q ₉ Q ₁₀	0.5478	0.5277	0.5376	0.7318
	Q ₈ Q ₁₀	0.5526	0.5307	0.5415	0.7352
	Q ₉	0.5455	0.5266	0.5359	0.7318
	Q ₉ Q ₁₀	0.5526	0.5309	0.5415	0.7332
	Q ₁₀	0.5502	0.5289	0.5393	0.7329
	Q ₁₁	0.5455	0.5236	0.5343	0.7321
	Q ₁₁ Q ₈	0.5502	0.531	0.5404	0.7346
	Q ₁₁ Q ₉	0.5502	0.5311	0.5405	0.7347
	Q ₁₁ Q ₁₀	0.5478	0.5313	0.5394	0.7304
	x	0.5145	0.5333	0.5237	0.7055
	xy	0.5145	0.5298	0.5220	0.7049
	xyz	0.5058	0.5432	0.5238	0.7023
	xz	0.5058	0.5473	0.5257	0.7008
	y	0.5174	0.5267	0.5220	0.7075
	yz	0.5058	0.5285	0.5169	0.7029
	z	0.4913	0.5229	0.5066	0.6969

Πίνακας IV.10: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - ΗΒ-Σκωτία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
ΗΒ-Σκωτία		0.5686	0.5802	0.5743	0.7549
	Q_3	0.6294	0.6226	0.626	0.7894
	Q_3Q_8	0.6331	0.6249	0.629	0.7913
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6367	0.6294	0.633	0.7938
	Q_3Q_9	0.6367	0.6284	0.6325	0.7933
	Q_3Q_{10}	0.6352	0.6283	0.6317	0.7924
	Q_3Q_{11}	0.6389	0.6304	0.6346	0.7936
	Q_8	0.6461	0.6383	0.6422	0.7977
	Q_8Q_9	0.6432	0.6344	0.6388	0.7957
	$Q_8Q_9Q_{10}$	0.6425	0.6345	0.6385	0.7957
	Q_8Q_{10}	0.6468	0.6387	0.6427	0.7991
	Q_9	0.6396	0.6327	0.6362	0.7942
	Q_9Q_{10}	0.6302	0.6231	0.6266	0.7900
	Q_{10}	0.6410	0.6337	0.6373	0.7967
	Q_{11}	0.6360	0.6280	0.6320	0.793
	$Q_{11}Q_8$	0.6331	0.6262	0.6296	0.7921
	$Q_{11}Q_9$	0.6345	0.6266	0.6305	0.7918
	$Q_{11}Q_{10}$	0.6367	0.6284	0.6325	0.7933
	x	0.6000	0.5934	0.5967	0.7770
	xy	0.6070	0.5995	0.6032	0.7805
	xyz	0.6202	0.6130	0.6165	0.7865
	xz	0.6061	0.5990	0.6025	0.7786
	y	0.6114	0.6057	0.6085	0.7840
	yz	0.6193	0.6115	0.6154	0.7867
	z	0.6096	0.6046	0.6071	0.7813

Πίνακας IV.11: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - HB-Ουαλία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
HB-Ουαλία		0.4973	0.5043	0.5008	0.6977
	Q_3	0.6190	0.6250	0.6220	0.7740
	Q_3Q_8	0.6230	0.6297	0.6263	0.7752
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6190	0.6247	0.6218	0.7732
	Q_3Q_9	0.6230	0.6289	0.6259	0.7746
	Q_3Q_{10}	0.6217	0.6285	0.6250	0.7757
	Q_3Q_{11}	0.6270	0.6324	0.6297	0.7785
	Q_8	0.6257	0.6285	0.6271	0.7775
	Q_8Q_9	0.6297	0.6438	0.6366	0.7793
	$Q_8Q_9Q_{10}$	0.6310	0.6395	0.6352	0.7805
	Q_8Q_{10}	0.6337	0.6485	0.6410	0.7822
	Q_9	0.6337	0.6446	0.6391	0.7815
	Q_9Q_{10}	0.6257	0.6347	0.6301	0.7764
	Q_{10}	0.6230	0.6299	0.6264	0.7762
	Q_{11}	0.6324	0.6433	0.6378	0.7814
	$Q_{11}Q_8$	0.6283	0.6418	0.6350	0.7796
	$Q_{11}Q_9$	0.6257	0.6382	0.6319	0.7784
	$Q_{11}Q_{10}$	0.6230	0.6289	0.6259	0.7746
	x	0.6153	0.6050	0.6101	0.7680
	xy	0.6071	0.5922	0.5996	0.7650
	xyz	0.5990	0.5858	0.5924	0.7598
	xz	0.5925	0.5764	0.5843	0.7578
	y	0.6201	0.6101	0.6151	0.7743
	yz	0.6104	0.5941	0.6021	0.7686
	z	0.6234	0.6078	0.6155	0.7752

Πίνακας IV.12: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ισπανία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ισπανία		0.3718	0.3761	0.3739	0.5923
	Q_3	0.4261	0.4174	0.4217	0.6343
	Q_3Q_8	0.4273	0.4178	0.4225	0.6349
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.4271	0.4187	0.4229	0.6356
	Q_3Q_9	0.4272	0.4180	0.4225	0.6353
	Q_3Q_{10}	0.4315	0.4231	0.4273	0.6384
	Q_3Q_{11}	0.4408	0.4331	0.4369	0.6437
	Q_8	0.4424	0.4346	0.4385	0.6450
	Q_8Q_9	0.4413	0.4328	0.4370	0.6443
	$Q_8Q_9Q_{10}$	0.4413	0.4333	0.4373	0.6439
	Q_8Q_{10}	0.4476	0.4396	0.4436	0.6496
	Q_9	0.4444	0.4371	0.4408	0.6467
	Q_9Q_{10}	0.4260	0.4167	0.4213	0.6338
	Q_{10}	0.4253	0.4158	0.4205	0.6339
	Q_{11}	0.4306	0.4219	0.4262	0.6375
	$Q_{11}Q_8$	0.4293	0.4269	0.4281	0.6375
	$Q_{11}Q_9$	0.4304	0.422	0.4261	0.6369
	$Q_{11}Q_{10}$	0.4272	0.4180	0.4225	0.6353
	x	0.4225	0.4289	0.4257	0.6324
	xy	0.4208	0.4263	0.4235	0.6323
	xyz	0.4222	0.4269	0.4246	0.6332
	xz	0.4229	0.4269	0.4249	0.6325
	y	0.4172	0.4233	0.4202	0.6290
	yz	0.4168	0.4211	0.4189	0.6289
	z	0.4126	0.4104	0.4115	0.6252

Πίνακας IV.13: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ελλάδα

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ελλάδα		0.5581	0.5437	0.5508	0.7202
	Q_3	0.5940	0.5821	0.5880	0.7479
	Q_3Q_8	0.5947	0.5853	0.5900	0.7491
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5948	0.5847	0.5897	0.7490
	Q_3Q_9	0.5951	0.5843	0.5897	0.7495
	Q_3Q_{10}	0.5992	0.5889	0.5940	0.7527
	Q_3Q_{11}	0.5985	0.5887	0.5936	0.7516
	Q_8	0.6001	0.5881	0.5940	0.7522
	Q_8Q_9	0.6021	0.5950	0.5985	0.7539
	$Q_8Q_9Q_{10}$	0.6002	0.5893	0.5947	0.7522
	Q_8Q_{10}	0.6059	0.5980	0.6019	0.7569
	Q_9	0.6036	0.5905	0.5970	0.7547
	Q_9Q_{10}	0.5943	0.5824	0.5883	0.7484
	Q_{10}	0.5927	0.5818	0.5872	0.7472
	Q_{11}	0.5967	0.5843	0.5905	0.7505
	$Q_{11}Q_8$	0.5972	0.5855	0.5913	0.7506
	$Q_{11}Q_9$	0.5954	0.5802	0.5877	0.7494
	$Q_{11}Q_{10}$	0.5951	0.5843	0.5897	0.7495
	x	0.5840	0.5660	0.5748	0.7432
	xy	0.5931	0.5778	0.5854	0.7479
	xyz	0.5953	0.5793	0.5872	0.7489
	xz	0.5905	0.5731	0.5816	0.7463
	y	0.5849	0.5654	0.5750	0.7417
	yz	0.5884	0.5700	0.5791	0.7437
	z	0.5830	0.5631	0.5729	0.7405

Πίνακας IV.14: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Κροατία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Κροατία		0.5920	0.5522	0.5714	0.7609
	Q_3	0.6209	0.6261	0.6235	0.7845
	Q_3Q_8	0.6209	0.6222	0.6215	0.7838
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6237	0.6246	0.6242	0.7849
	Q_3Q_9	0.6209	0.6180	0.6194	0.7841
	Q_3Q_{10}	0.6223	0.6236	0.6229	0.7834
	Q_3Q_{11}	0.6423	0.6435	0.6429	0.7946
	Q_8	0.6495	0.6433	0.6464	0.8005
	Q_8Q_9	0.6538	0.6439	0.6488	0.8014
	$Q_8Q_9Q_{10}$	0.6438	0.6412	0.6425	0.7968
	Q_8Q_{10}	0.6524	0.6297	0.6408	0.8010
	Q_9	0.6552	0.6463	0.6507	0.8001
	Q_9Q_{10}	0.6180	0.6138	0.6159	0.7828
	Q_{10}	0.6209	0.6233	0.6221	0.7845
	Q_{11}	0.6195	0.6221	0.6208	0.7826
	$Q_{11}Q_8$	0.6237	0.6253	0.6245	0.7852
	$Q_{11}Q_9$	0.6166	0.6291	0.6228	0.7809
	$Q_{11}Q_{10}$	0.6209	0.6180	0.6194	0.7841
	x	0.5846	0.5826	0.5836	0.7619
	xy	0.5827	0.6005	0.5915	0.7626
	xyz	0.5901	0.5839	0.5870	0.7658
	xz	0.5882	0.5804	0.5843	0.7644
	y	0.5754	0.5252	0.5492	0.7544
	yz	0.5790	0.5587	0.5687	0.7580
	z	0.5827	0.5757	0.5792	0.7578

Πίνακας IV.15: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ουγγαρία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ουγγαρία		0.6238	0.5892	0.6060	0.7783
	Q_3	0.5951	0.5549	0.5743	0.7683
	Q_3Q_8	0.6039	0.5836	0.5936	0.7727
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6068	0.5694	0.5875	0.7732
	Q_3Q_9	0.6049	0.5810	0.5927	0.7733
	Q_3Q_{10}	0.6205	0.6214	0.6209	0.7844
	Q_3Q_{11}	0.6039	0.5949	0.5993	0.7726
	Q_8	0.6010	0.5773	0.5889	0.7708
	Q_8Q_9	0.6107	0.5878	0.5991	0.7753
	$Q_8Q_9Q_{10}$	0.6078	0.5862	0.5968	0.7748
	Q_8Q_{10}	0.6283	0.6225	0.6254	0.7891
	Q_9	0.6234	0.6200	0.6217	0.7862
	Q_9Q_{10}	0.5990	0.5861	0.5925	0.7704
	Q_{10}	0.6010	0.5630	0.5814	0.7710
	Q_{11}	0.6146	0.6215	0.6181	0.7812
	$Q_{11}Q_8$	0.6176	0.6240	0.6208	0.7825
	$Q_{11}Q_9$	0.6185	0.6255	0.6220	0.7840
	$Q_{11}Q_{10}$	0.6049	0.5810	0.5927	0.7733
	x	0.6225	0.6269	0.6247	0.7809
	xy	0.6308	0.6463	0.6385	0.7858
	xyz	0.6332	0.6423	0.6377	0.7854
	xz	0.6296	0.6354	0.6325	0.7821
	y	0.6332	0.6094	0.6211	0.7860
	yz	0.6284	0.5958	0.6117	0.7840
	z	0.6332	0.6048	0.6187	0.7848

Πίνακας IV.16: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ιρλανδία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ιρλανδία		0.3919	0.3749	0.3832	0.6034
	Q_3	0.4816	0.4539	0.4674	0.6647
	Q_3Q_8	0.4825	0.4575	0.4697	0.6657
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.4825	0.4579	0.4699	0.6652
	Q_3Q_9	0.4851	0.4570	0.4706	0.6669
	Q_3Q_{10}	0.4869	0.4572	0.4716	0.6686
	Q_3Q_{11}	0.4948	0.4703	0.4822	0.6746
	Q_8	0.4939	0.4692	0.4812	0.6747
	Q_8Q_9	0.4965	0.4752	0.4856	0.6764
	$Q_8Q_9Q_{10}$	0.5009	0.4776	0.4890	0.6783
	Q_8Q_{10}	0.4948	0.4716	0.4829	0.6755
	Q_9	0.4983	0.4751	0.4864	0.6782
	Q_9Q_{10}	0.4843	0.4482	0.4655	0.6654
	Q_{10}	0.4808	0.4528	0.4664	0.6648
	Q_{11}	0.4869	0.4589	0.4725	0.6689
	$Q_{11}Q_8$	0.4913	0.4703	0.4806	0.6711
	$Q_{11}Q_9$	0.4878	0.4558	0.4713	0.6694
	$Q_{11}Q_{10}$	0.4851	0.4570	0.4706	0.6669
	x	0.4204	0.3926	0.4060	0.6250
	xy	0.4087	0.3802	0.3939	0.6192
	xyz	0.4178	0.3893	0.4030	0.6249
	xz	0.4254	0.3972	0.4108	0.6300
	y	0.3945	0.3611	0.3770	0.6070
	yz	0.4012	0.3698	0.3848	0.6148
	z	0.4187	0.3879	0.4027	0.6241

Πίνακας IV.17: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ιταλία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ιταλία		0.6022	0.5974	0.5998	0.7631
	Q_3	0.6538	0.6435	0.6486	0.7974
	Q_3Q_8	0.6602	0.6494	0.6548	0.8003
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6591	0.6483	0.6536	0.8002
	Q_3Q_9	0.6598	0.6485	0.6541	0.8006
	Q_3Q_{10}	0.6619	0.6506	0.6562	0.8017
	Q_3Q_{11}	0.6664	0.6569	0.6616	0.8044
	Q_8	0.6676	0.6571	0.6623	0.8056
	Q_8Q_9	0.6710	0.6616	0.6663	0.8075
	$Q_8Q_9Q_{10}$	0.6703	0.6608	0.6655	0.8067
	Q_8Q_{10}	0.6745	0.6650	0.6697	0.8094
	Q_9	0.6703	0.6609	0.6656	0.8072
	Q_9Q_{10}	0.6600	0.6489	0.6544	0.8005
	Q_{10}	0.6577	0.6463	0.6520	0.7997
	Q_{11}	0.6616	0.6504	0.6560	0.8016
	$Q_{11}Q_8$	0.6605	0.6489	0.6547	0.8015
	$Q_{11}Q_9$	0.6593	0.6478	0.6535	0.8005
	$Q_{11}Q_{10}$	0.6598	0.6485	0.6541	0.8006
	x	0.6365	0.6293	0.6329	0.7843
	xy	0.6363	0.6291	0.6326	0.7855
	xyz	0.6321	0.6216	0.6268	0.7831
	xz	0.6346	0.6252	0.6299	0.7832
	y	0.6291	0.6195	0.6242	0.7805
	yz	0.6315	0.6212	0.6263	0.7815
	z	0.6249	0.6150	0.6199	0.7777

Πίνακας IV.18: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Λιθουανία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Λιθουανία		0.6142	0.5953	0.6046	0.7843
	Q_3	0.6560	0.6445	0.6502	0.8103
	Q_3Q_8	0.6486	0.6307	0.6396	0.8069
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6552	0.6385	0.6467	0.8097
	Q_3Q_9	0.6511	0.6344	0.6427	0.8077
	Q_3Q_{10}	0.6593	0.6580	0.6587	0.8134
	Q_3Q_{11}	0.6667	0.6599	0.6633	0.8166
	Q_8	0.6724	0.6594	0.6658	0.8198
	Q_8Q_9	0.6732	0.6600	0.6666	0.8204
	$Q_8Q_9Q_{10}$	0.6675	0.6563	0.6618	0.8170
	Q_8Q_{10}	0.6847	0.6836	0.6841	0.8267
	Q_9	0.6781	0.6816	0.6799	0.8234
	Q_9Q_{10}	0.6478	0.6358	0.6418	0.8067
	Q_{10}	0.6503	0.6359	0.6430	0.8083
	Q_{11}	0.6511	0.6547	0.6529	0.8099
	$Q_{11}Q_8$	0.6585	0.6509	0.6547	0.8134
	$Q_{11}Q_9$	0.6486	0.6502	0.6494	0.8088
	$Q_{11}Q_{10}$	0.6511	0.6344	0.6427	0.8077
	x	0.6589	0.6454	0.6521	0.8146
	xy	0.7031	0.6850	0.6939	0.8377
	xyz	0.7055	0.6873	0.6963	0.8385
	xz	0.6531	0.6383	0.6456	0.8114
	y	0.7078	0.6901	0.6988	0.8393
	yz	0.6997	0.6810	0.6902	0.8350
	z	0.6589	0.6405	0.6496	0.8140

Πίνακας IV.19: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Λετονία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Λετονία		0.7367	0.7084	0.7223	0.8643
	Q_3	0.7446	0.7273	0.7358	0.8709
	Q_3Q_8	0.7473	0.7359	0.7415	0.8723
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.7500	0.7360	0.7429	0.8736
	Q_3Q_9	0.7446	0.7286	0.7365	0.8709
	Q_3Q_{10}	0.7500	0.7407	0.7453	0.8736
	Q_3Q_{11}	0.7690	0.7655	0.7673	0.8827
	Q_8	0.7717	0.7653	0.7685	0.8841
	Q_8Q_9	0.7745	0.7705	0.7725	0.8854
	$Q_8Q_9Q_{10}$	0.7799	0.7740	0.7770	0.8877
	Q_8Q_{10}	0.7745	0.7659	0.7702	0.8850
	Q_9	0.7772	0.7741	0.7756	0.8868
	Q_9Q_{10}	0.7446	0.7295	0.7370	0.8709
	Q_{10}	0.7446	0.7271	0.7357	0.8709
	Q_{11}	0.7473	0.7341	0.7407	0.8718
	$Q_{11}Q_8$	0.7582	0.7468	0.7524	0.8777
	$Q_{11}Q_9$	0.7473	0.7372	0.7422	0.8723
	$Q_{11}Q_{10}$	0.7446	0.7286	0.7365	0.8709
	x	0.7415	0.7362	0.7388	0.8668
	xy	0.7245	0.7156	0.7200	0.8583
	xyz	0.7279	0.7215	0.7247	0.8594
	xz	0.7551	0.7514	0.7532	0.8741
	y	0.7551	0.7514	0.7532	0.8741
	yz	0.7449	0.7443	0.7446	0.8685
	z	0.7483	0.7451	0.7467	0.8702

Πίνακας IV.20: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ολλανδία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ολλανδία		0.6810	0.6855	0.6832	0.8233
	Q_3	0.6900	0.6997	0.6948	0.8331
	Q_3Q_8	0.6812	0.6840	0.6826	0.8295
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6754	0.6772	0.6763	0.8258
	Q_3Q_9	0.6798	0.6818	0.6808	0.8287
	Q_3Q_{10}	0.6754	0.6777	0.6766	0.8247
	Q_3Q_{11}	0.6856	0.688	0.6868	0.8302
	Q_8	0.6900	0.6984	0.6941	0.8331
	Q_8Q_9	0.6769	0.6784	0.6776	0.8280
	$Q_8Q_9Q_{10}$	0.6812	0.6878	0.6845	0.8280
	Q_8Q_{10}	0.6812	0.6849	0.6830	0.8264
	Q_9	0.6769	0.6822	0.6795	0.8235
	Q_9Q_{10}	0.6900	0.6970	0.6935	0.8324
	Q_{10}	0.6856	0.6897	0.6877	0.8309
	Q_{11}	0.6725	0.6752	0.6739	0.8226
	$Q_{11}Q_8$	0.6856	0.6876	0.6866	0.8287
	$Q_{11}Q_9$	0.6769	0.6838	0.6803	0.8251
	$Q_{11}Q_{10}$	0.6812	0.684	0.6826	0.8295
	x	0.6647	0.6722	0.6684	0.8227
	xy	0.6705	0.6835	0.6769	0.8212
	xyz	0.6879	0.6940	0.6909	0.8260
	xz	0.6705	0.6801	0.6753	0.8196
	y	0.6474	0.6516	0.6495	0.8072
yz	0.6590	0.6668	0.6629	0.8135	
z	0.6512	0.6687	0.6598	0.8064	

Πίνακας IV.21: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Πολωνία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Πολωνία		0.6442	0.6088	0.6260	0.7769
	Q_3	0.6887	0.6741	0.6813	0.8094
	Q_3Q_8	0.6902	0.6716	0.6807	0.8102
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.6906	0.6715	0.6809	0.8108
	Q_3Q_9	0.6913	0.6724	0.6817	0.8110
	Q_3Q_{10}	0.6936	0.6780	0.6857	0.8128
	Q_3Q_{11}	0.6981	0.6796	0.6887	0.8147
	Q_8	0.6962	0.6785	0.6872	0.8141
	Q_8Q_9	0.6991	0.6805	0.6897	0.8153
	$Q_8Q_9Q_{10}$	0.6986	0.6818	0.6901	0.8150
	Q_8Q_{10}	0.7019	0.6854	0.6935	0.8180
	Q_9	0.7011	0.6846	0.6928	0.8169
	Q_9Q_{10}	0.6904	0.6761	0.6832	0.8102
	Q_{10}	0.6907	0.6745	0.6825	0.8106
	Q_{11}	0.6928	0.6805	0.6866	0.8124
	$Q_{11}Q_8$	0.6935	0.6792	0.6863	0.8132
	$Q_{11}Q_9$	0.6918	0.6764	0.684	0.8118
	$Q_{11}Q_{10}$	0.6913	0.6724	0.6817	0.8110
	x	0.6863	0.6634	0.6747	0.8091
	xy	0.6888	0.6672	0.6778	0.8102
	xyz	0.6925	0.6751	0.6837	0.8125
	xz	0.6872	0.6652	0.676	0.8088
	y	0.6887	0.6655	0.6769	0.8093
	yz	0.6887	0.6673	0.6778	0.8088
	z	0.6887	0.6741	0.6813	0.8094

Πίνακας IV.22: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Πορτογαλία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Πορτογαλία		0.4609	0.4075	0.4325	0.6407
	Q_3	0.5507	0.5147	0.5321	0.7085
	Q_3Q_8	0.5497	0.5083	0.5282	0.7092
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5521	0.5136	0.5321	0.7100
	Q_3Q_9	0.5490	0.5069	0.5271	0.7088
	Q_3Q_{10}	0.5535	0.5112	0.5315	0.7113
	Q_3Q_{11}	0.5686	0.5429	0.5554	0.7260
	Q_8	0.5688	0.5504	0.5595	0.7263
	Q_8Q_9	0.5669	0.5405	0.5534	0.7256
	$Q_8Q_9Q_{10}$	0.5683	0.5478	0.5578	0.7260
	Q_8Q_{10}	0.5740	0.5434	0.5583	0.7305
	Q_9	0.5709	0.5540	0.5623	0.7274
	Q_9Q_{10}	0.5493	0.5107	0.5293	0.7084
	Q_{10}	0.5512	0.5110	0.5303	0.7091
	Q_{11}	0.5523	0.5145	0.5328	0.7101
	$Q_{11}Q_8$	0.5542	0.5152	0.5340	0.7115
	$Q_{11}Q_9$	0.5521	0.5129	0.5318	0.7097
	$Q_{11}Q_{10}$	0.5490	0.5069	0.5271	0.7088
	x	0.5424	0.4912	0.5155	0.7057
	xy	0.5447	0.5027	0.5228	0.7071
	xyz	0.5450	0.5026	0.5229	0.7069
	xz	0.5434	0.4923	0.5166	0.7068
	y	0.5275	0.4781	0.5016	0.6950
	yz	0.5334	0.4976	0.5149	0.6982
	z	0.5290	0.4769	0.5016	0.6947

Πίνακας IV.23: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Ρουμανία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Ρουμανία		0.5185	0.4754	0.4960	0.7103
	Q_3	0.5816	0.5432	0.5617	0.7535
	Q_3Q_8	0.5842	0.5430	0.5628	0.7542
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.5825	0.5499	0.5657	0.7539
	Q_3Q_9	0.5825	0.5421	0.5616	0.7540
	Q_3Q_{10}	0.5791	0.5353	0.5563	0.7525
	Q_3Q_{11}	0.5901	0.5577	0.5735	0.7578
	Q_8	0.5859	0.5565	0.5708	0.7563
	Q_8Q_9	0.5884	0.5562	0.5718	0.7568
	$Q_8Q_9Q_{10}$	0.5910	0.5607	0.5754	0.7589
	Q_8Q_{10}	0.5969	0.5710	0.5837	0.7616
	Q_9	0.5901	0.5602	0.5748	0.7588
	Q_9Q_{10}	0.5833	0.5418	0.5618	0.7550
	Q_{10}	0.5876	0.5460	0.5660	0.7571
	Q_{11}	0.5799	0.5360	0.5571	0.7529
	$Q_{11}Q_8$	0.5867	0.5517	0.5687	0.7574
	$Q_{11}Q_9$	0.5791	0.5381	0.5578	0.7532
	$Q_{11}Q_{10}$	0.5825	0.5421	0.5616	0.7540
	x	0.6034	0.5787	0.5908	0.7607
	xy	0.5917	0.5521	0.5712	0.7539
	xyz	0.5917	0.5629	0.5770	0.7547
	xz	0.5970	0.5688	0.5826	0.7591
	y	0.5663	0.5299	0.5475	0.7418
	yz	0.5546	0.5232	0.5385	0.7340
	z	0.5652	0.5394	0.5520	0.7419

Πίνακας IV.24: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Σουηδία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Σουηδία		0.4589	0.4499	0.4544	0.6511
	Q_3	0.4854	0.4903	0.4879	0.6758
	Q_3Q_8	0.4854	0.4950	0.4902	0.6759
	$Q_3Q_8Q_9Q_{10}Q_{11}$	0.4871	0.4938	0.4904	0.6754
	Q_3Q_9	0.4889	0.4885	0.4887	0.6761
	Q_3Q_{10}	0.4820	0.4804	0.4812	0.6727
	Q_3Q_{11}	0.4923	0.5042	0.4981	0.6791
	Q_8	0.4871	0.4988	0.4929	0.6767
	Q_8Q_9	0.4871	0.5011	0.4940	0.6759
	$Q_8Q_9Q_{10}$	0.4871	0.4908	0.4889	0.6772
	Q_8Q_{10}	0.4768	0.4807	0.4788	0.6709
	Q_9	0.4906	0.4989	0.4947	0.6766
	Q_9Q_{10}	0.4889	0.4890	0.4889	0.6769
	Q_{10}	0.4889	0.4953	0.4920	0.6778
	Q_{11}	0.4820	0.4792	0.4806	0.6731
	$Q_{11}Q_8$	0.4803	0.4786	0.4794	0.6733
	$Q_{11}Q_9$	0.4889	0.4879	0.4884	0.6759
	$Q_{11}Q_{10}$	0.4889	0.4885	0.4887	0.6761
	x	0.5182	0.5241	0.5212	0.6959
	xy	0.5240	0.5373	0.5306	0.7018
	xyz	0.5336	0.5550	0.5441	0.7079
	xz	0.5259	0.5442	0.5349	0.7024
	y	0.5010	0.5172	0.5089	0.6852
	yz	0.5144	0.5543	0.5336	0.6932
	z	0.4875	0.5263	0.5062	0.6767

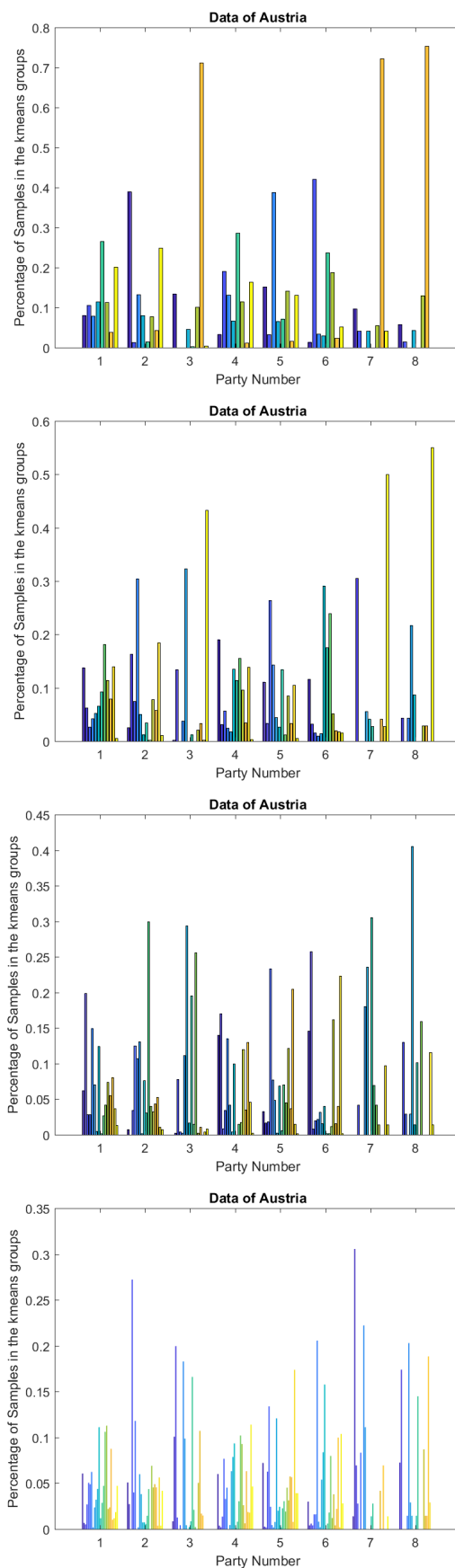
Πίνακας IV.25: Η απόδοση της πρόβλεψης ψήφου ενισχύοντας κάθε φορά τις δηλώσεις πολιτικής με μια ή περισσότερες συμπληρωματικές ερωτήσεις - Σλοβακία

Σύνολο Δεδομένων	δηλώσεις πολιτικής +	Recall	Precision	F1	MAP
Σλοβακία		0.4204	0.3803	0.3993	0.6113
	Q ₃	0.4535	0.3909	0.4198	0.6423
	Q ₃ Q ₈	0.4582	0.3958	0.4247	0.6449
	Q ₃ Q ₈ Q ₉ Q ₁₀ Q ₁₁	0.4535	0.3929	0.4210	0.6434
	Q ₃ Q ₉	0.4594	0.3962	0.4255	0.6446
	Q ₃ Q ₁₀	0.4642	0.4124	0.4368	0.6500
	Q ₃ Q ₁₁	0.4630	0.4091	0.4344	0.6493
	Q ₈	0.4594	0.4037	0.4298	0.6480
	Q ₈ Q ₉	0.4642	0.4106	0.4358	0.6491
	Q ₈ Q ₉ Q ₁₀	0.4630	0.4115	0.4357	0.6495
	Q ₈ Q ₁₀	0.4547	0.4067	0.4294	0.6462
	Q ₉	0.4523	0.3983	0.4236	0.6455
	Q ₉ Q ₁₀	0.4570	0.3928	0.4225	0.6441
	Q ₁₀	0.4487	0.3864	0.4152	0.6399
	Q ₁₁	0.4618	0.4086	0.4336	0.6481
	Q ₁₁ Q ₈	0.4511	0.3985	0.4231	0.6429
	Q ₁₁ Q ₉	0.4618	0.4124	0.4357	0.6491
	Q ₁₁ Q ₁₀	0.4594	0.3962	0.4255	0.6446
	x	0.4772	0.4401	0.4579	0.6612
	xy	0.4837	0.4582	0.4706	0.6641
	xyz	0.5336	0.5550	0.5441	0.7079
	xz	0.5259	0.5442	0.5349	0.7024
	y	0.4746	0.4387	0.4560	0.6554
	yz	0.4707	0.4349	0.4521	0.6540
	z	0.4759	0.4377	0.4560	0.6550

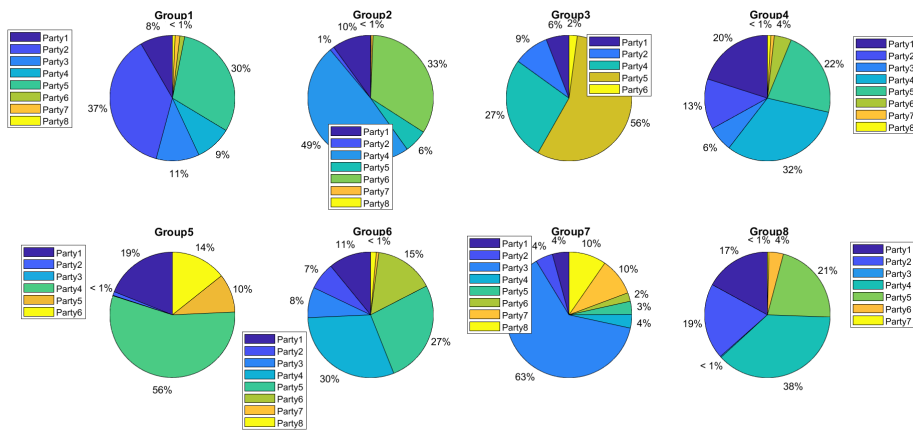
Παράρτημα V

Αποτελέσματα από τη διερεύνηση της ιδεολογικής ταύτισης κόμματος-χρήστη

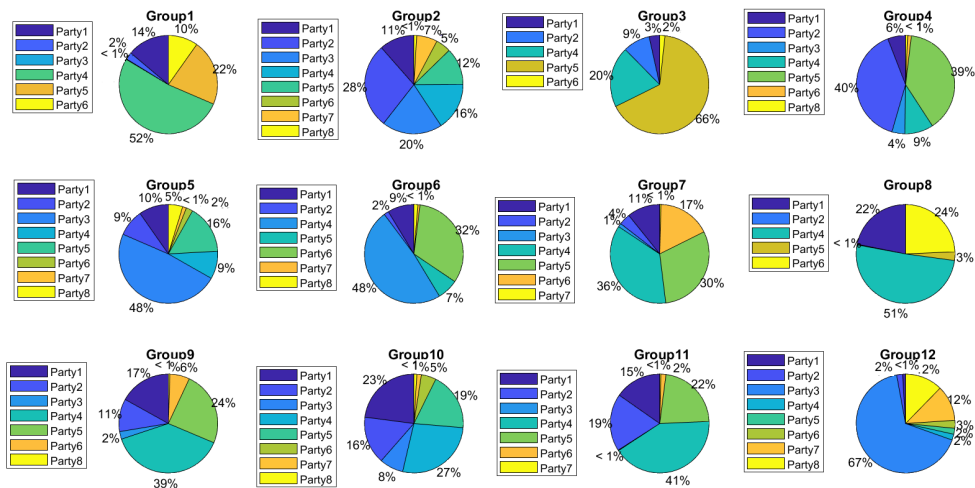
Σε αυτό το Παράρτημα και τα Παραρτήματα που ακολουθούν βρίσκεται η κατανομή χρηστών σε ομάδες σύμφωνα με το βαθμό ομοιότητας μεταξύ των απαντήσεων τους στις δηλώσεις πολιτικής.



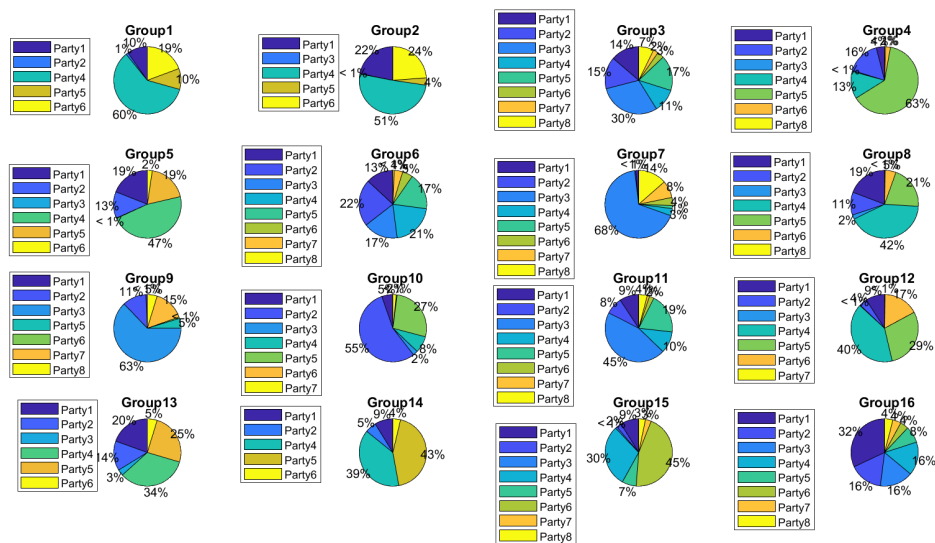
Σχήμα V.1: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Αυστρία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες



Σχήμα V.2: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 8 ομάδες

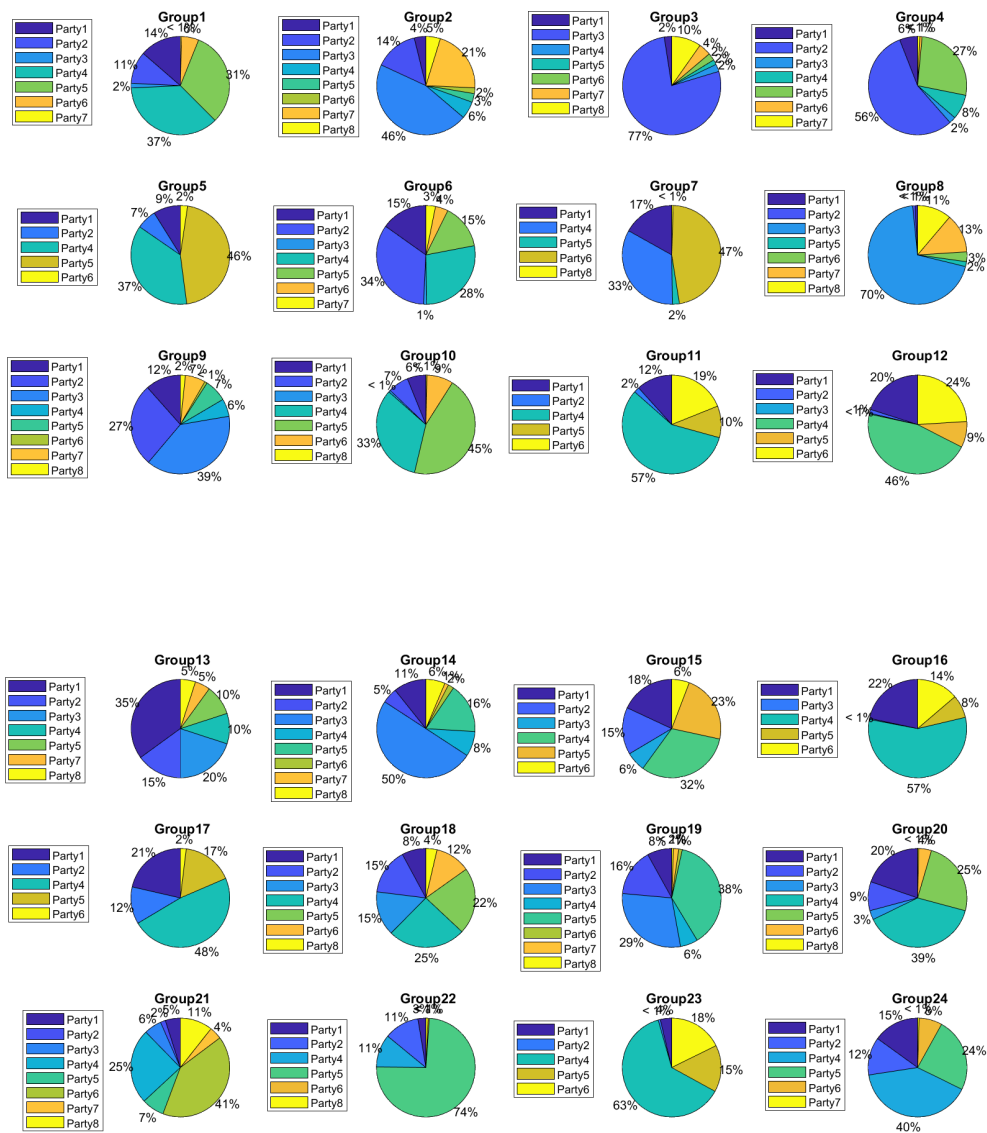


Σχήμα V.3: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 12 ομάδες

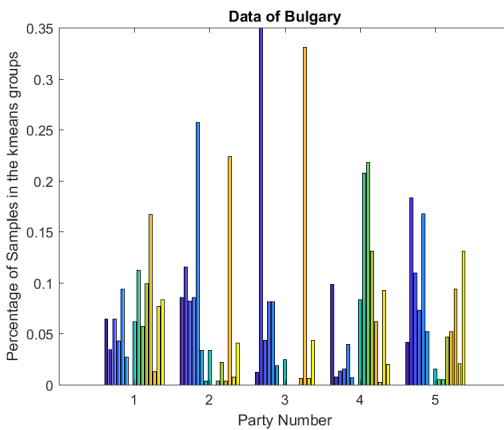
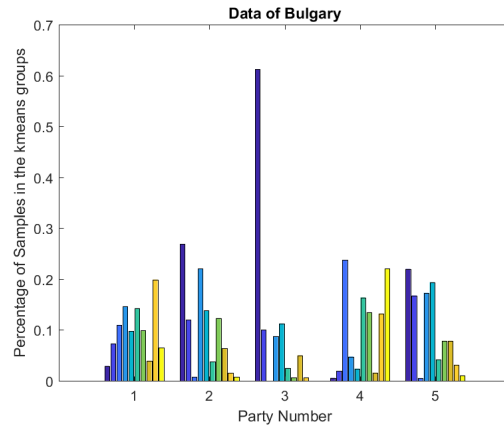
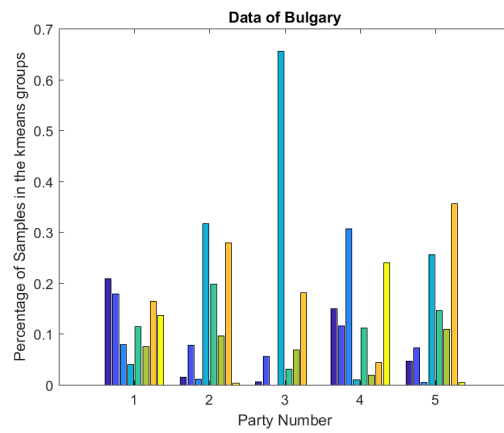
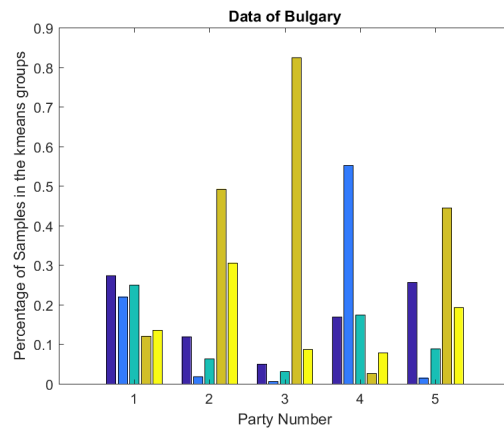


Σχήμα V.4: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 16 ομάδες

ΠΑΡΑΡΤΗΜΑ V. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ

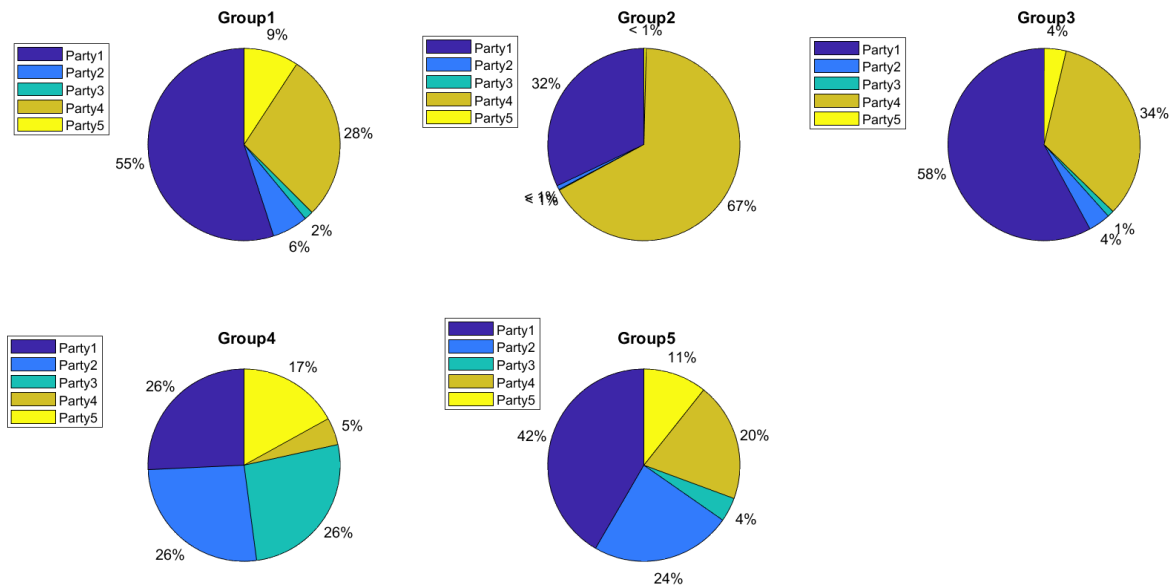


Σχήμα V.5: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Αυστρία – 24 ομάδες

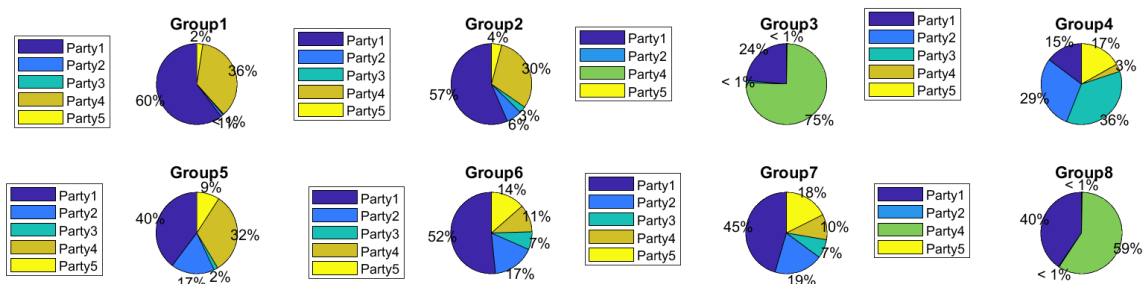


Σχήμα V.6: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Βουλγαρία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες

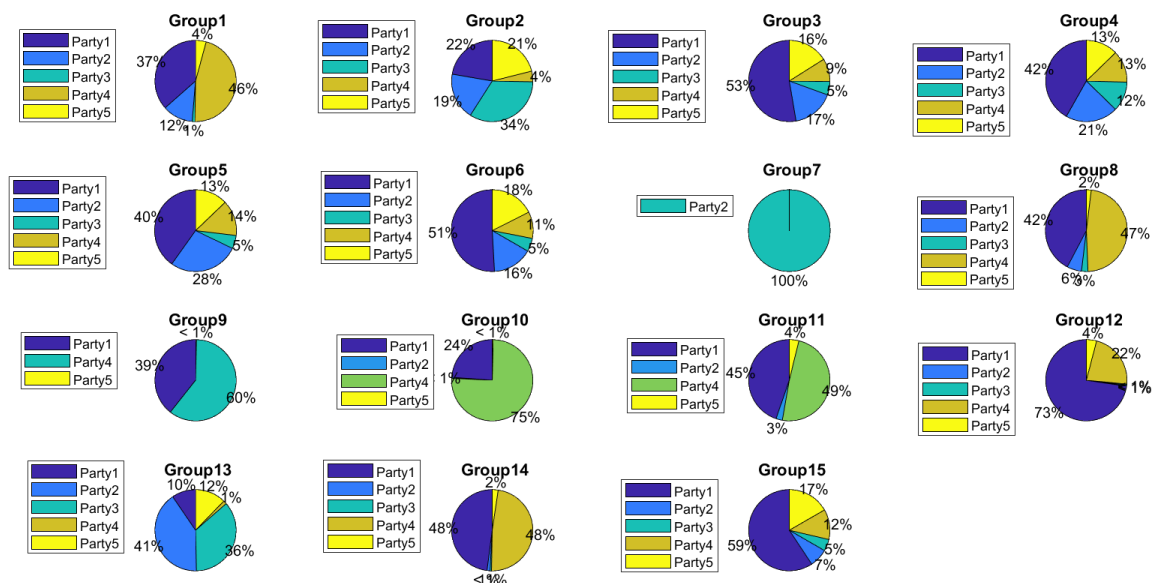
ΠΑΡΑΡΤΗΜΑ V. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ



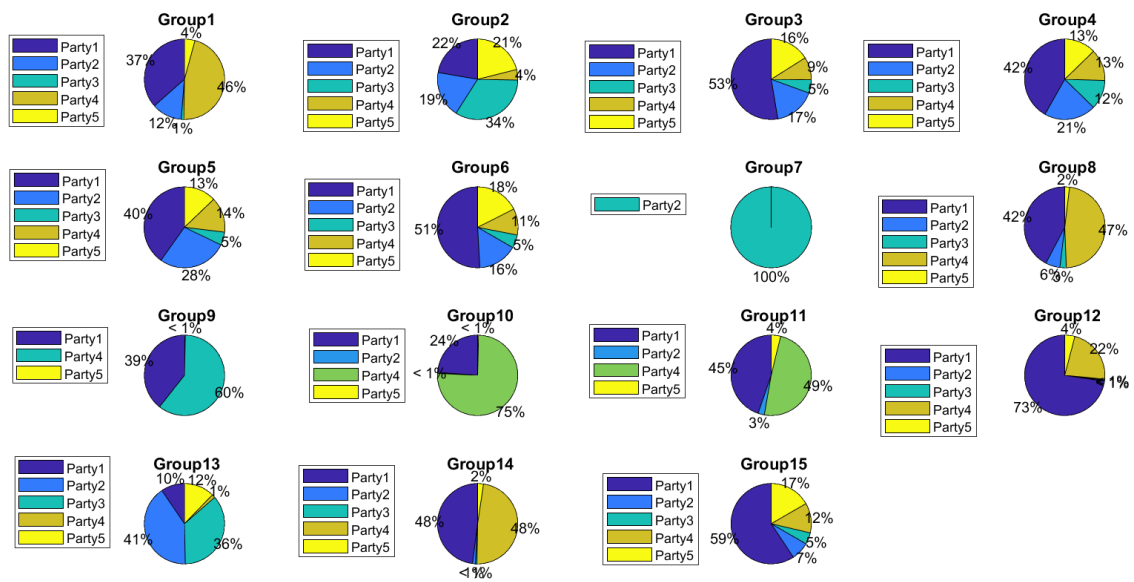
Σχήμα V.7: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 5 ομάδες



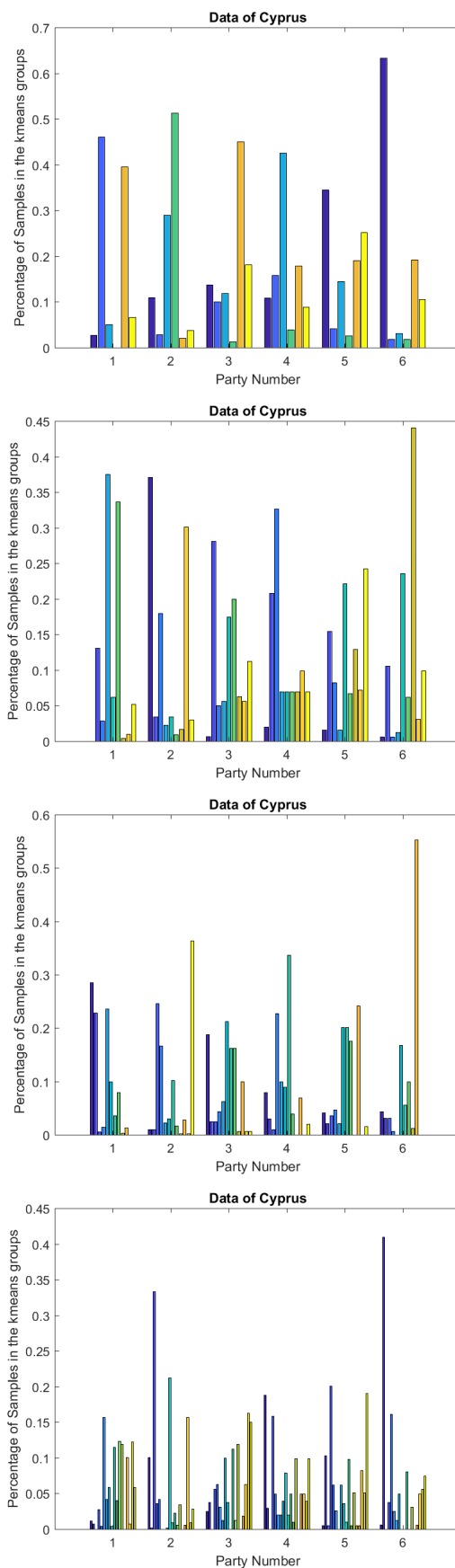
Σχήμα V.8: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 8 ομάδες



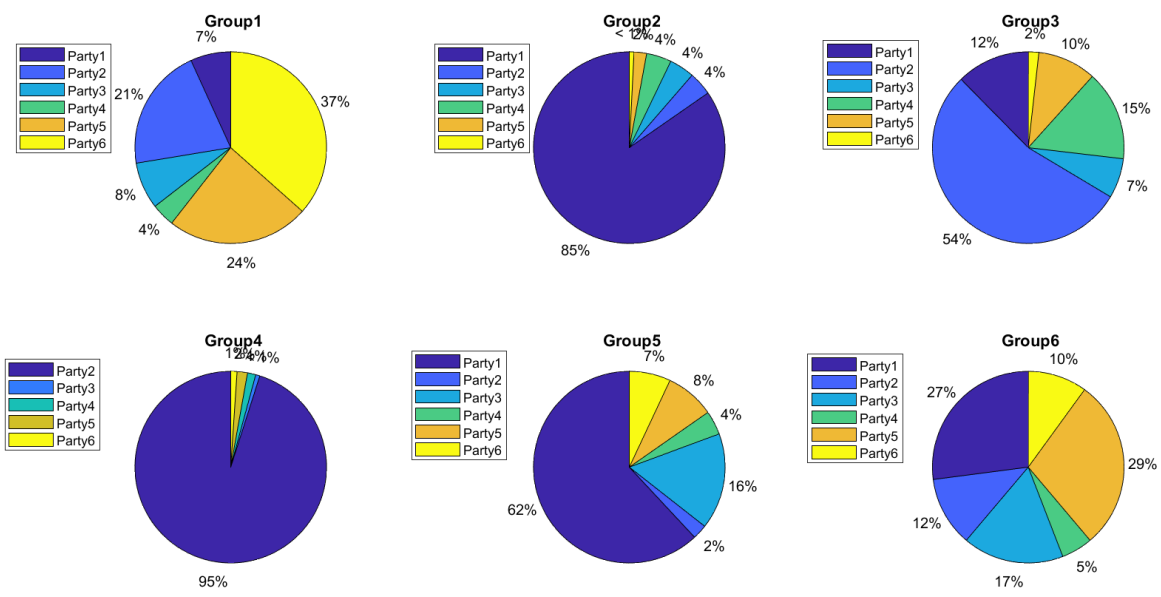
Σχήμα V.9: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 10 ομάδες



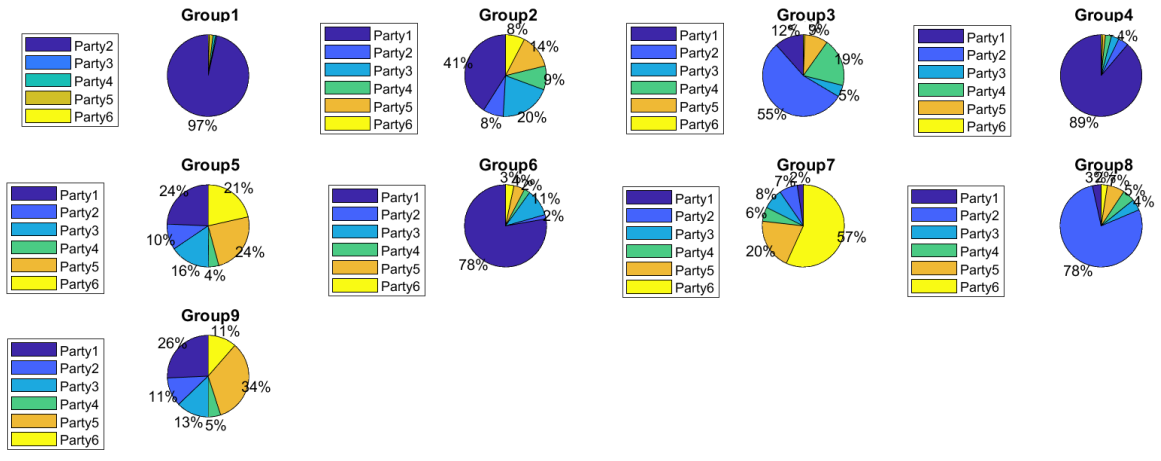
Σχήμα V.10: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Βουλγαρία – 15 ομάδες



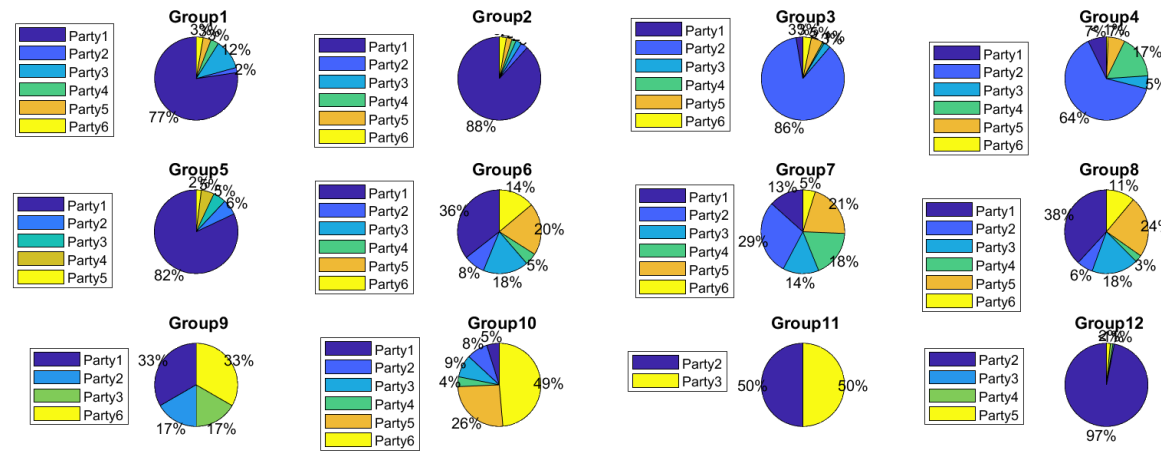
Σχήμα V.11: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Κύπρος (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες



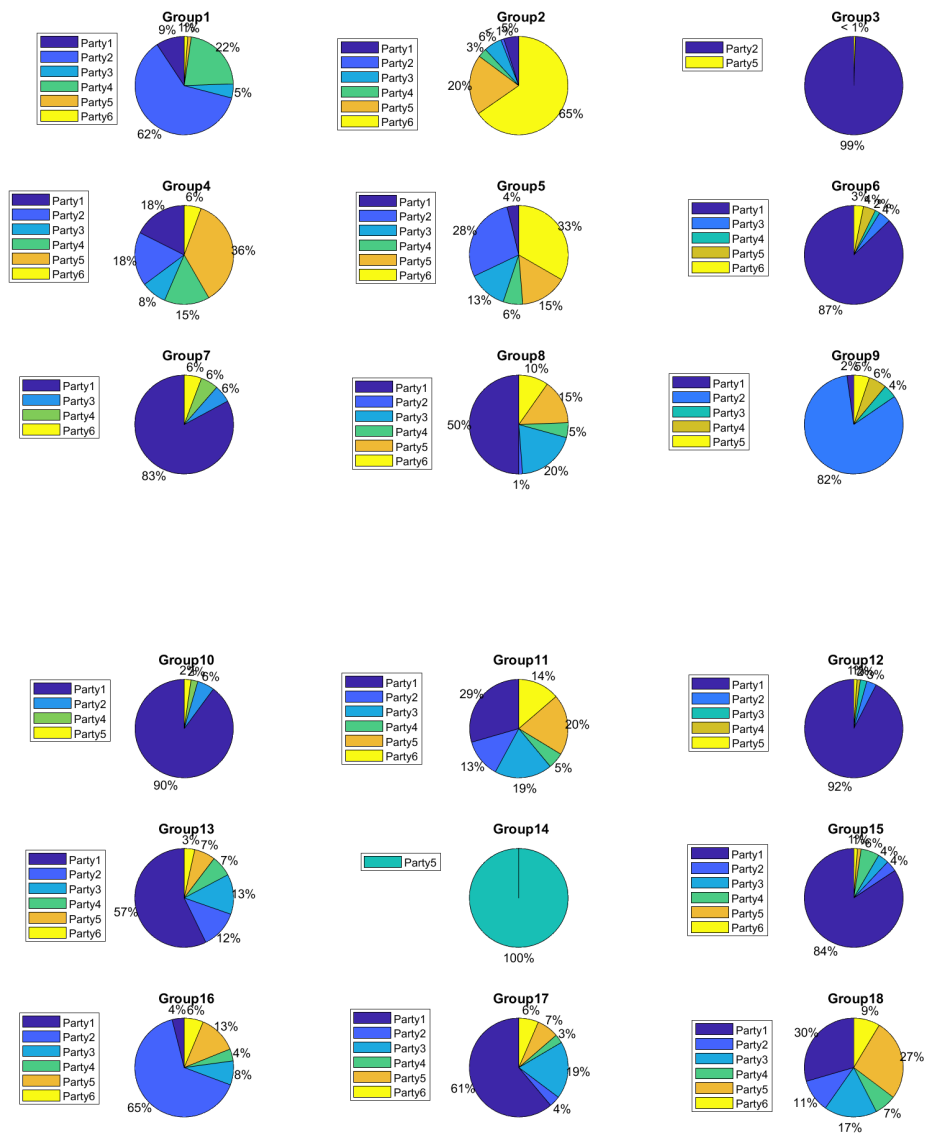
Σχήμα V.12: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 6 ομάδες



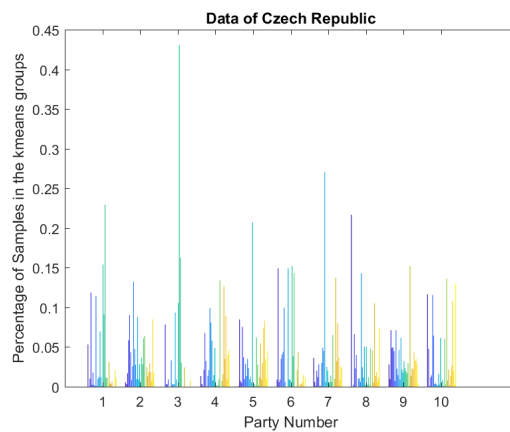
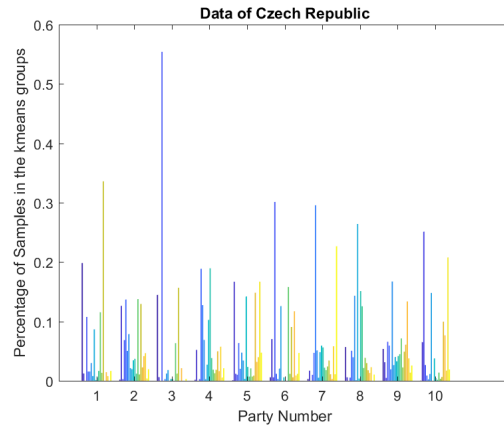
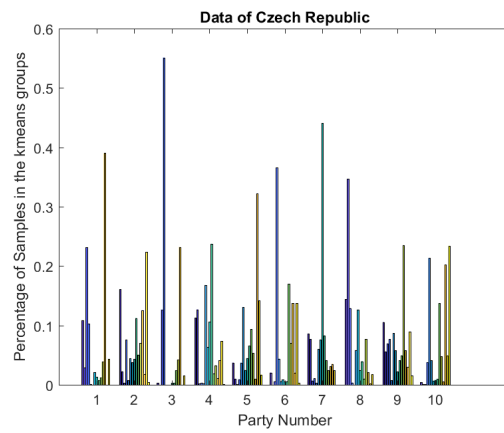
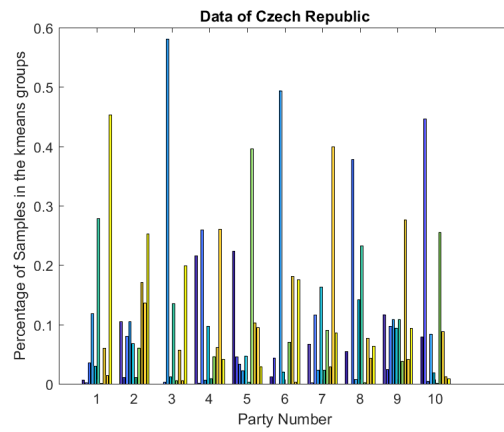
Σχήμα V.13: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 9 ομάδες



Σχήμα V.14: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 12 ομάδες

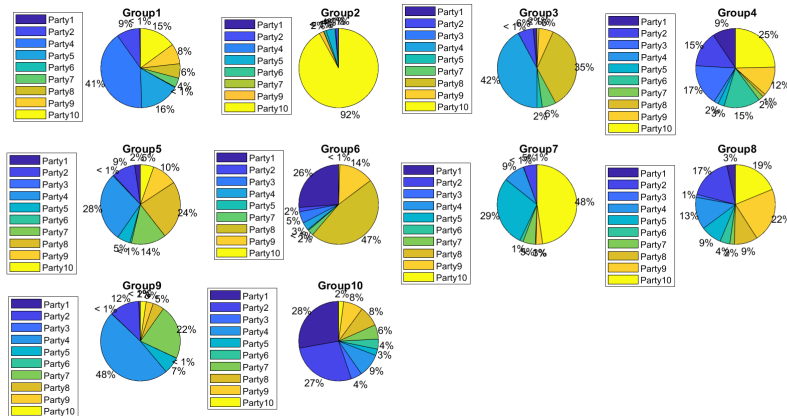


Σχήμα V.15: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κύπρος – 18 ομάδες

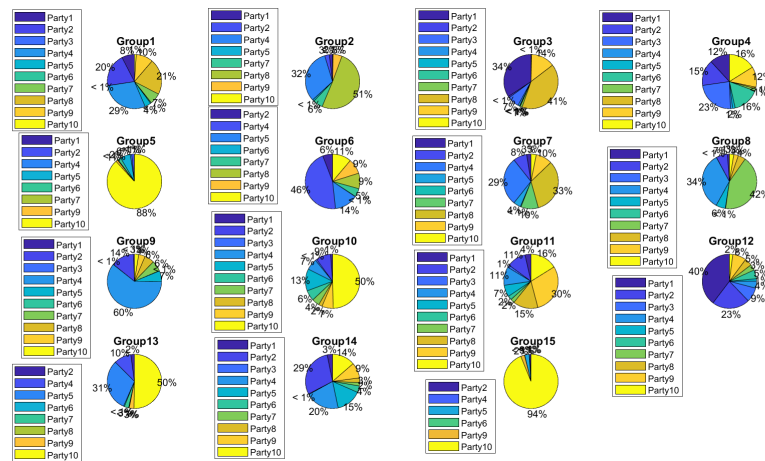


Σχήμα V.16: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Τσεχία (α) 10 ομάδες (β) 15 ομάδες (γ) 20 ομάδες (δ) 30 ομάδες

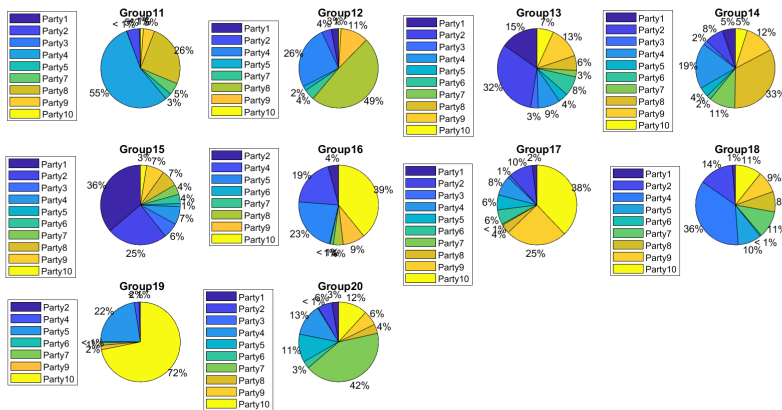
ΠΑΡΑΡΤΗΜΑ V. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ



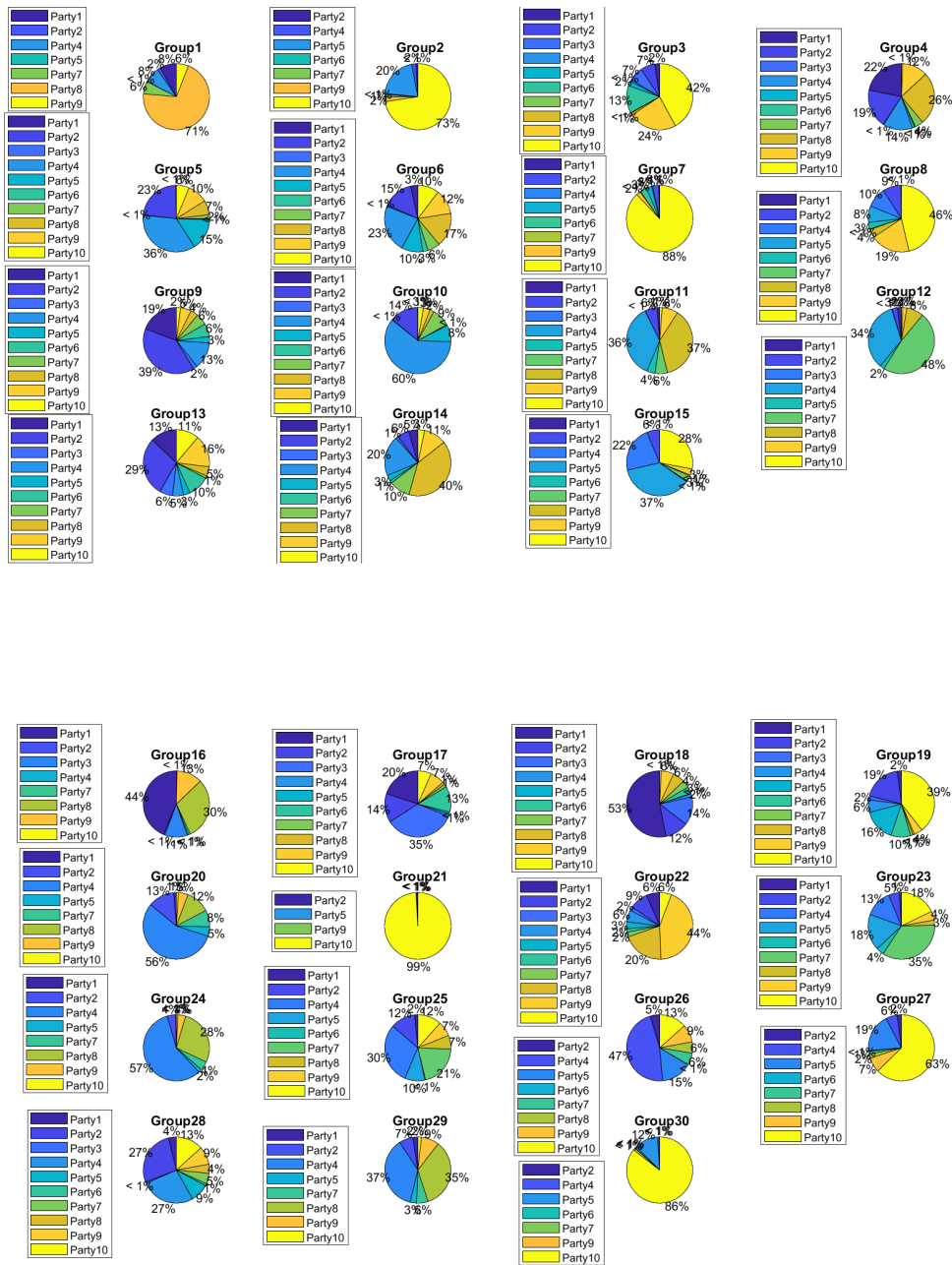
Σχήμα V.17: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 10 ομάδες



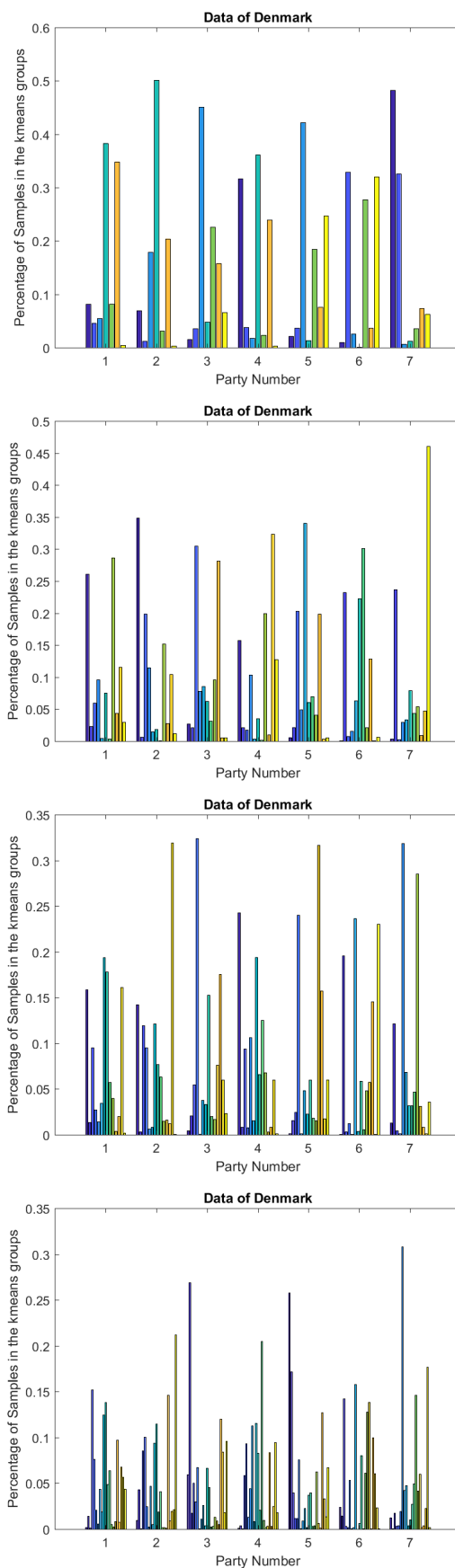
Σχήμα V.18: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 15 ομάδες



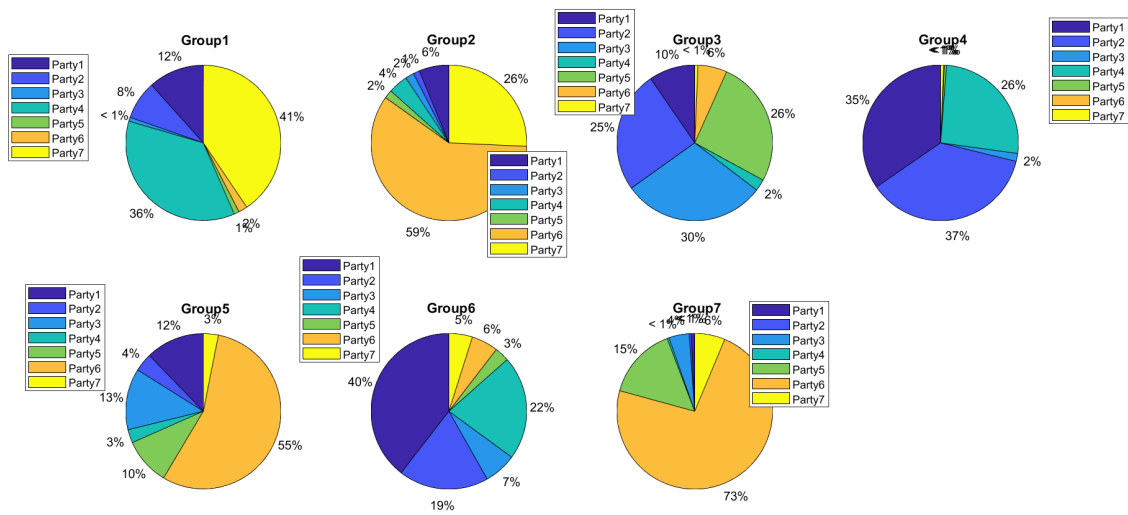
Σχήμα V.19: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Τσεχία – 20 ομάδες



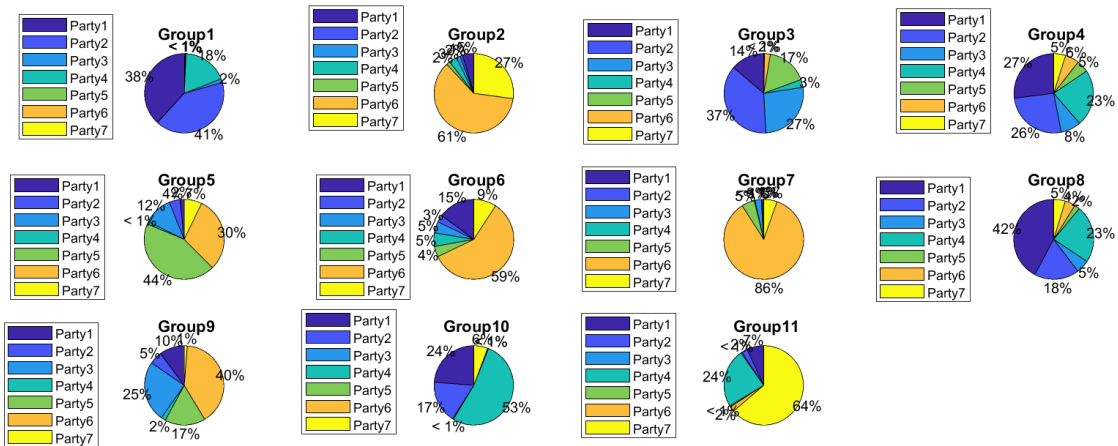
Σχήμα V.20: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα –Τσεχία – 30 ομάδες



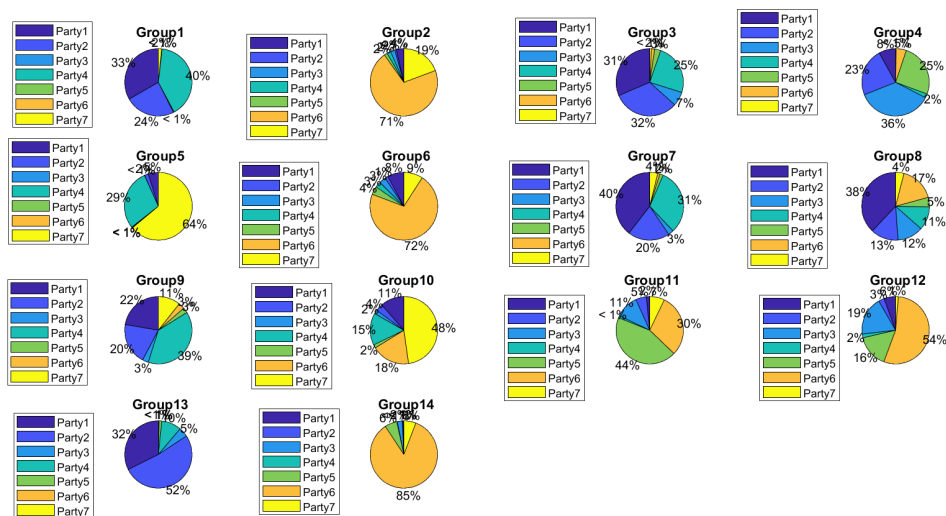
Σχήμα V.21: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Δανία (α) 7 ομάδες (β) 11 ομάδες (γ) 14 ομάδες (δ) 21 ομάδες



Σχήμα V.22: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 7 ομάδες

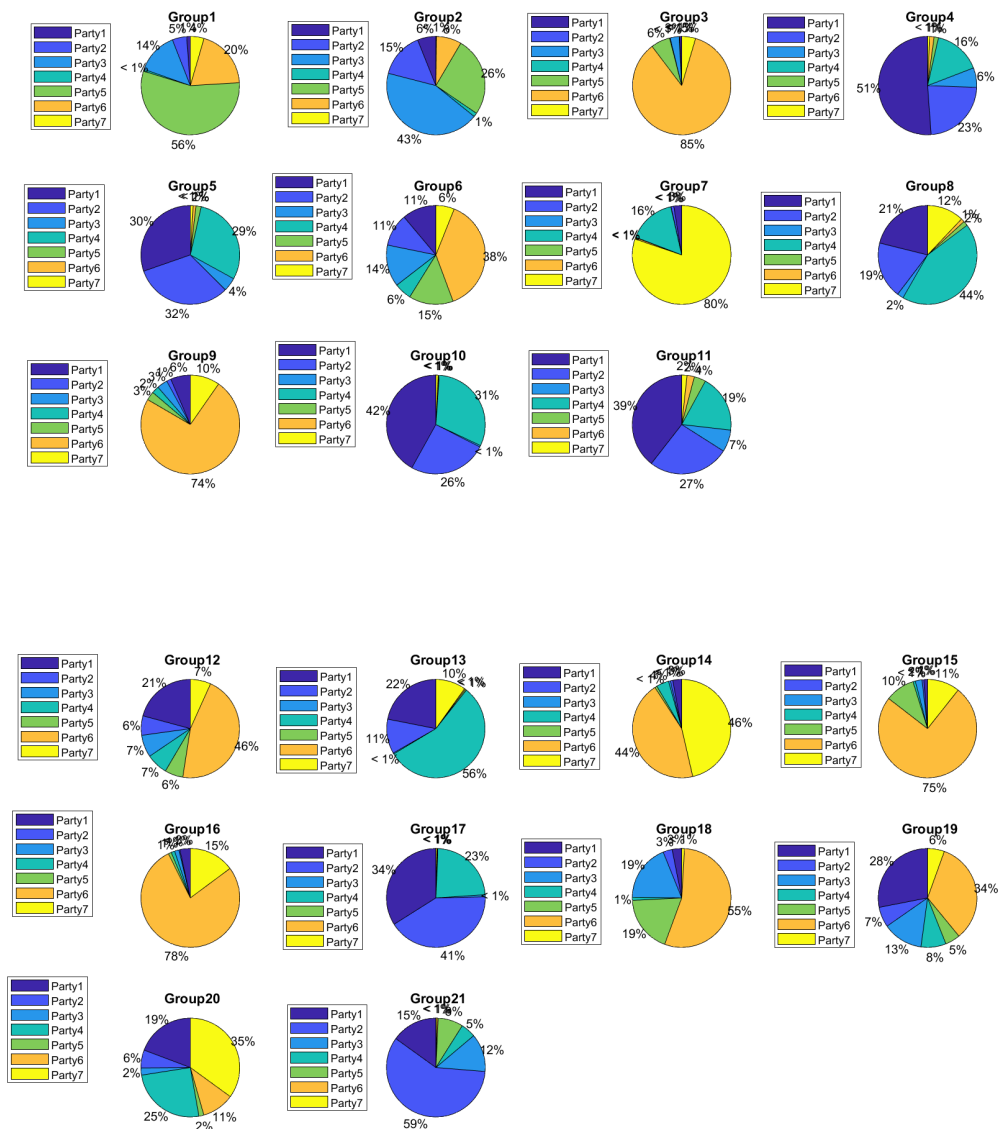


Σχήμα V.23: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 11 ομάδες

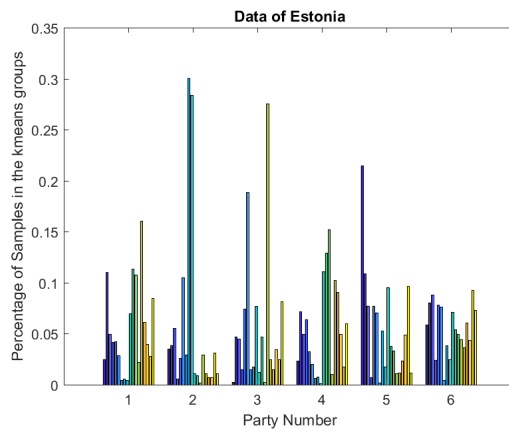
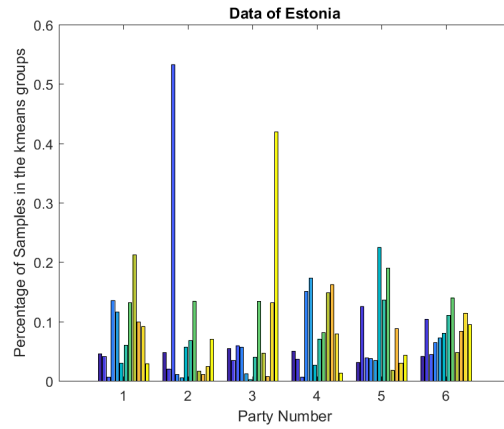
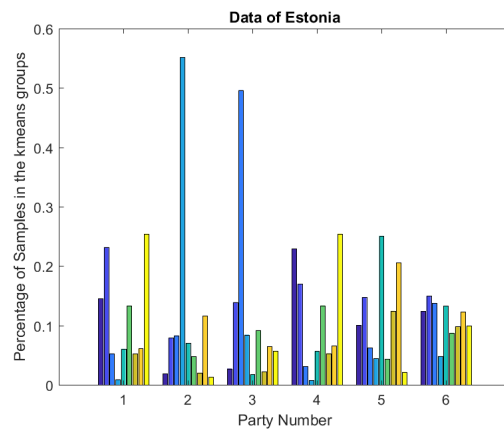
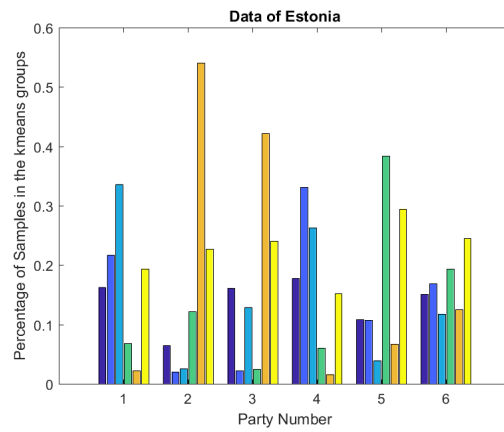


Σχήμα V.24: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 14 ομάδες

ΠΑΡΑΡΤΗΜΑ V. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ

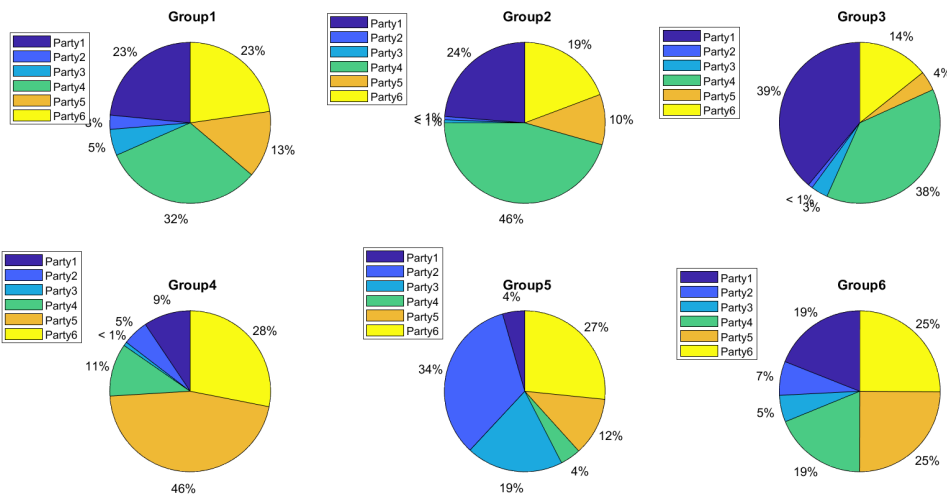


Σχήμα V.25: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Δανία – 21 ομάδες

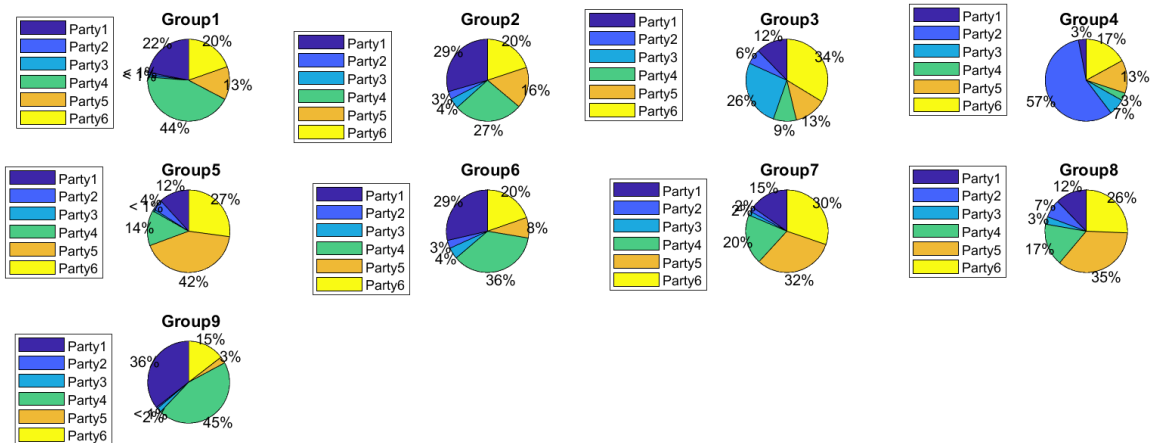


Σχήμα V.26: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Εσθονία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες

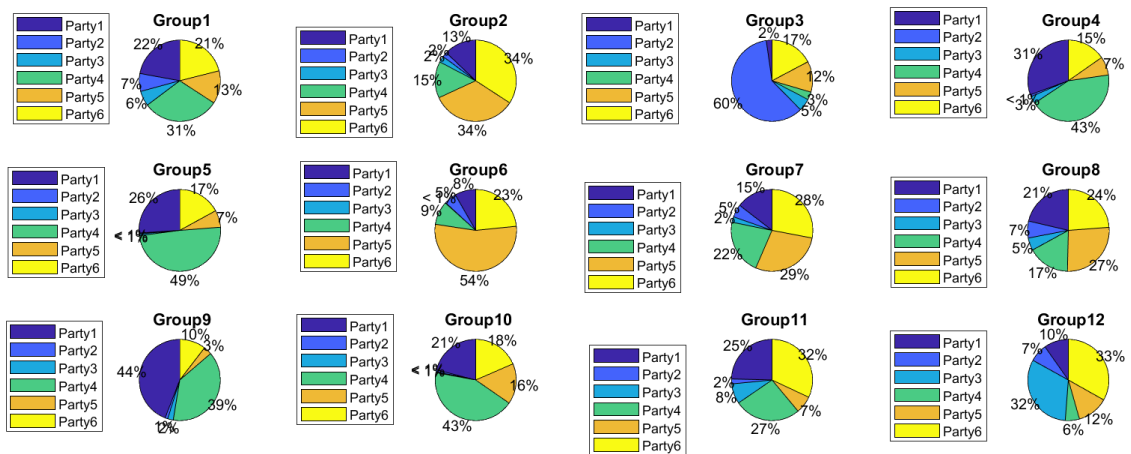
ΠΑΡΑΡΤΗΜΑ V. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ



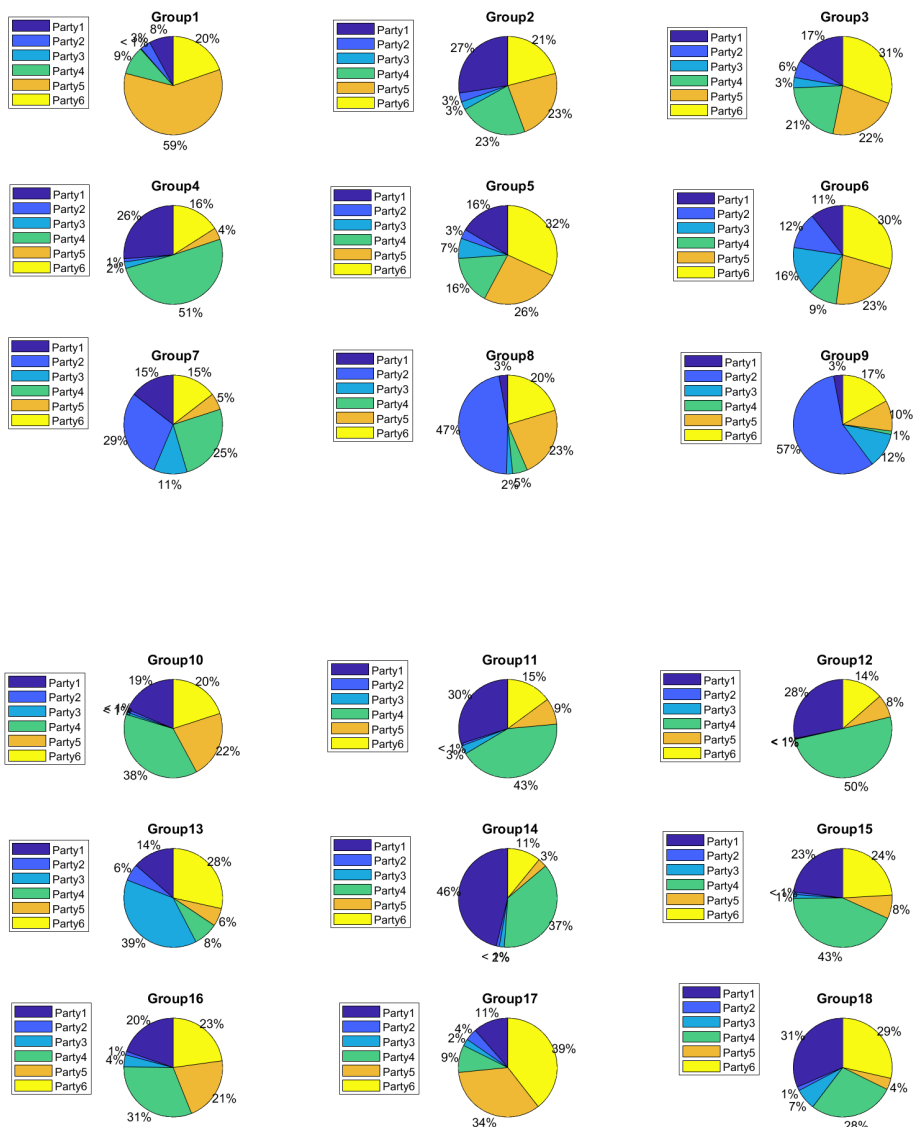
Σχήμα V.27: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 6 ομάδες



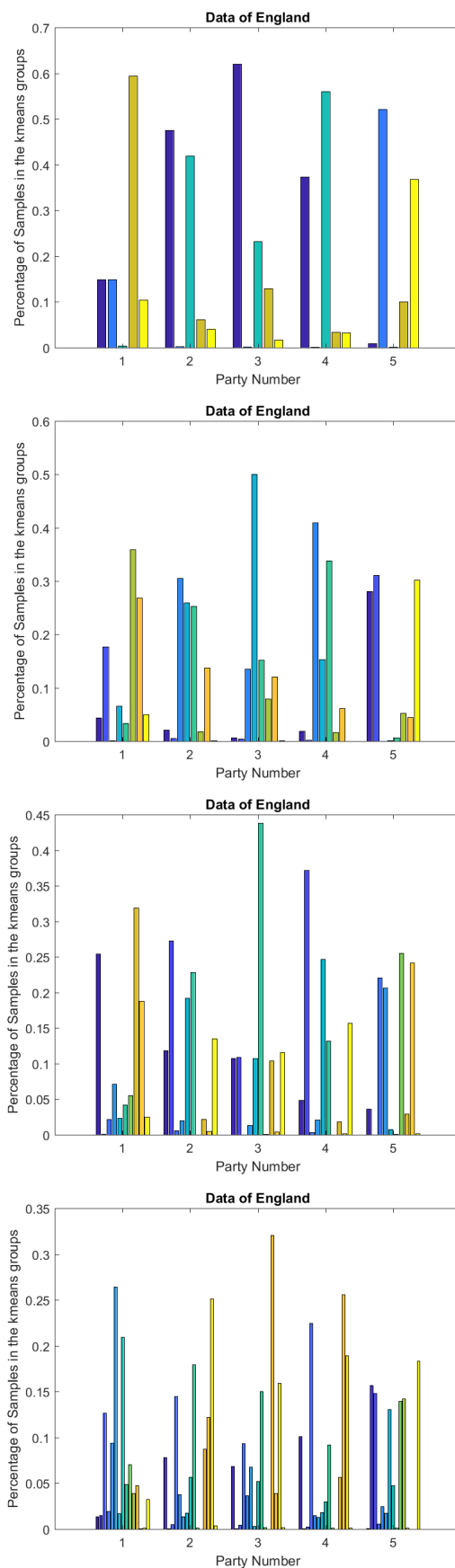
Σχήμα V.28: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 9 ομάδες



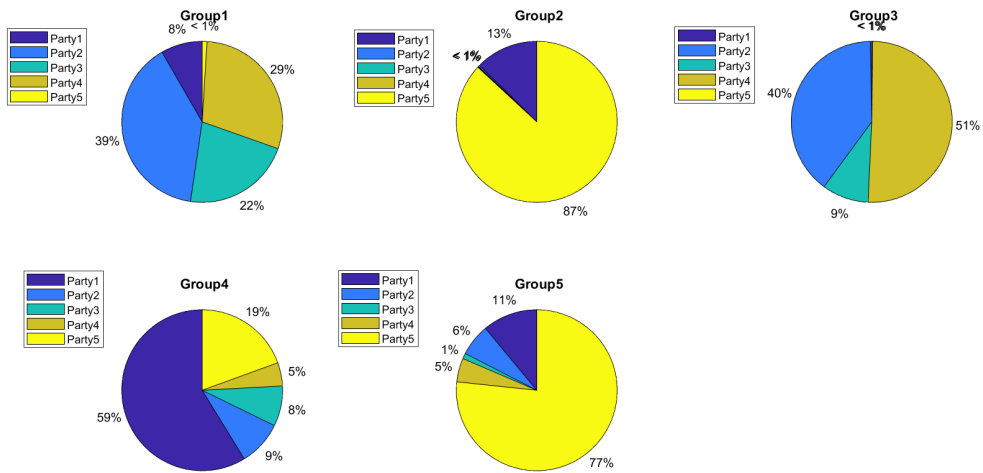
Σχήμα V.29: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Εσθονία – 12 ομάδες



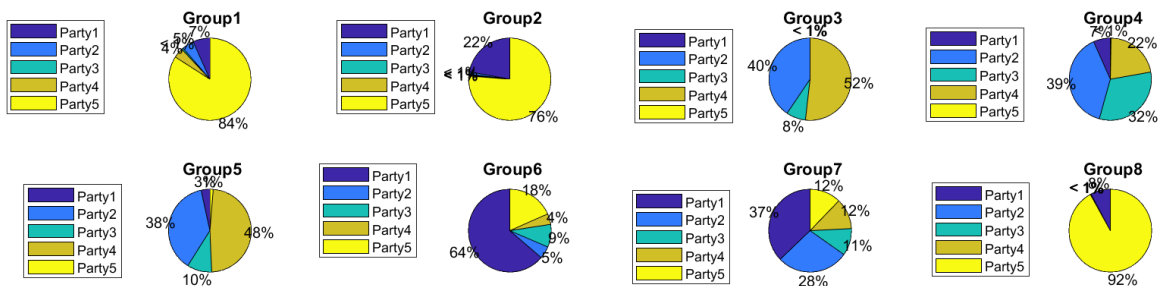
Σχήμα V.30: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα –Εσθονία – 18 ομάδες



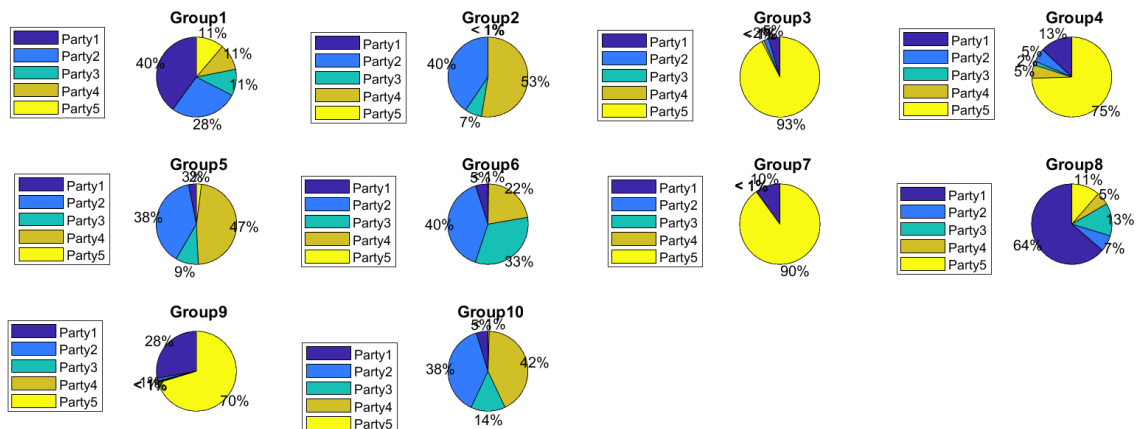
Σχήμα V.31: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες



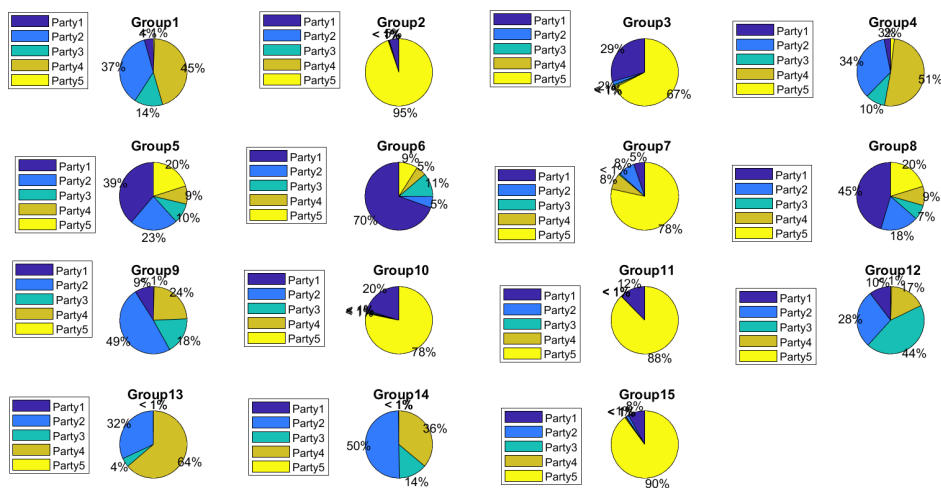
Σχήμα V.32: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 5 ομάδες



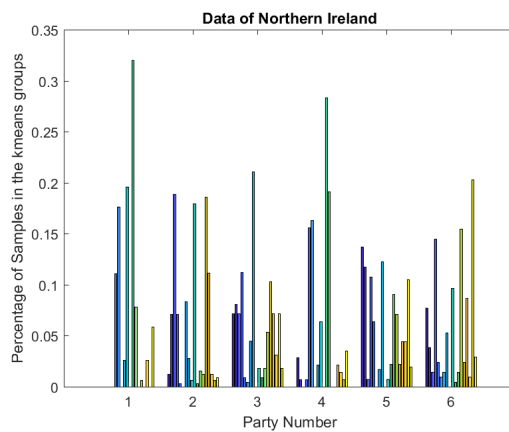
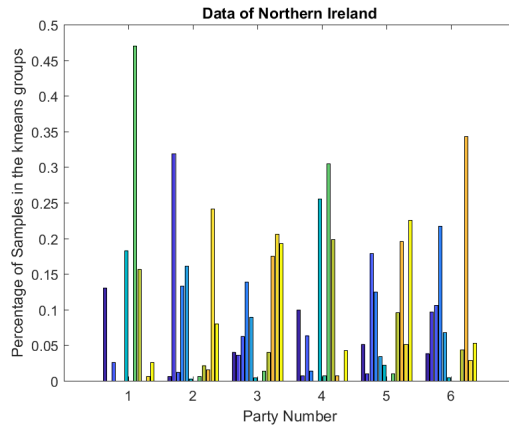
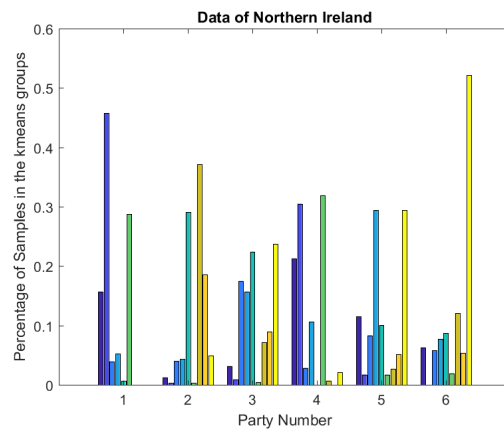
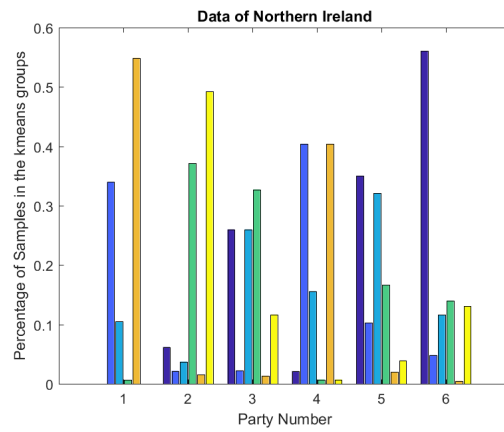
Σχήμα V.33: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 8 ομάδες



Σχήμα V.34: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 10 ομάδες

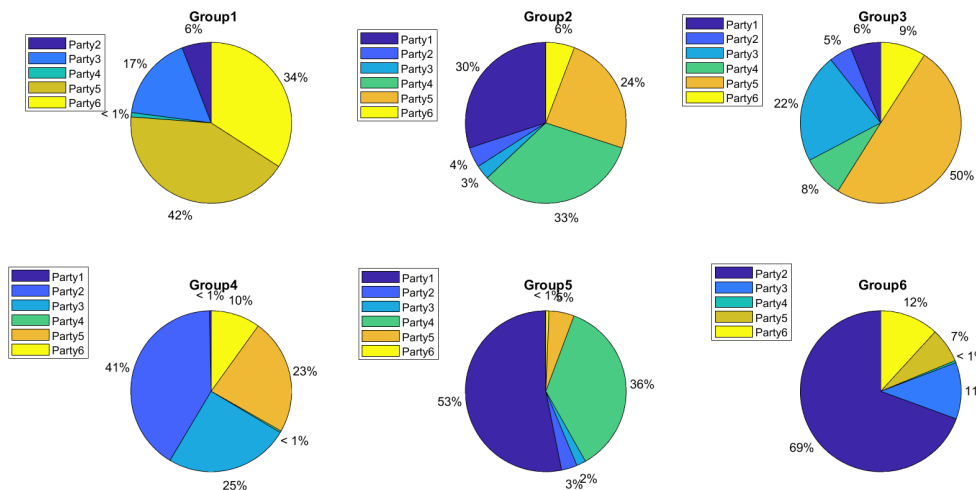


Σχήμα V.35: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Αγγλία – 15 ομάδες

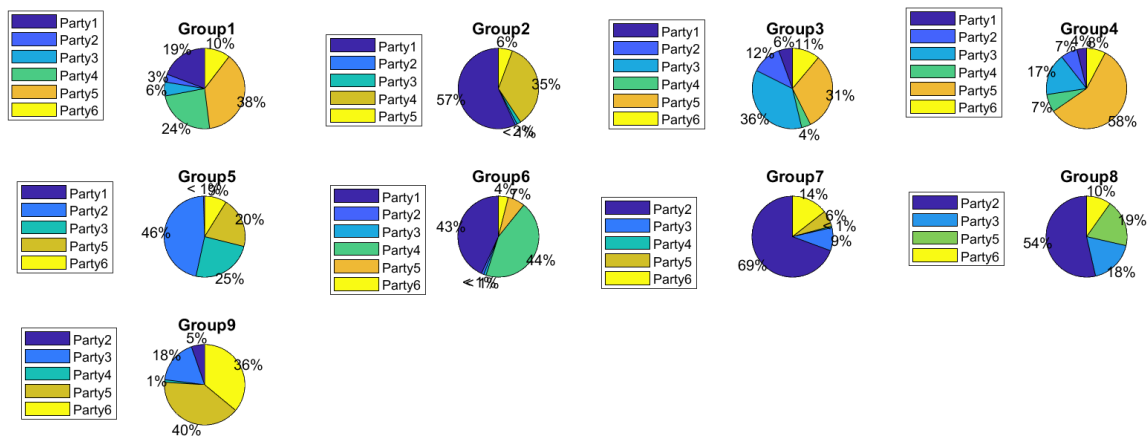


Σχήμα V.36: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες

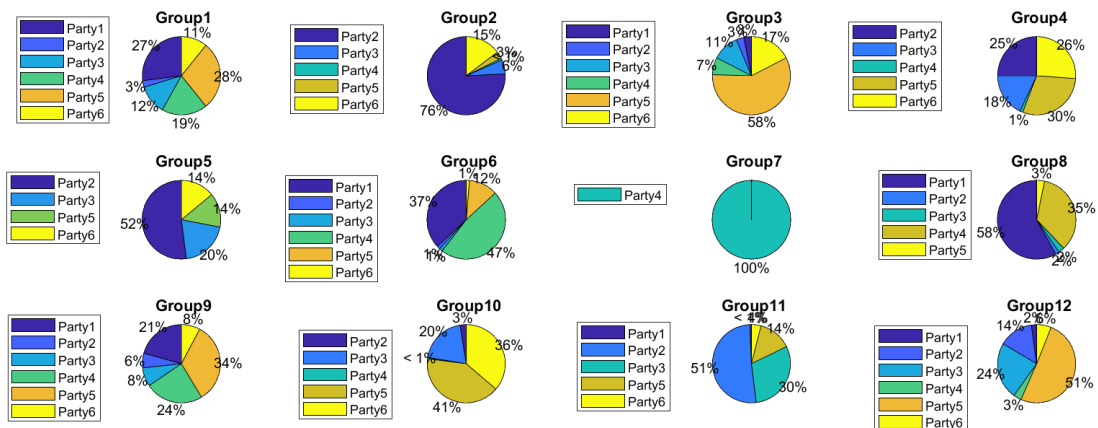
ΠΑΡΑΡΤΗΜΑ V. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ



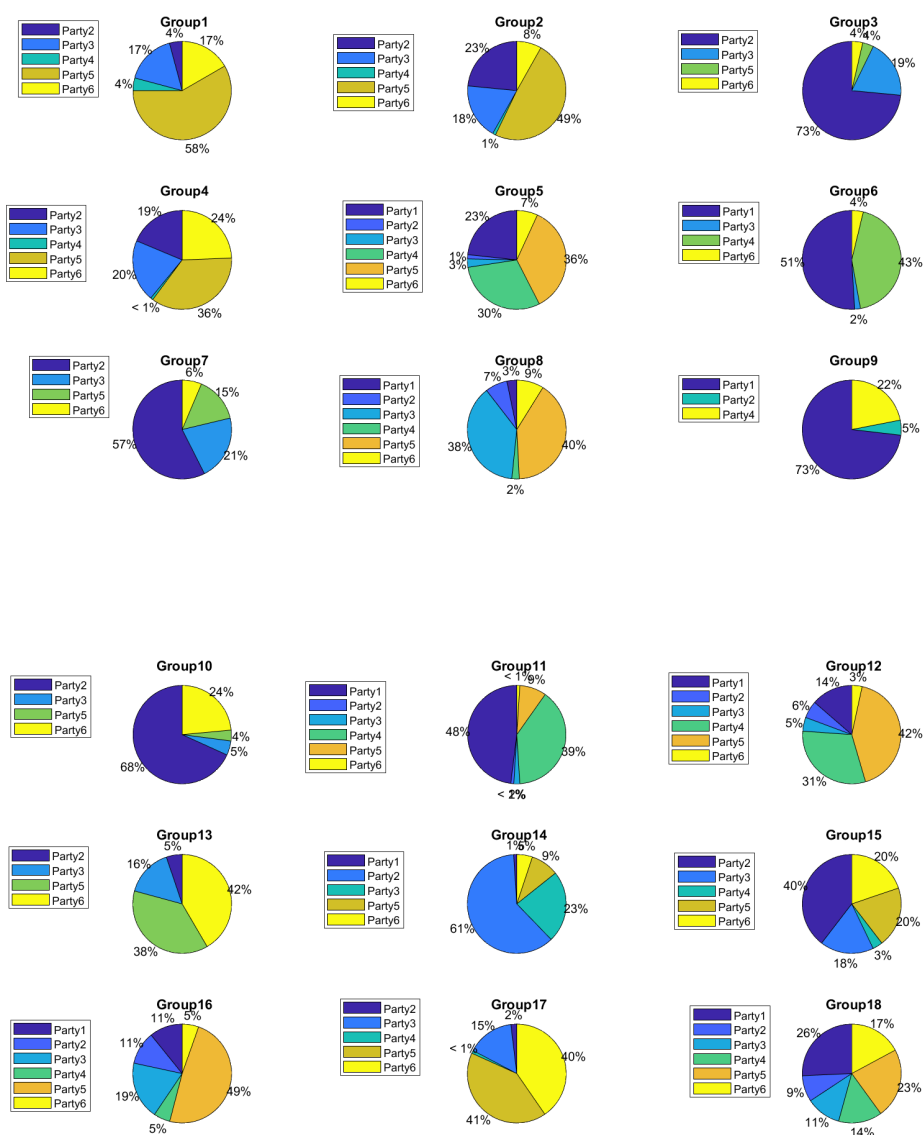
Σχήμα V.37: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 6 ομάδες



Σχήμα V.38: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 9 ομάδες



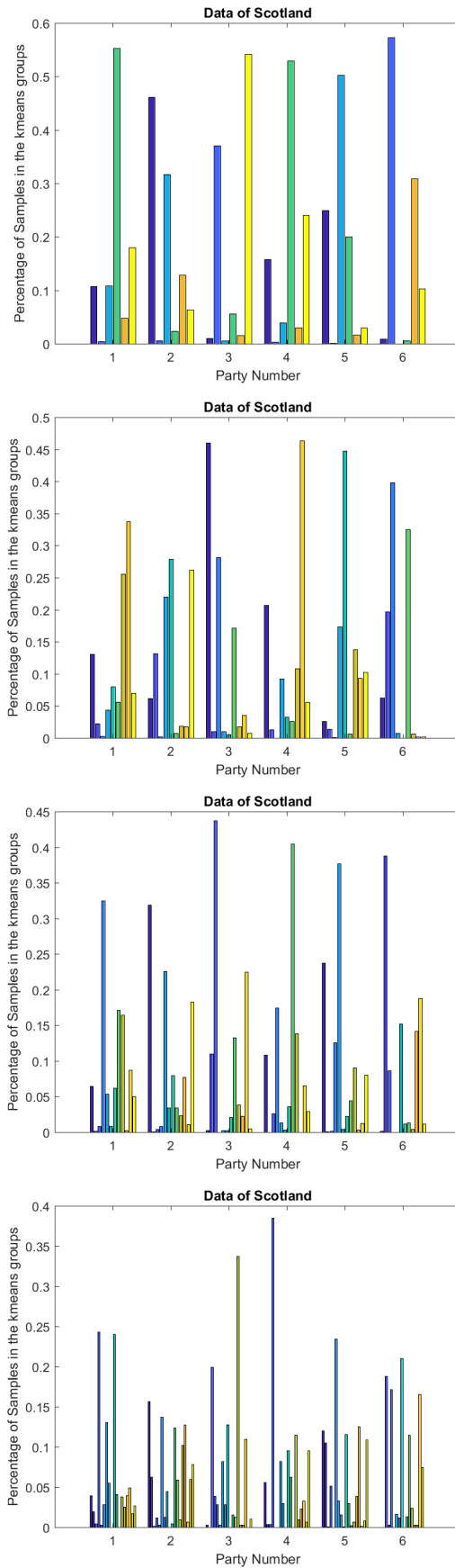
Σχήμα V.39: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 12 ομάδες



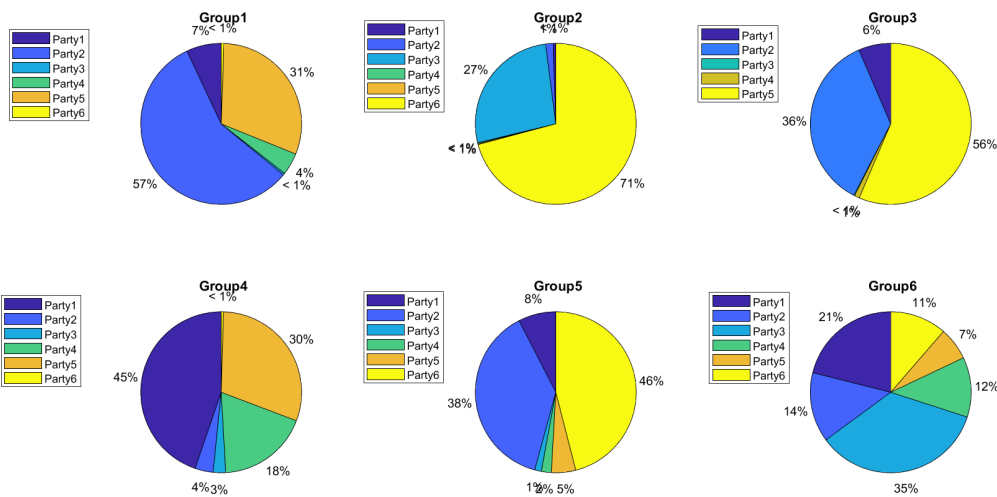
Σχήμα V.40: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Β. Ιρλανδία – 18 ομάδες

Παράρτημα VI

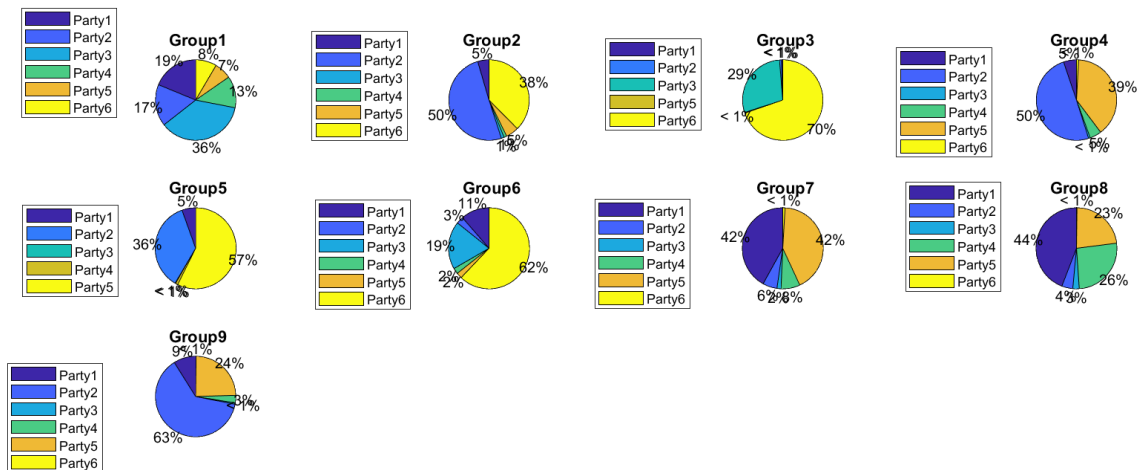
Αποτελέσματα από τη διερεύνηση της ιδεολογικής ταύτισης κόμματος-χρήστη - Μέρος Β



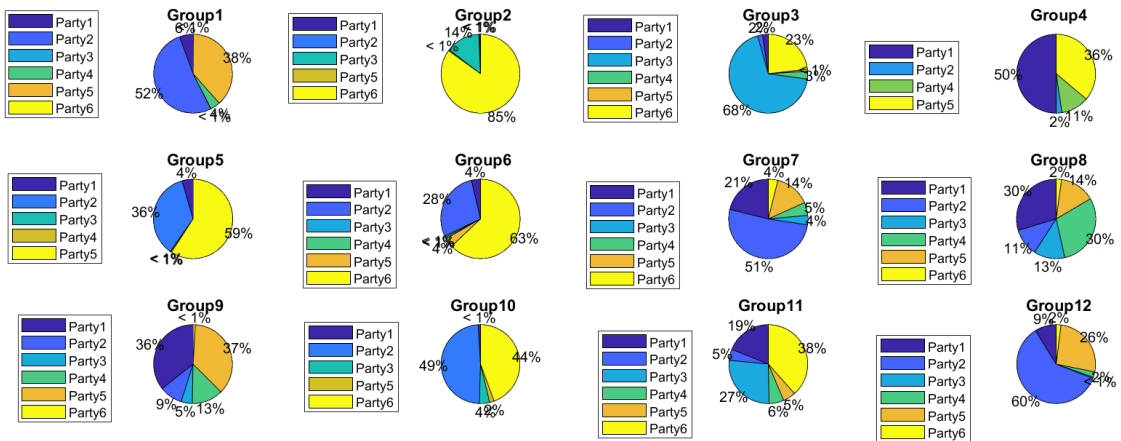
Σχήμα VI.1: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες



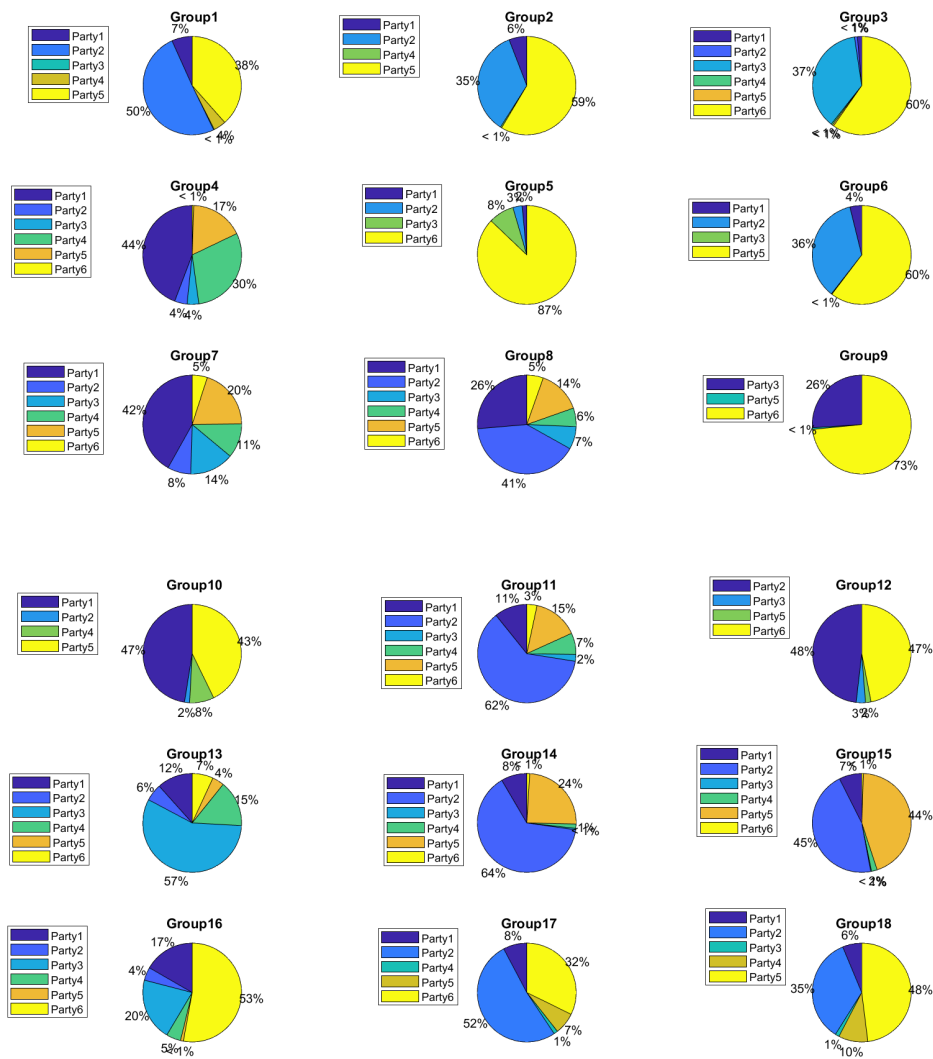
Σχήμα VI.2: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 6 ομάδες



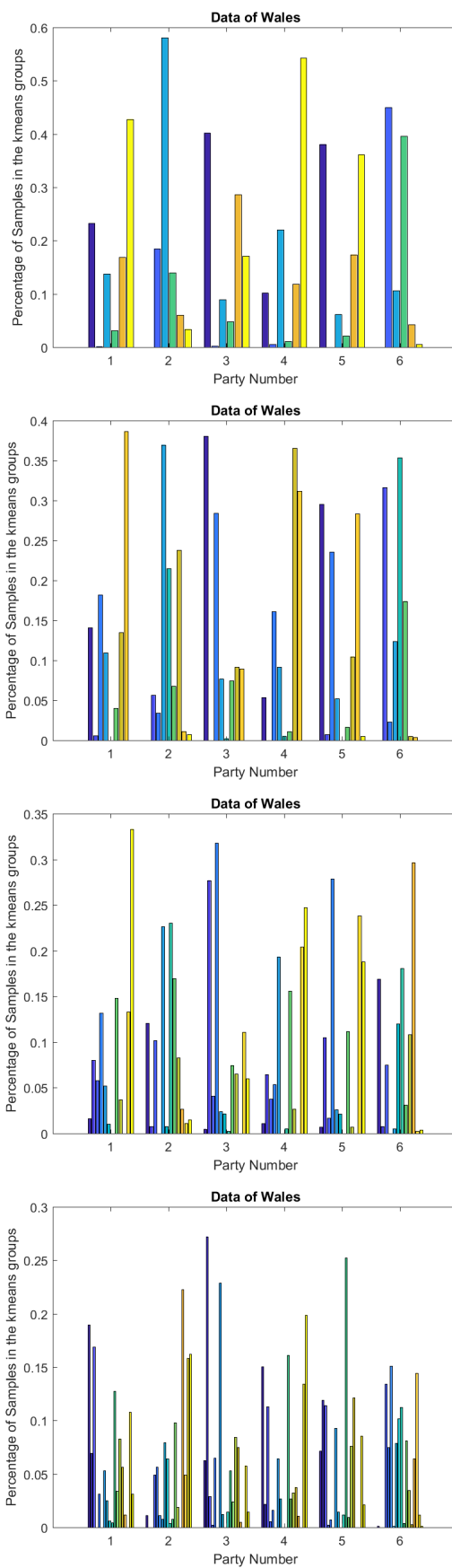
Σχήμα VI.3: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 9 ομάδες



Σχήμα VI.4: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 12 ομάδες

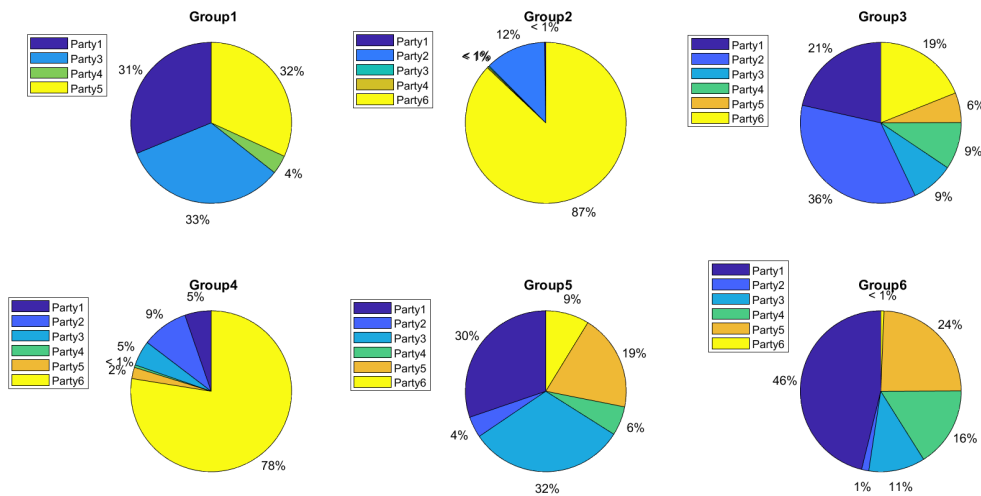


Σχήμα VI.5: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Σκωτία – 18 ομάδες

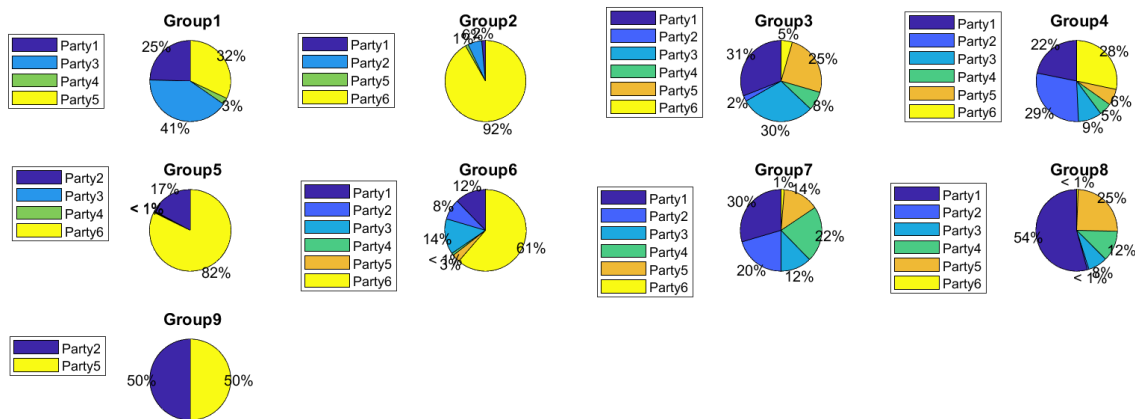


Σχήμα VI.6: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες

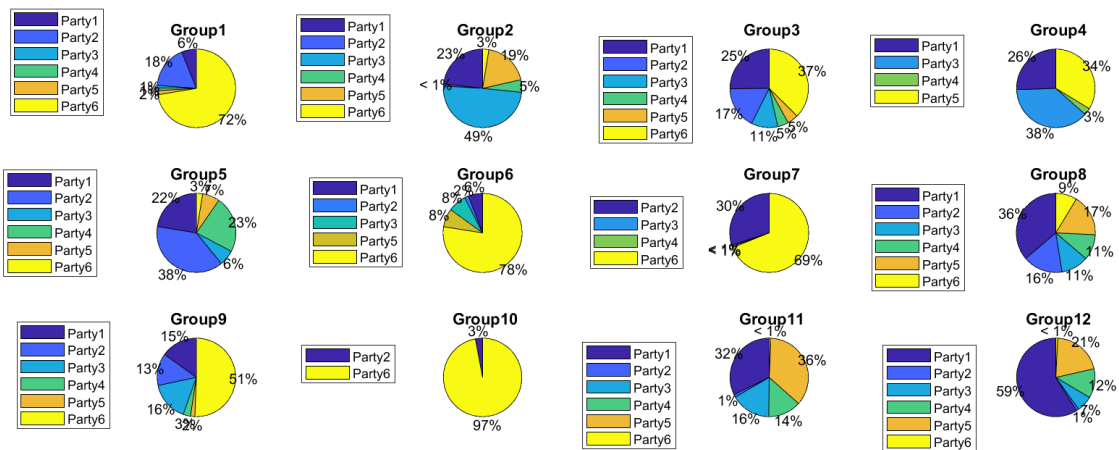
ΠΑΡΑΡΤΗΜΑ VI. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ - ΜΕΡΟΣ Β



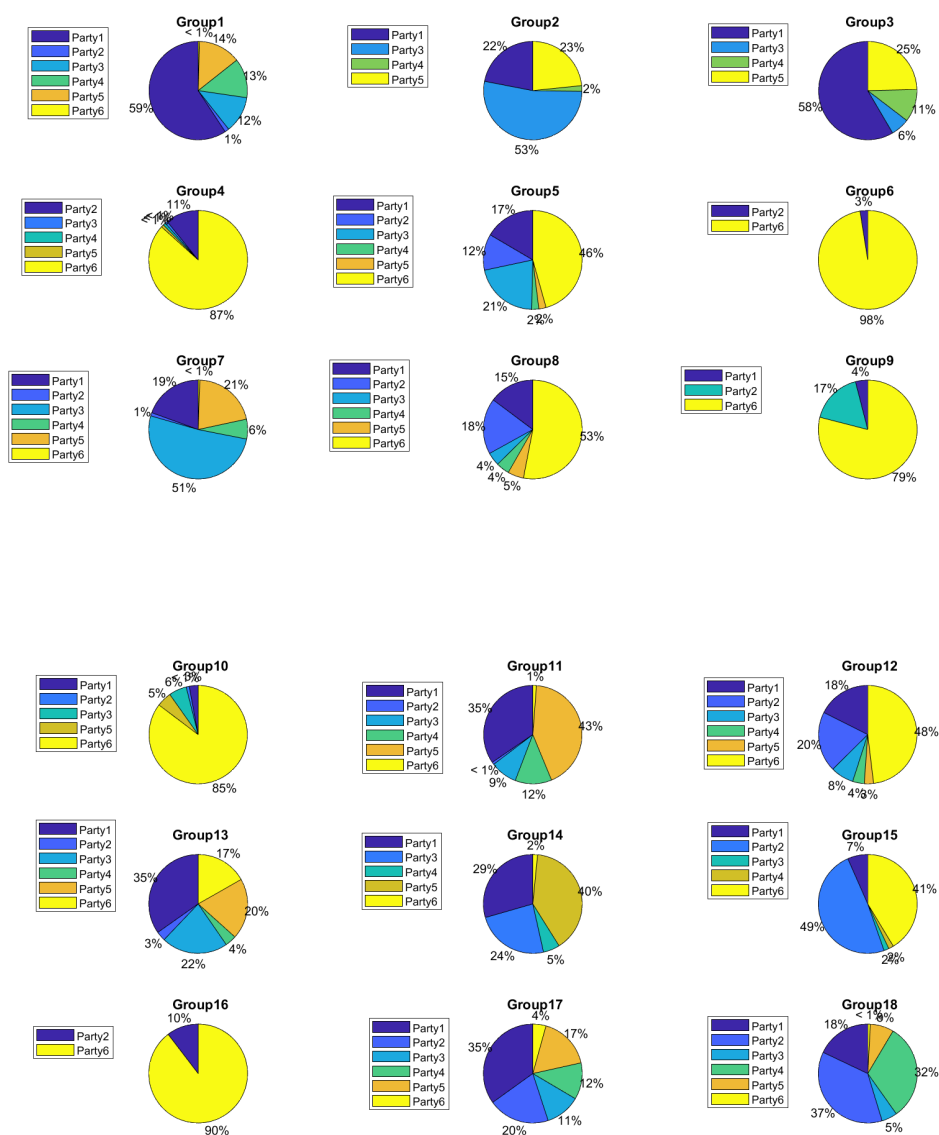
Σχήμα VI.7: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 6 ομάδες



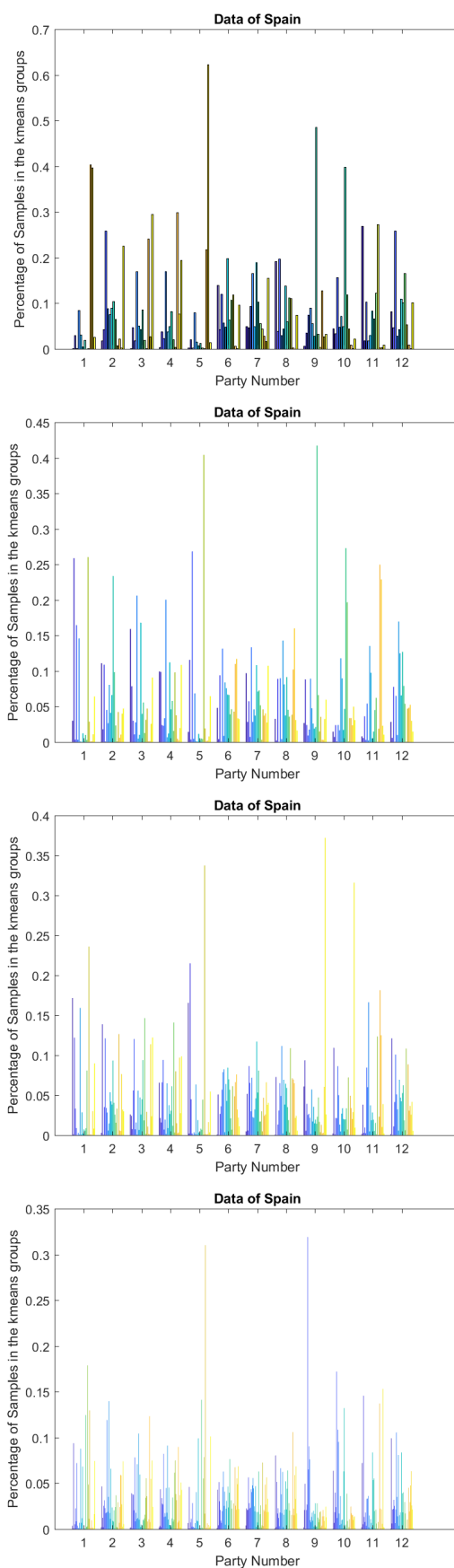
Σχήμα VI.8: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 9 ομάδες



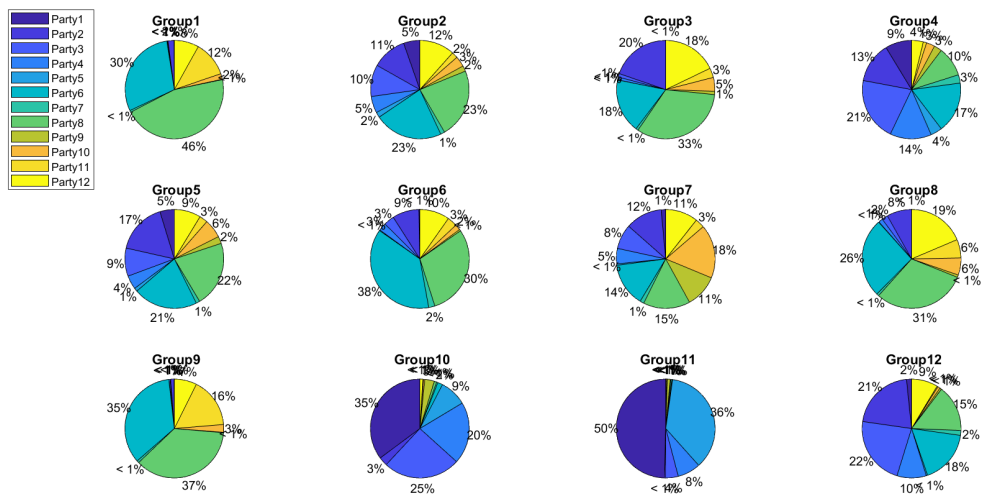
Σχήμα VI.9: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 12 ομάδες



Σχήμα VI.10: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ηνωμένο Βασίλειο / Ουαλία – 18 ομάδες

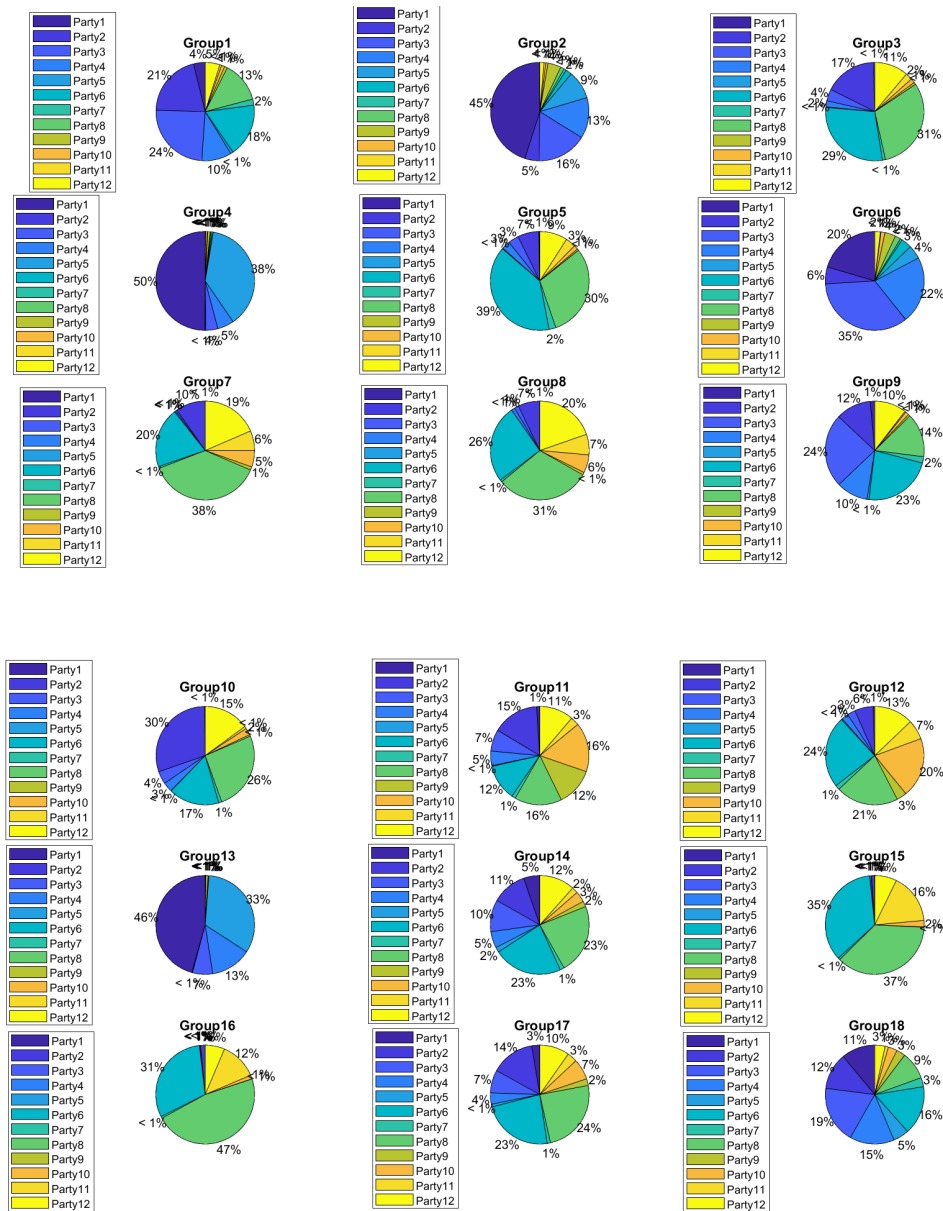


Σχήμα VI.11: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ισπανία (α) 12 ομάδες (β) 18 ομάδες (γ) 24 ομάδες (δ) 36 ομάδες

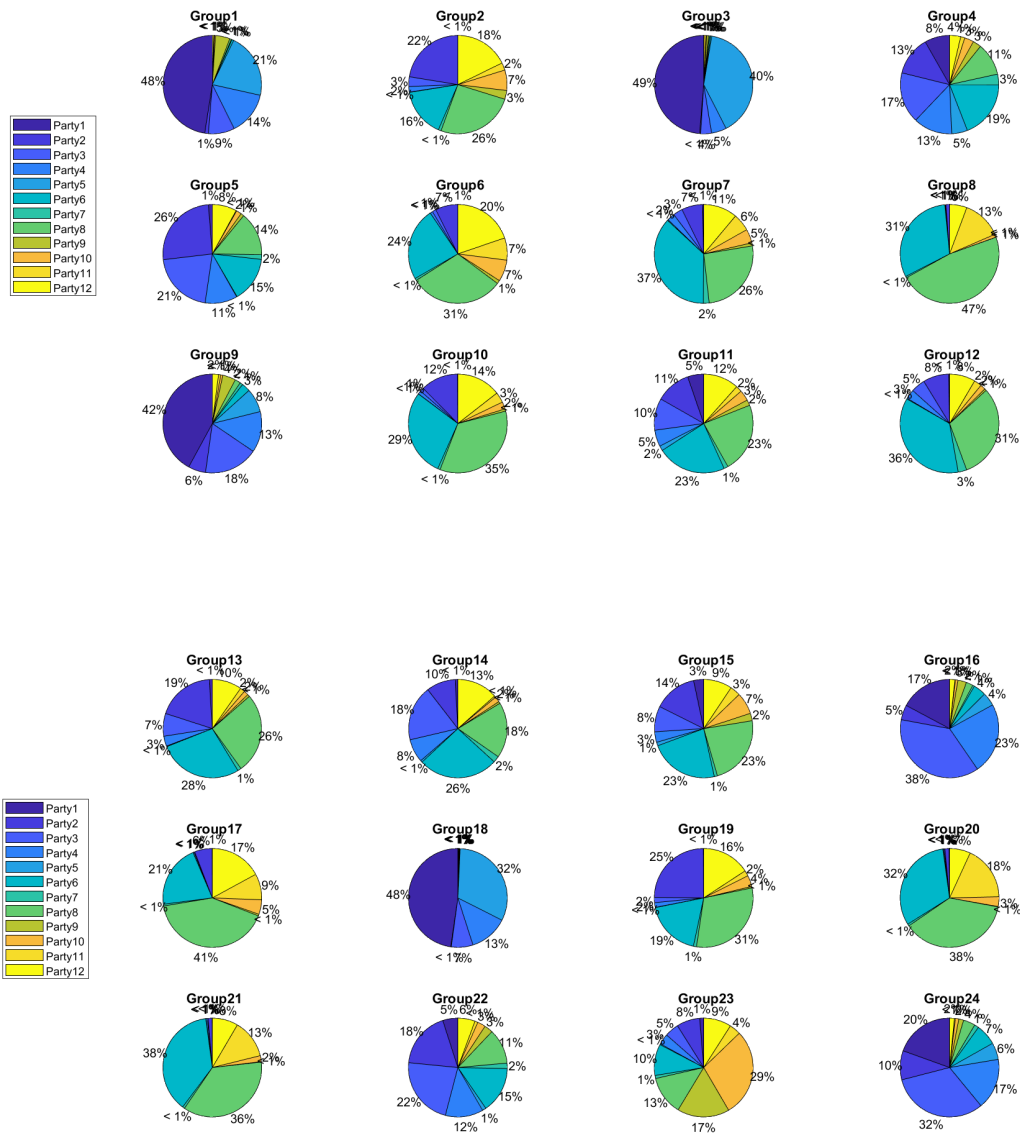


Σχήμα VI.12: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 12 ομάδες

ΠΑΡΑΡΤΗΜΑ VI. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ - ΜΕΡΟΣ Β

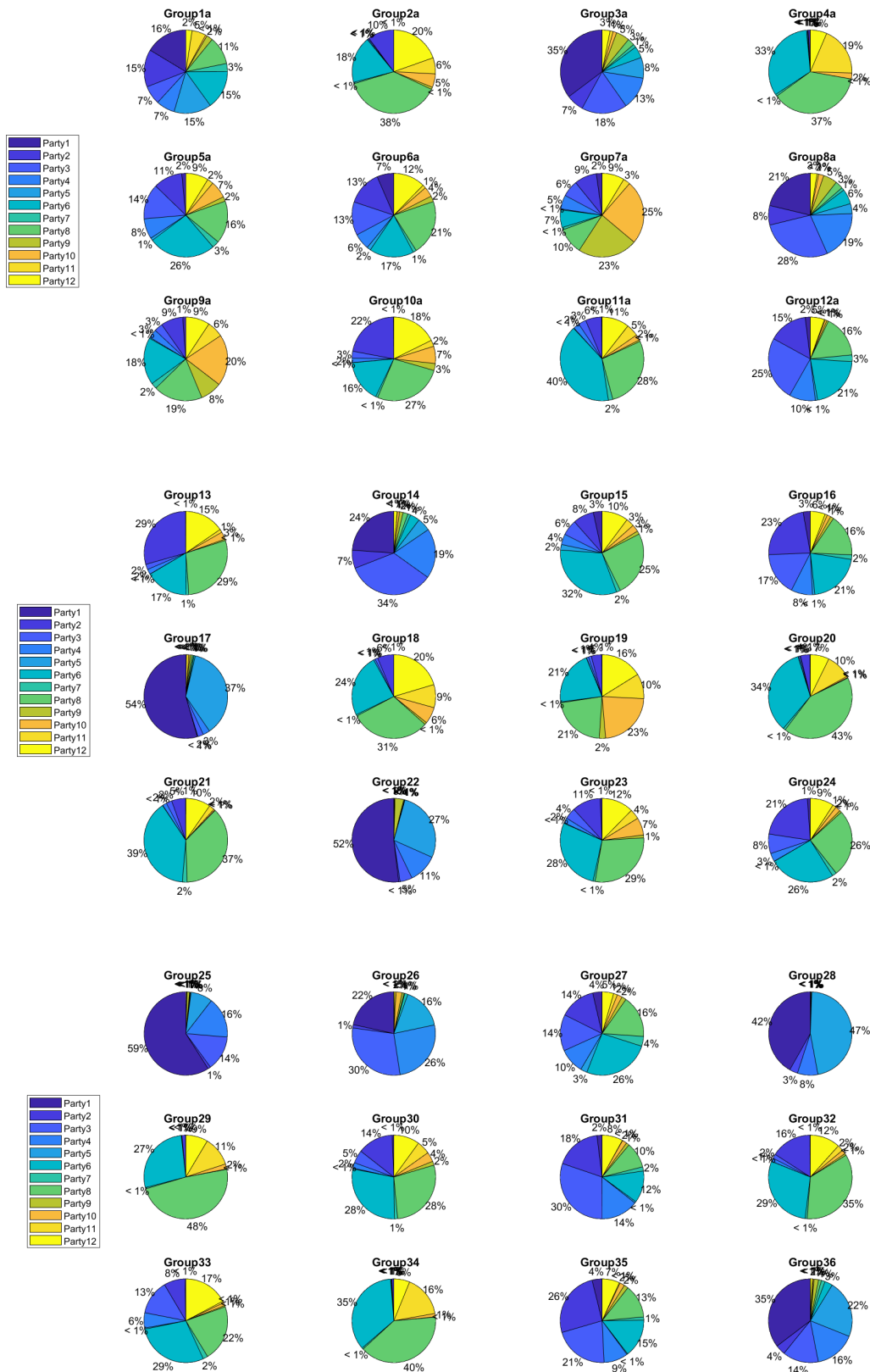


Σχήμα VI.13: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 18 ομάδες

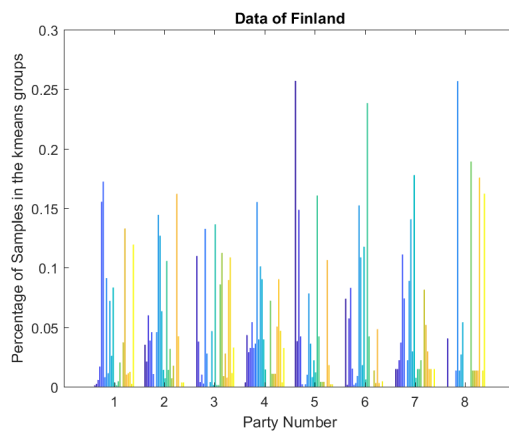
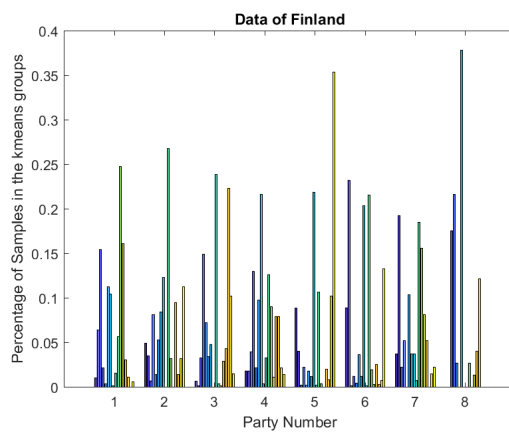
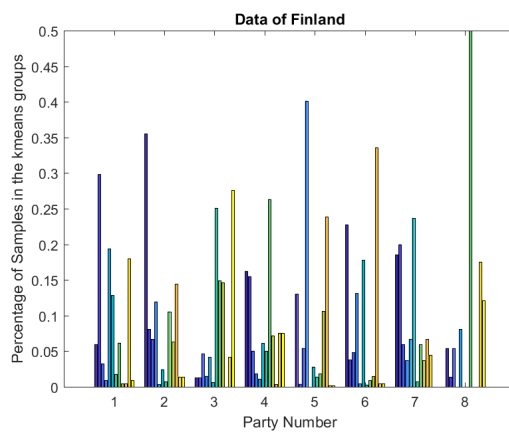
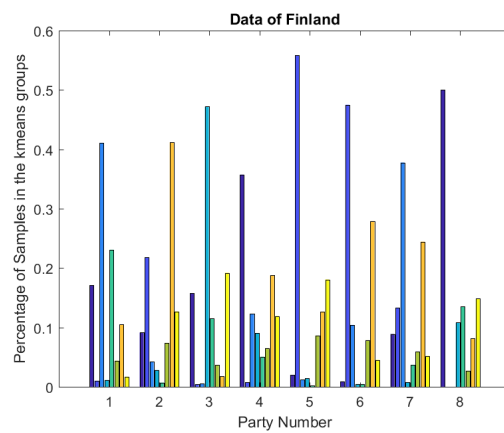


Σχήμα VI.14: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 24 ομάδες

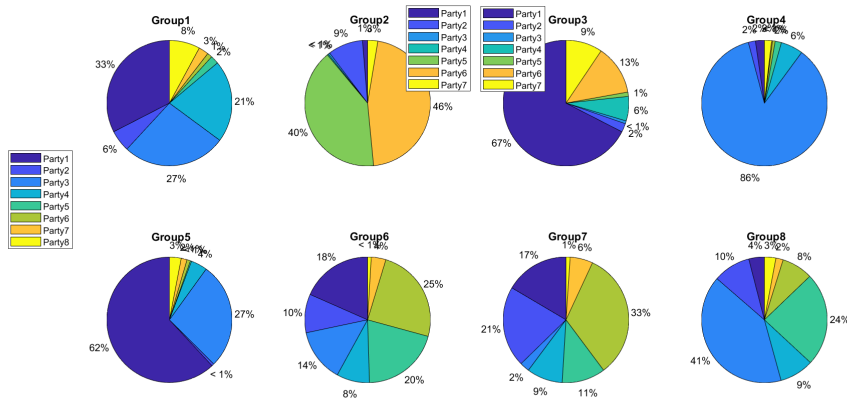
ΠΑΡΑΡΤΗΜΑ VI. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ - ΜΕΡΟΣ Β



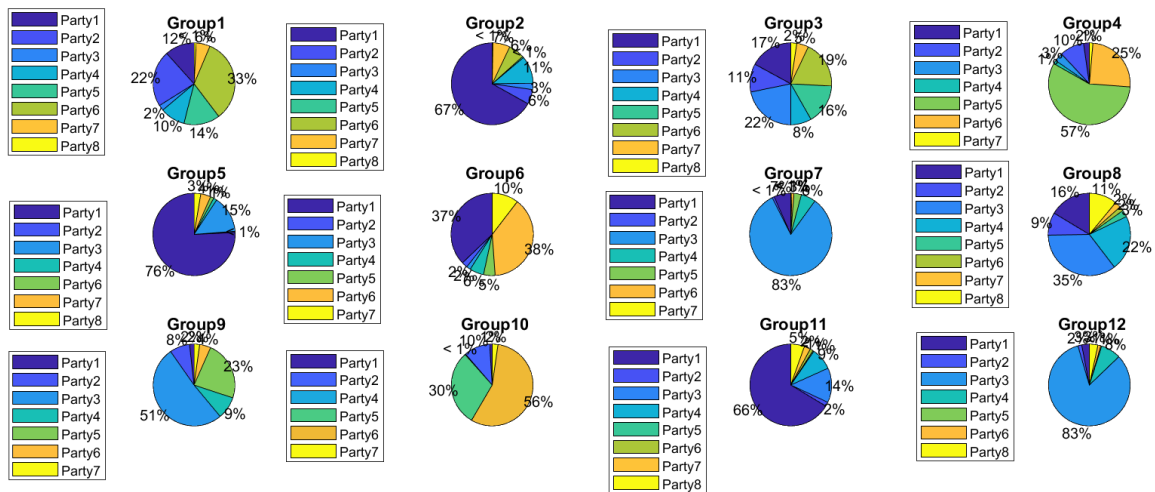
Σχήμα VI.15: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ισπανία – 36 ομάδες



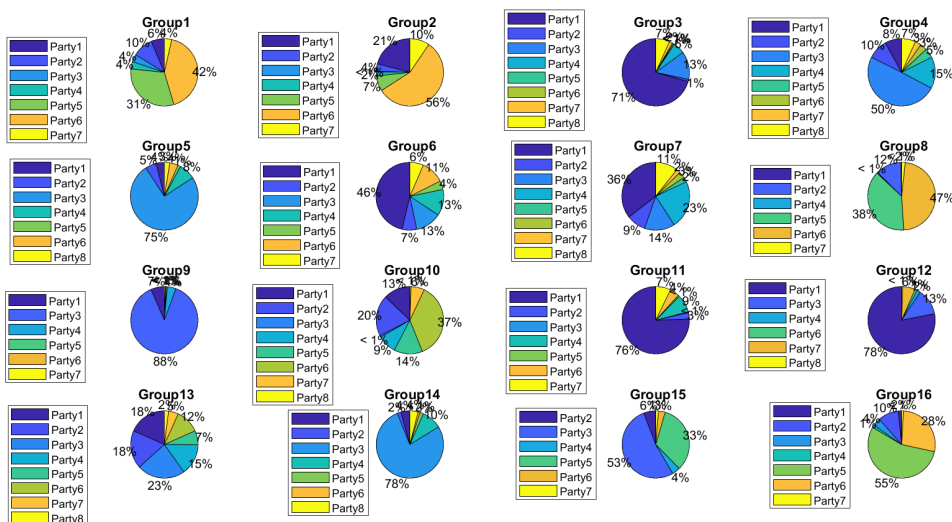
Σχήμα VI.16: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Φινλανδία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες



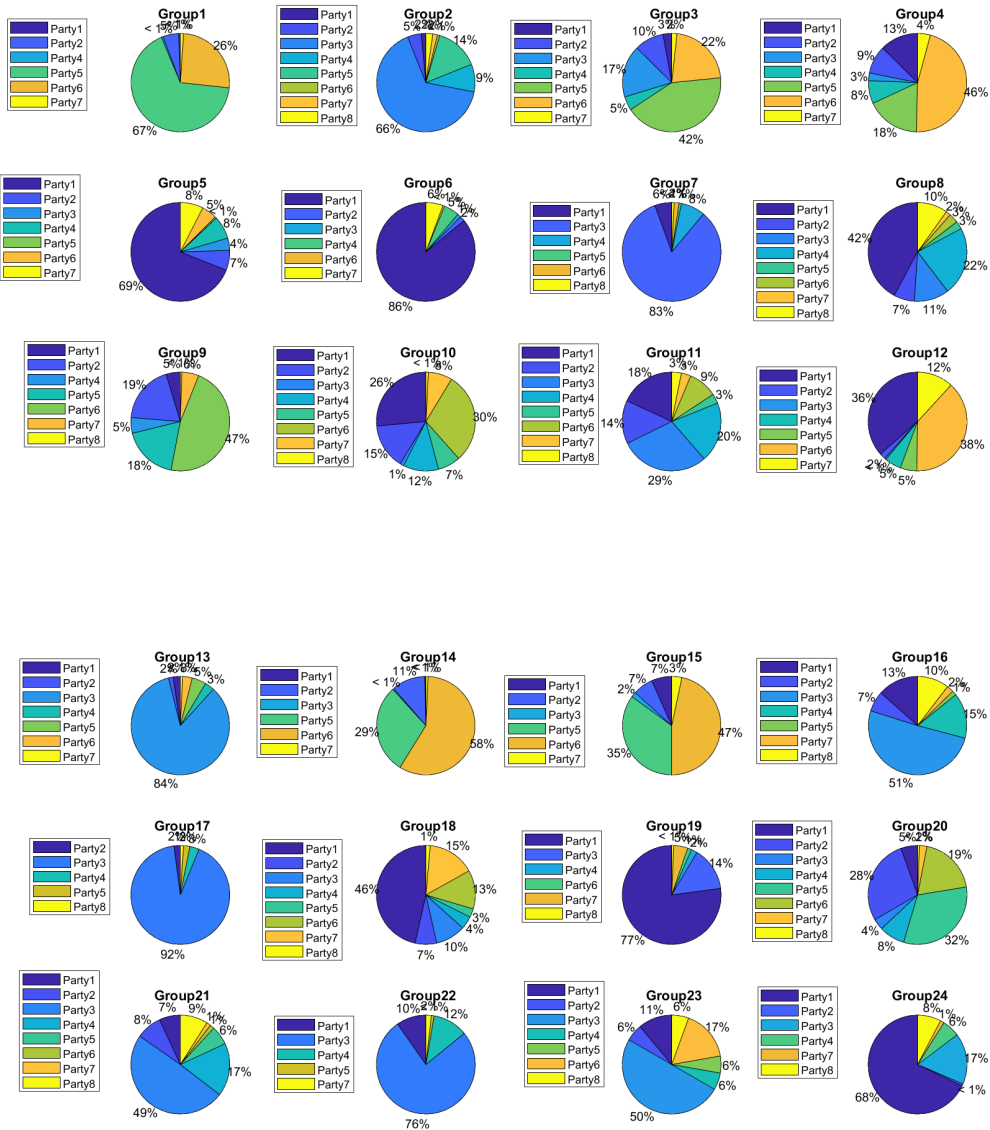
Σχήμα VI.17: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 8 ομάδες



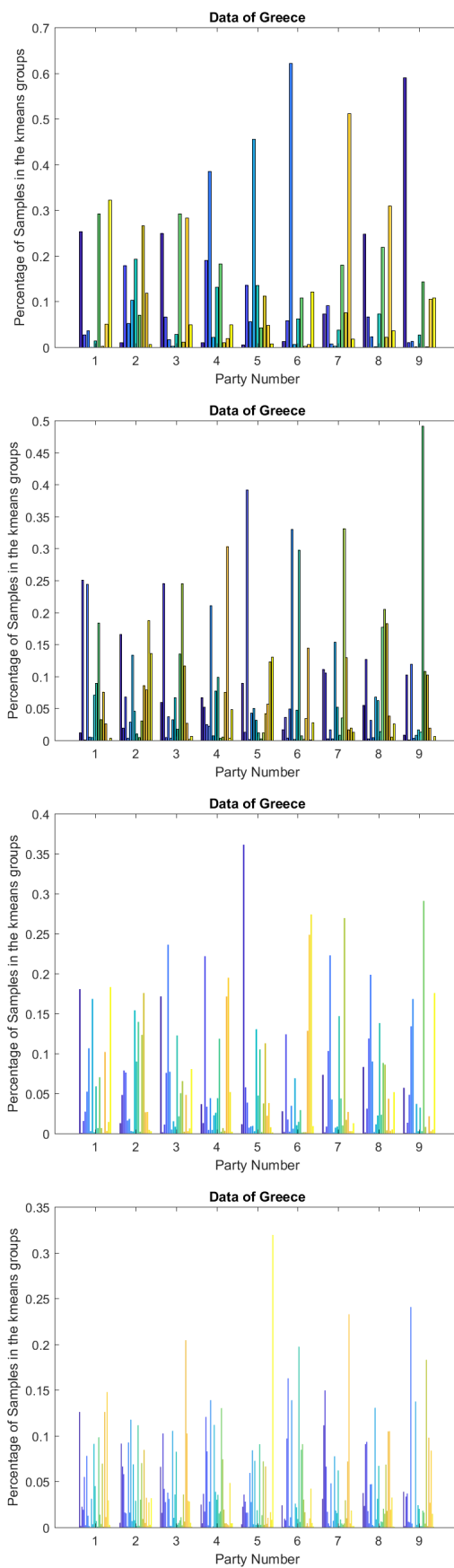
Σχήμα VI.18: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 12 ομάδες



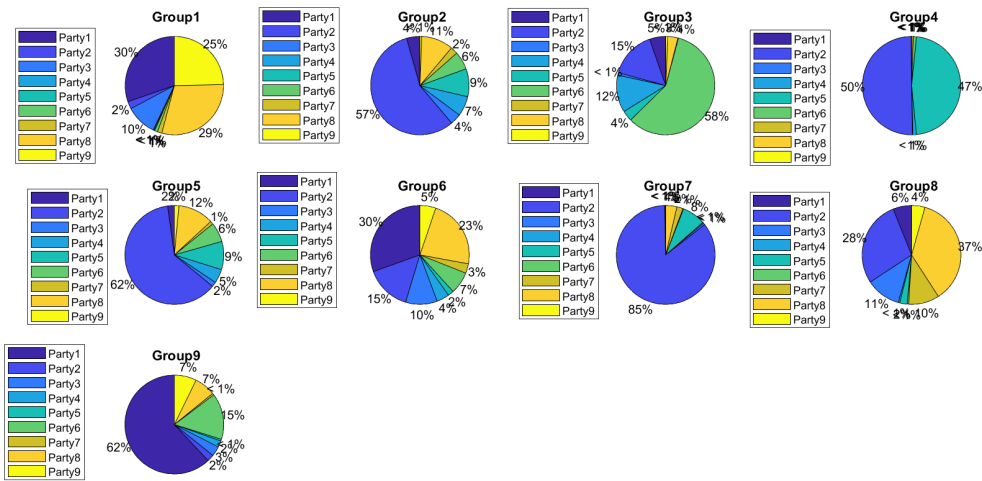
Σχήμα VI.19: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 16 ομάδες



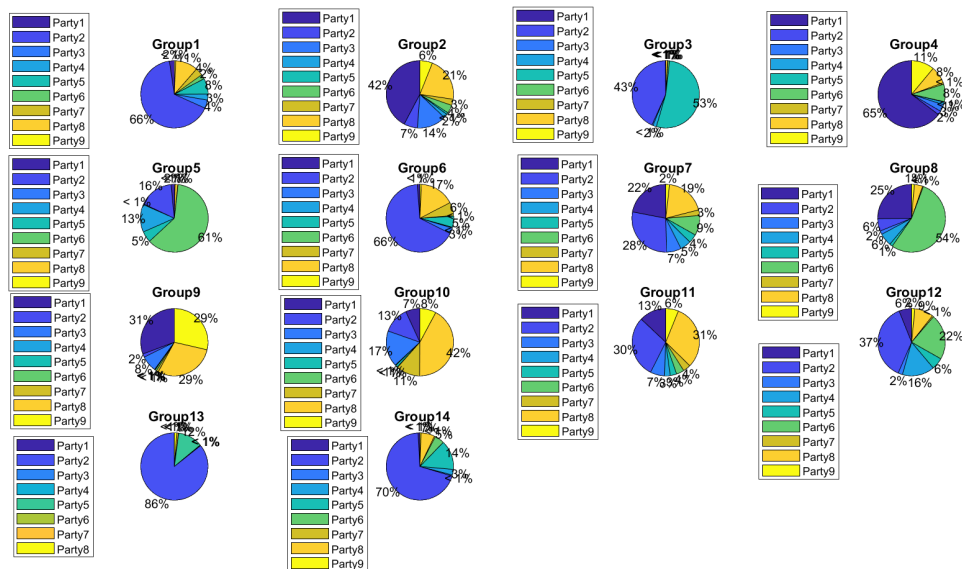
Σχήμα VI.20: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Φινλανδία – 24 ομάδες



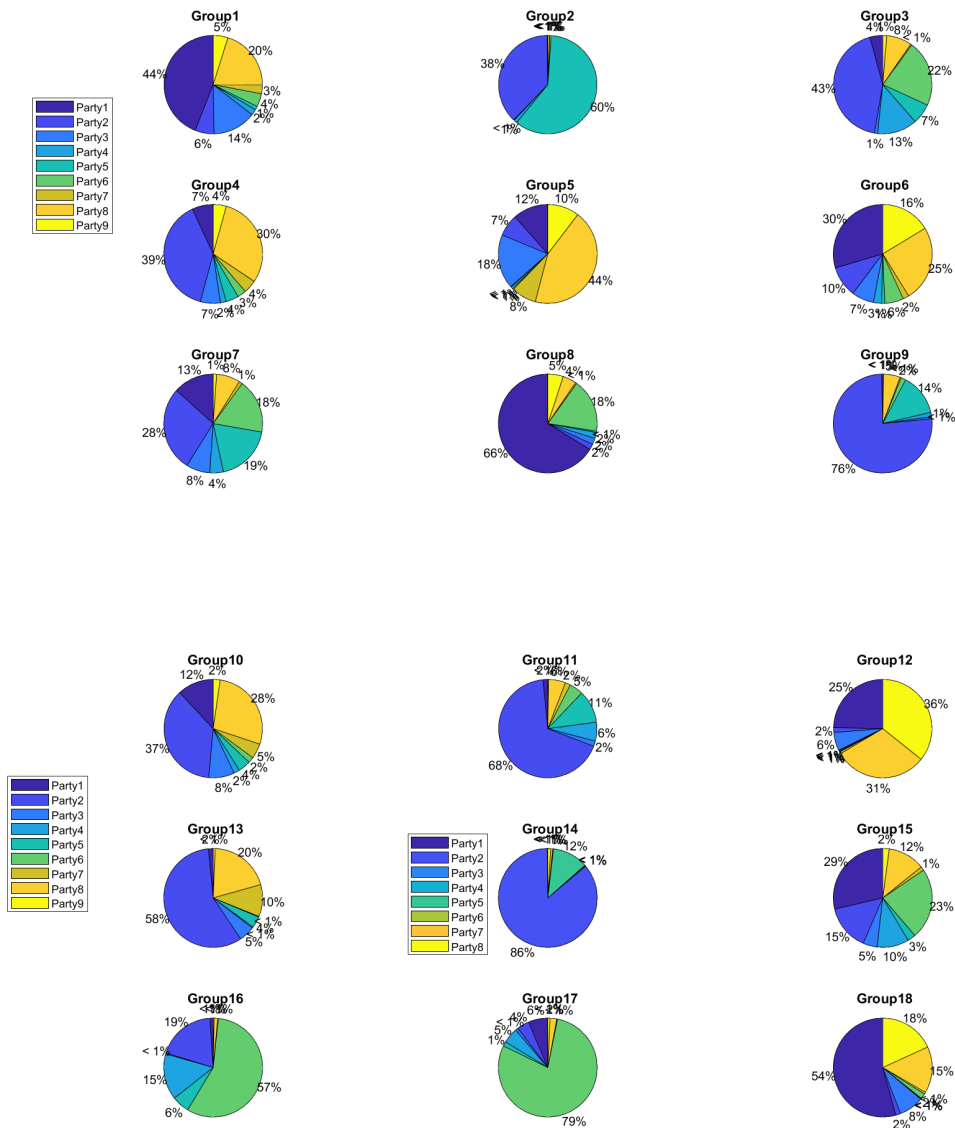
Σχήμα VI.21: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ελλάδα (α) 9 ομάδες (β) 14 ομάδες (γ) 18 ομάδες (δ) 27 ομάδες



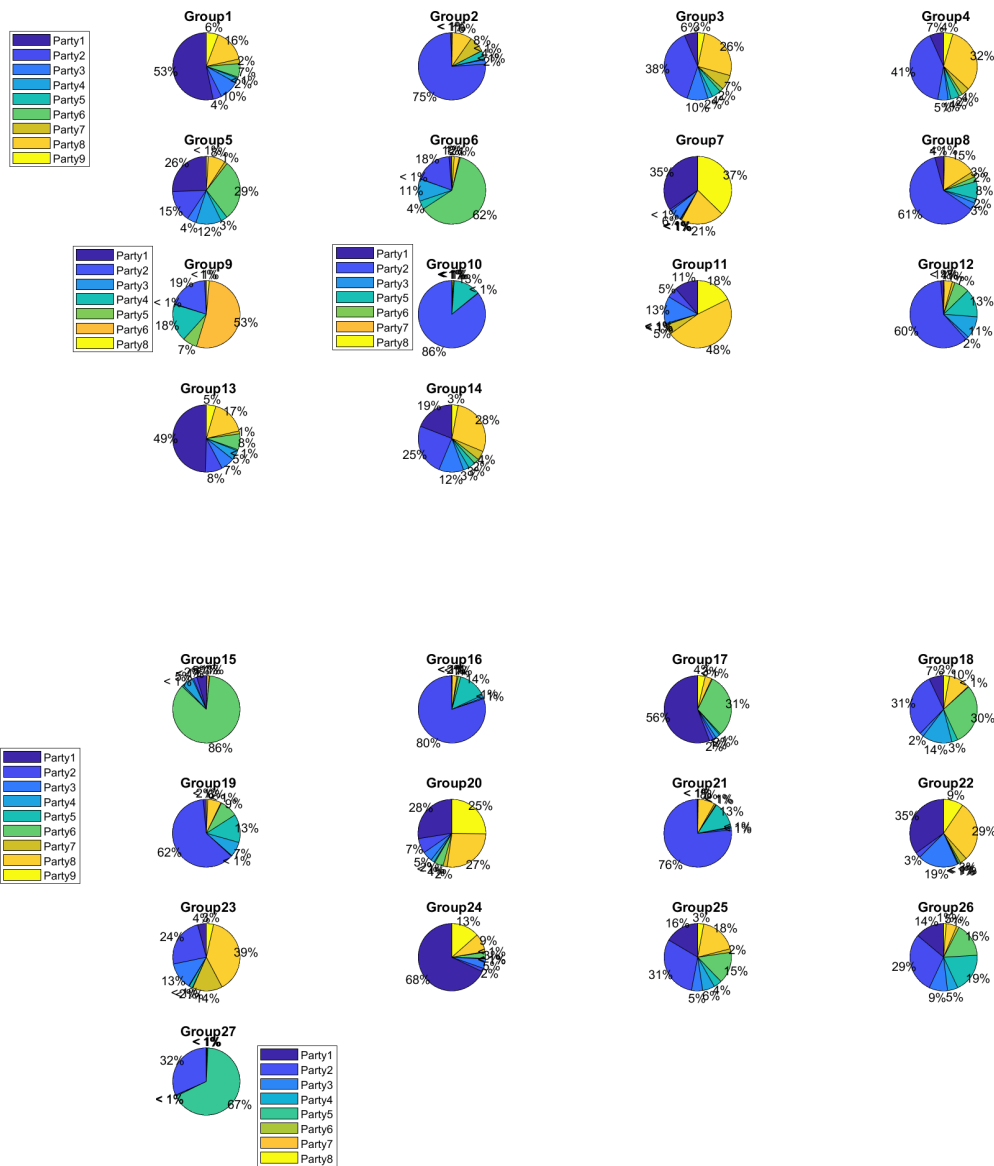
Σχήμα VI.22: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 9 ομάδες



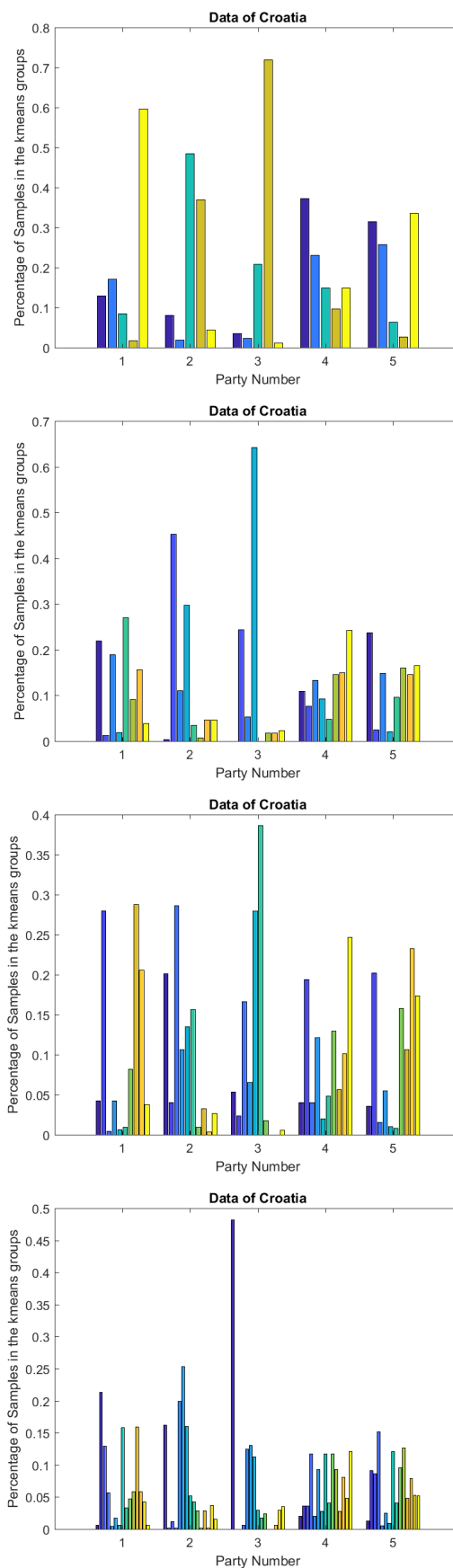
Σχήμα VI.23: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 14 ομάδες



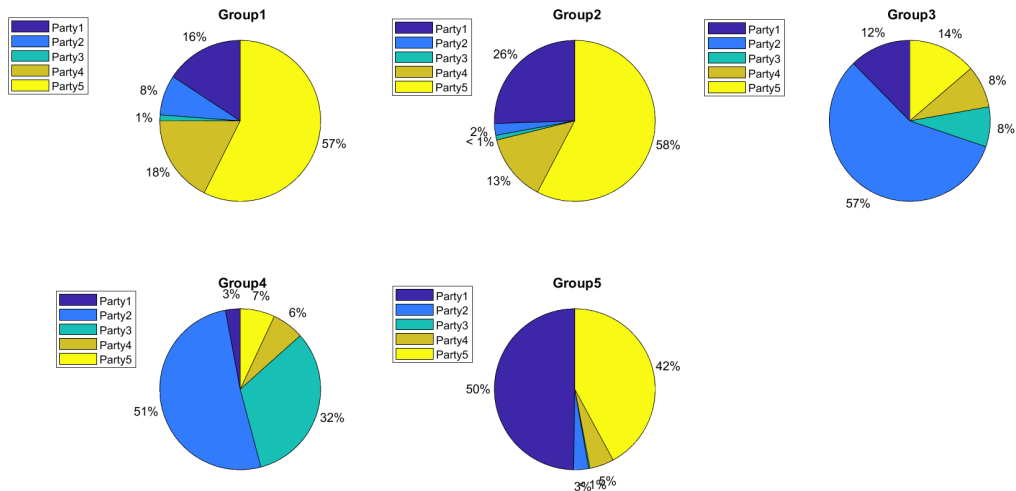
Σχήμα VI.24: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 18 ομάδες



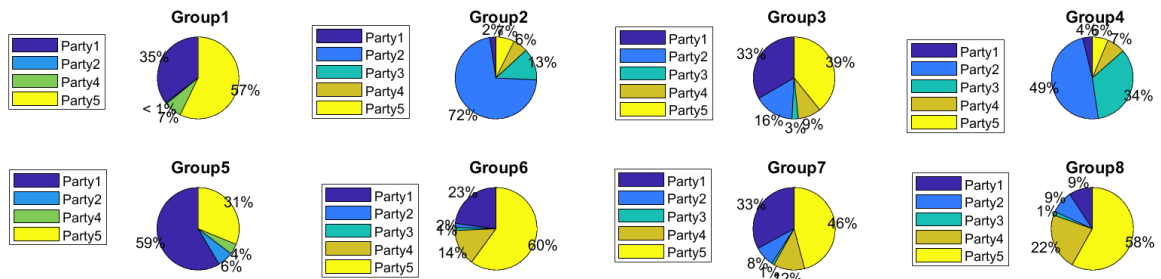
Σχήμα VI.25: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ελλάδα – 27 ομάδες



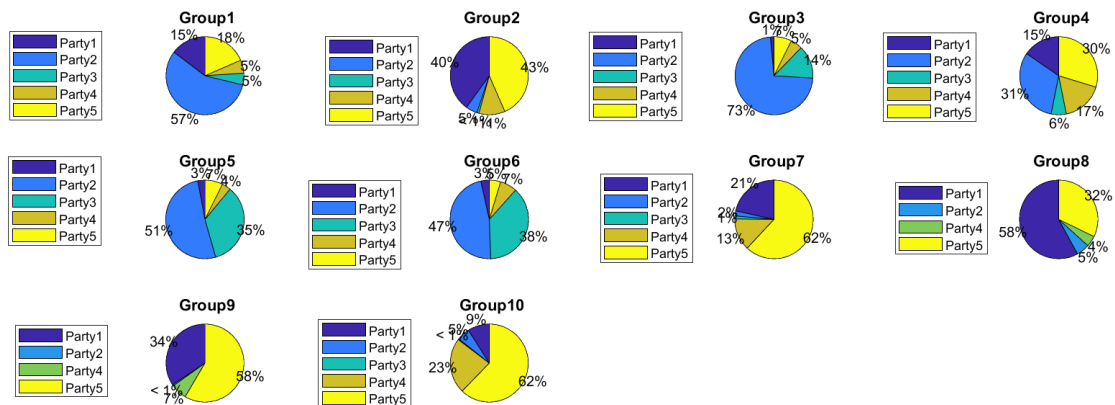
Σχήμα VI.26: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Κροατία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες



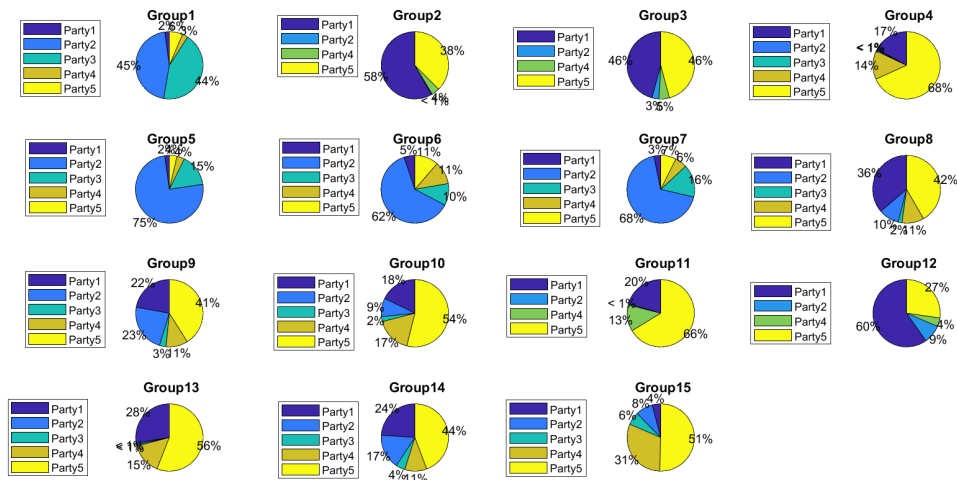
Σχήμα VI.27: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 5 ομάδες



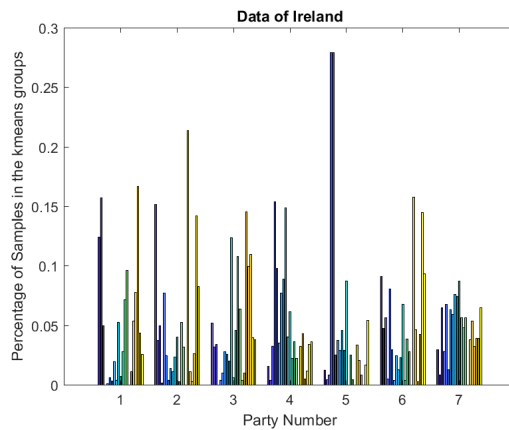
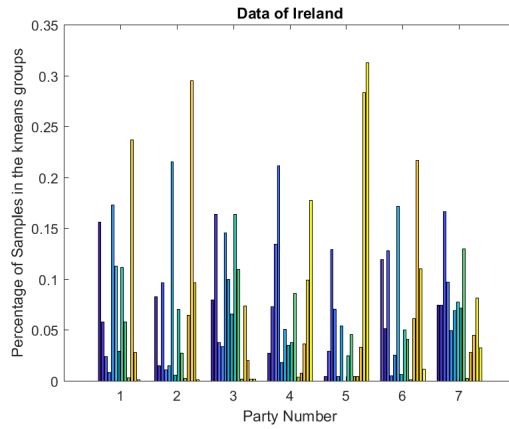
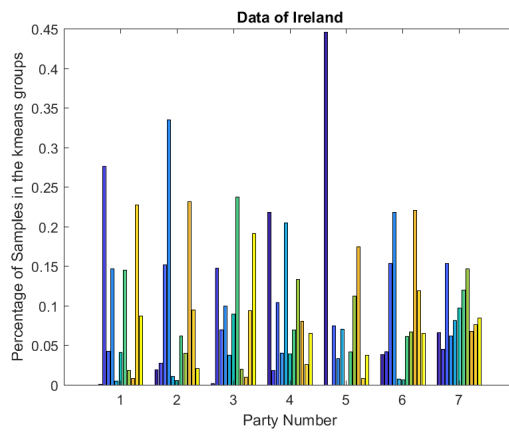
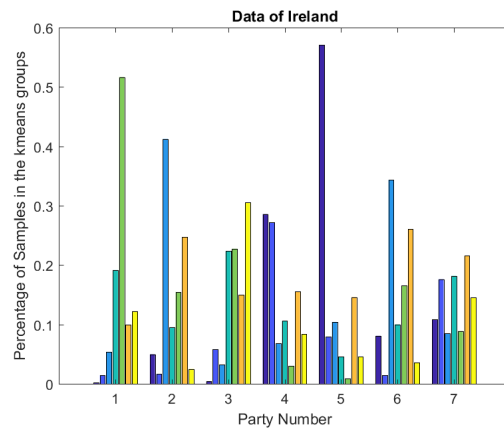
Σχήμα VI.28: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 8 ομάδες



Σχήμα VI.29: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 10 ομάδες

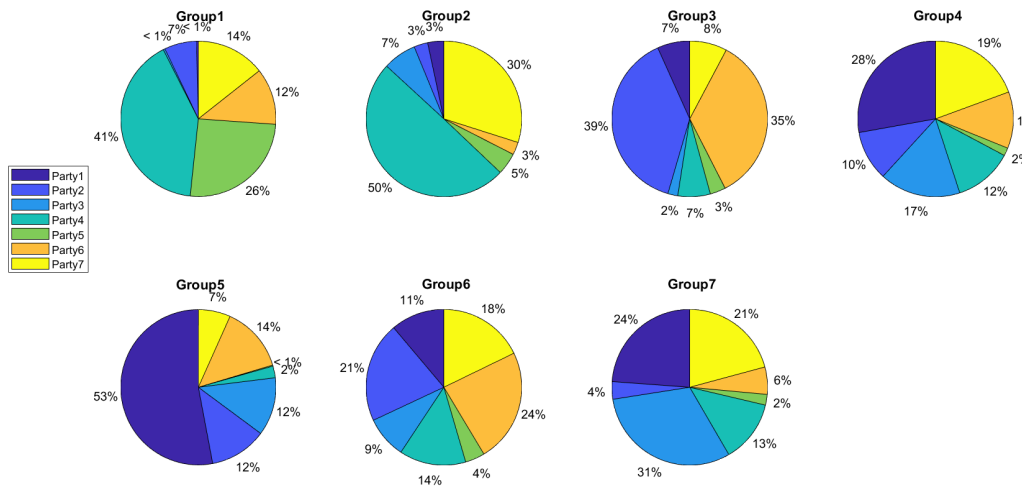


Σχήμα VI.30: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Κροατία – 15 ομάδες

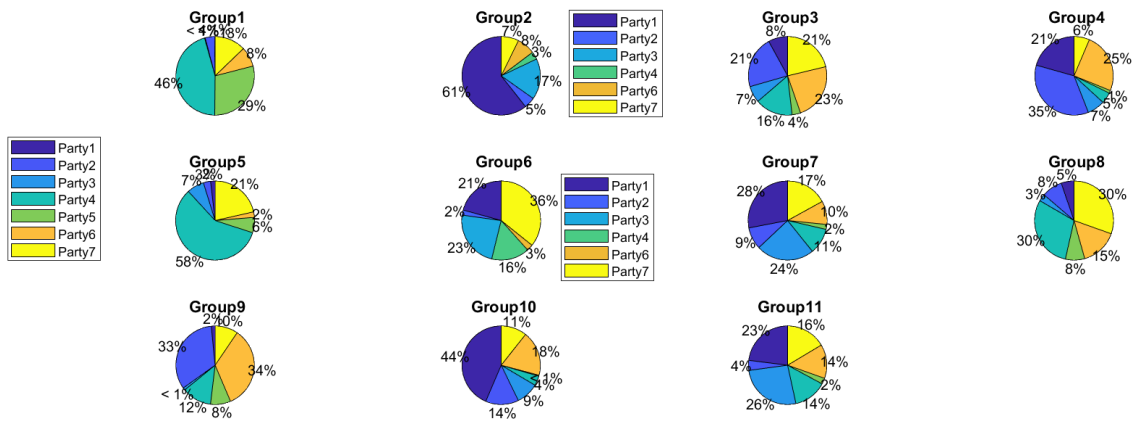


Σχήμα VI.31: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ιρλανδία (α) 7 ομάδες (β) 11 ομάδες (γ) 14 ομάδες (δ) 21 ομάδες

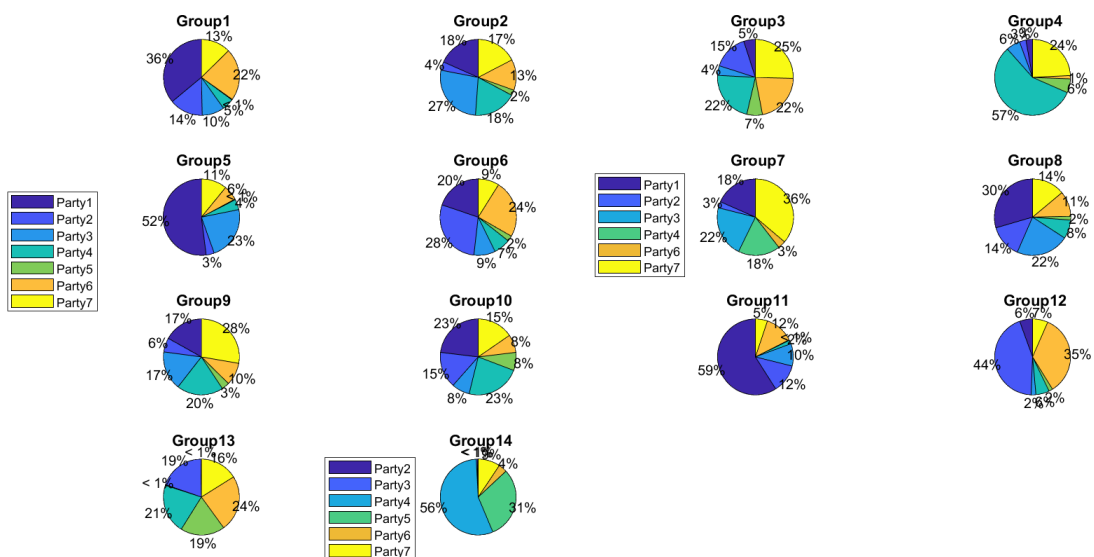
ΠΑΡΑΡΤΗΜΑ VI. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ - ΜΕΡΟΣ Β



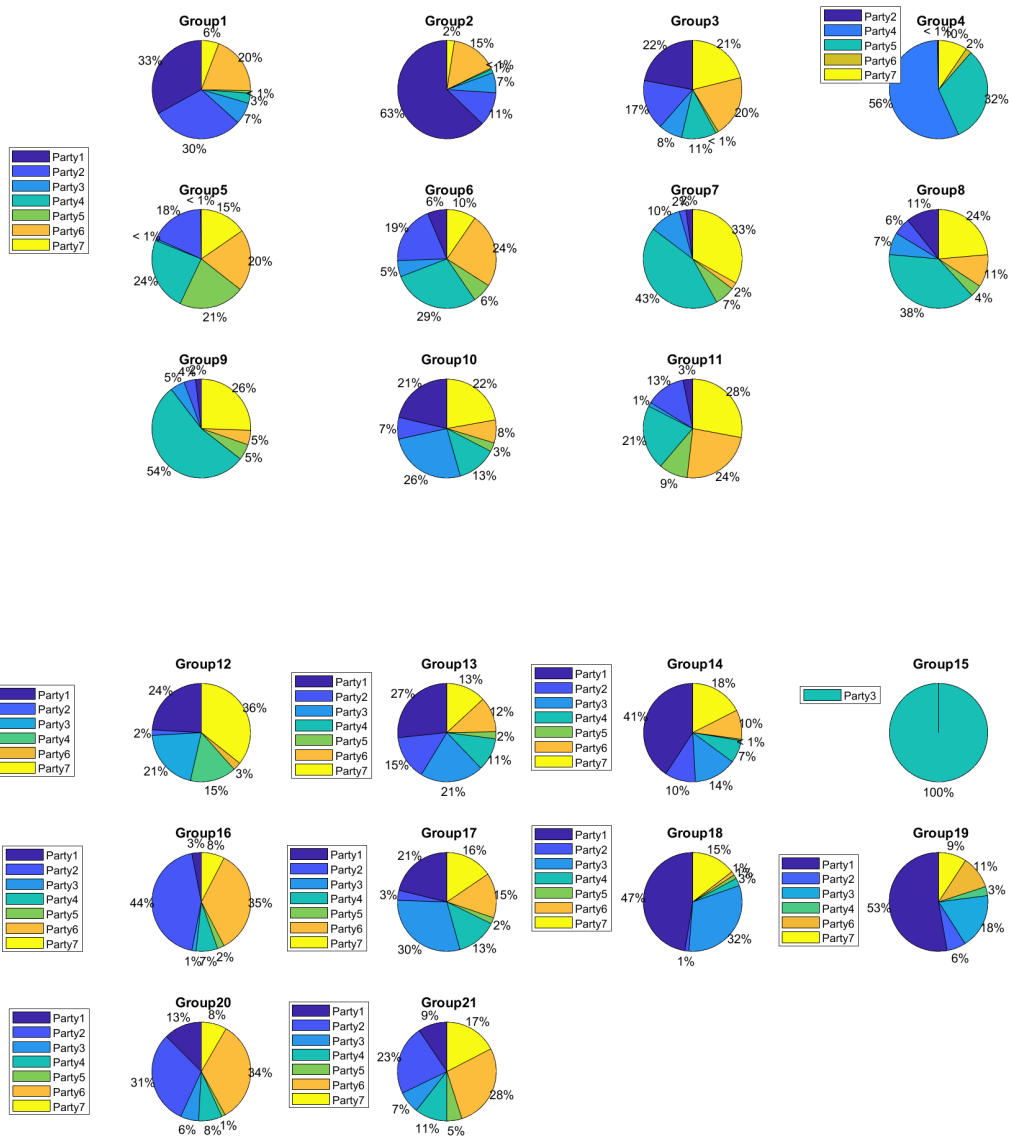
Σχήμα VI.32: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 7 ομάδες



Σχήμα VI.33: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 11 ομάδες



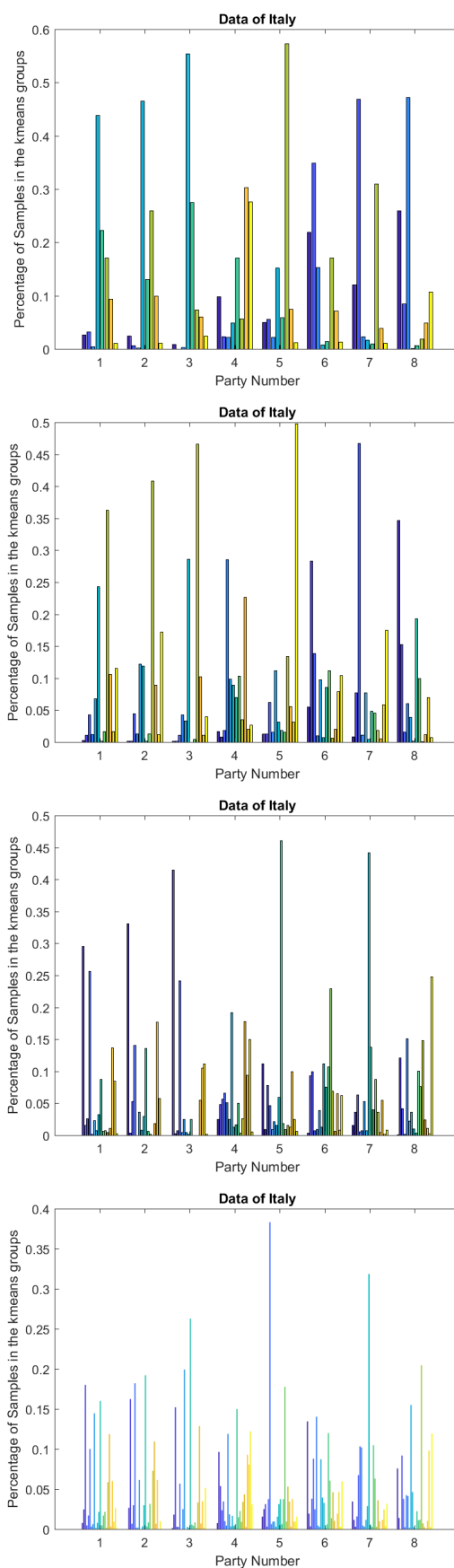
Σχήμα VI.34: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 14 ομάδες



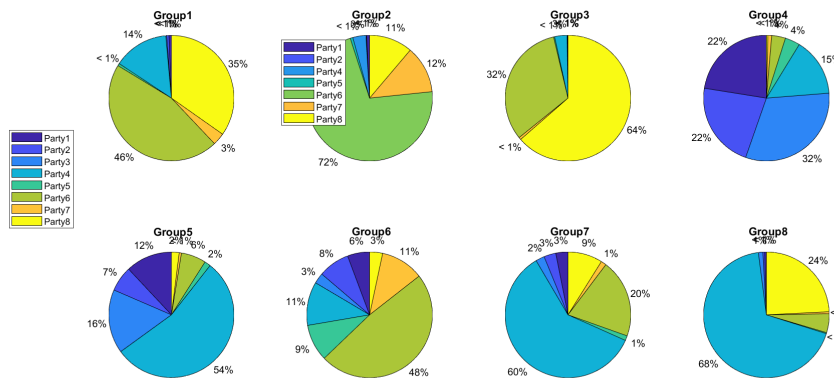
Σχήμα VI.35: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιρλανδία – 21 ομάδες

Παράρτημα VII

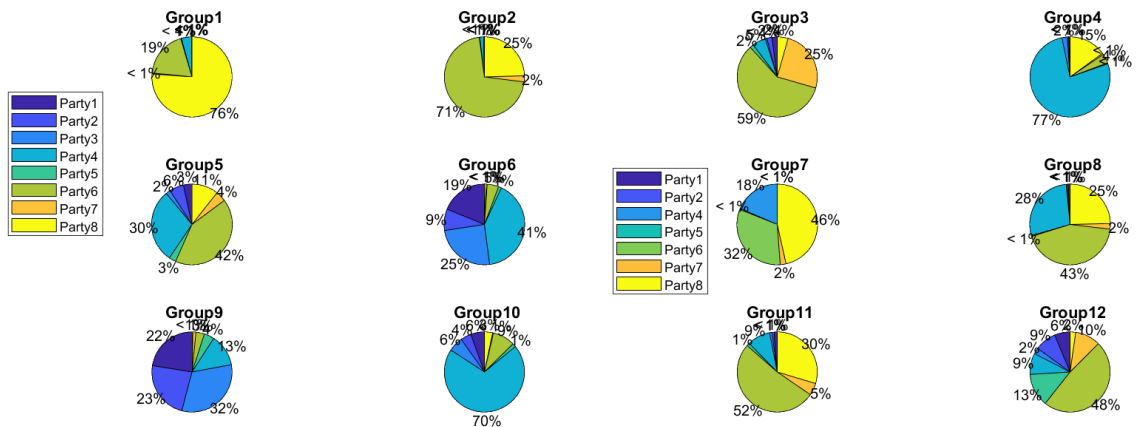
Αποτελέσματα από τη διερεύνηση της ιδεολογικής ταύτισης κόμματος-χρήστη - Μέρος Γ



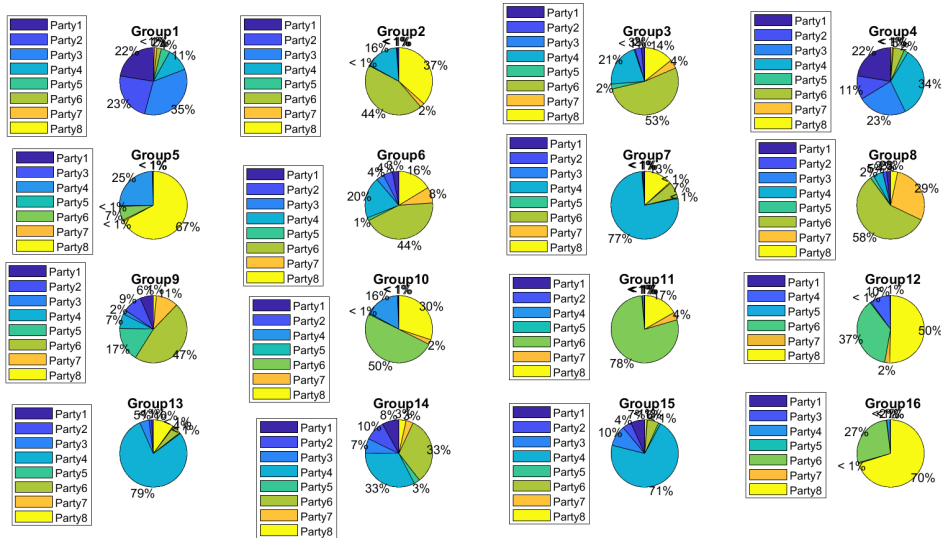
Σχήμα VII.1: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ιταλία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες



Σχήμα VII.2: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 8 ομάδες

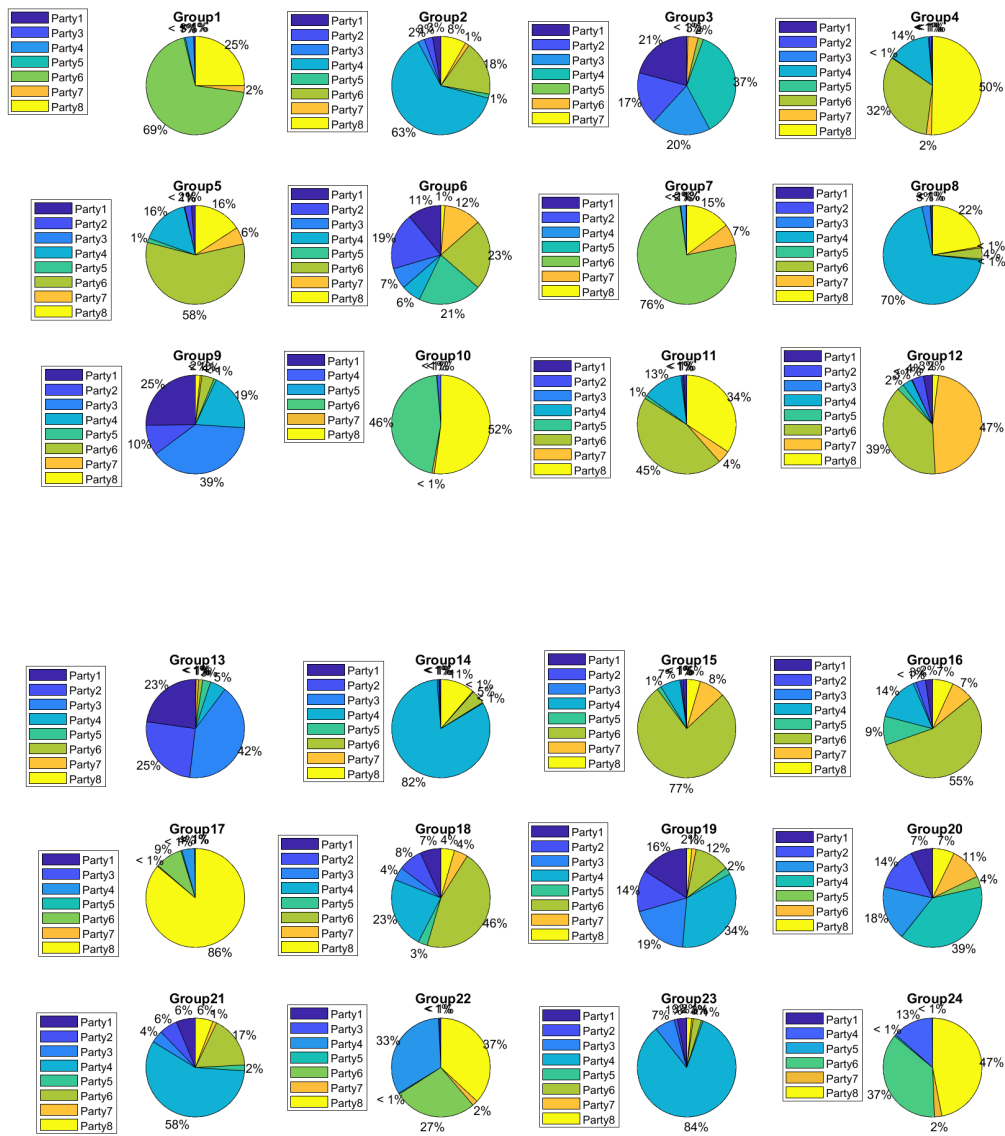


Σχήμα VII.3: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 12 ομάδες

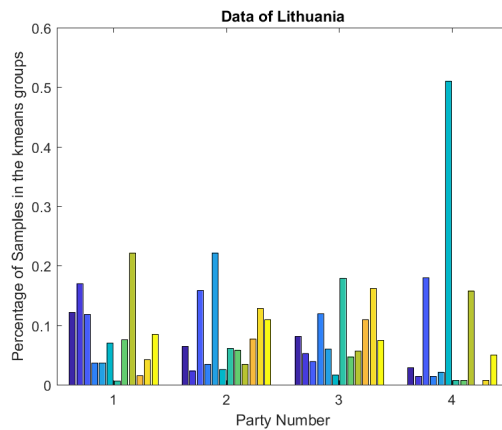
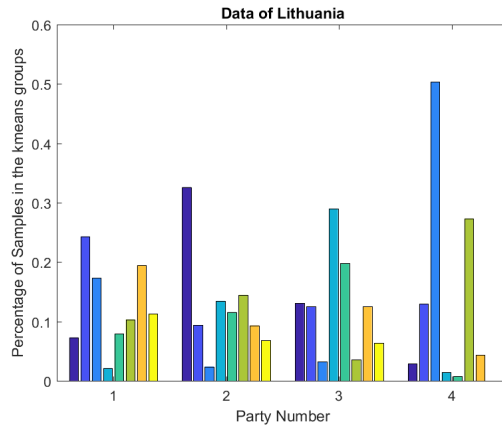
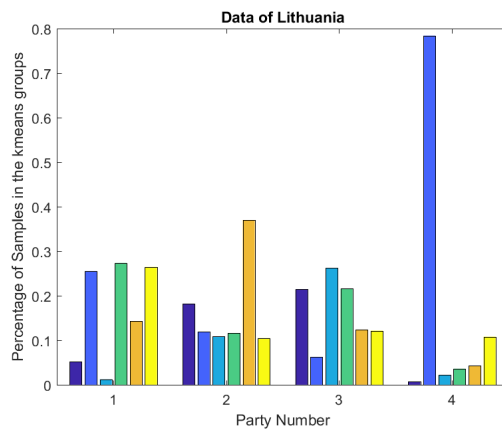
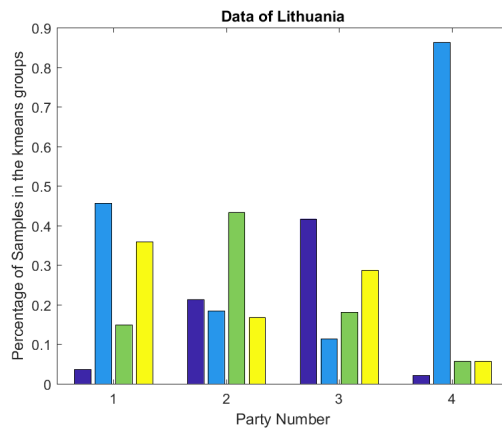


Σχήμα VII.4: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 16 ομάδες

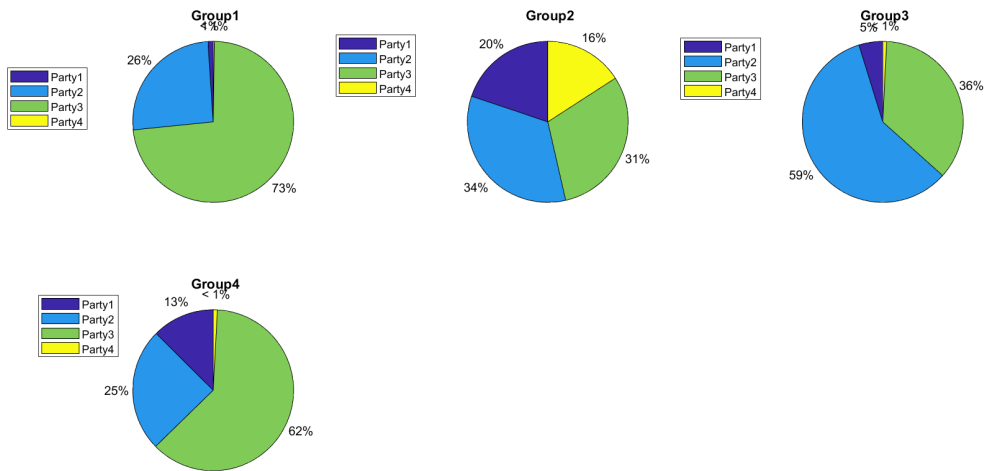
ΠΑΡΑΡΤΗΜΑ VII. ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΠΟ ΤΗ ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΙΔΕΟΛΟΓΙΚΗΣ ΤΑΥΤΙΣΗΣ ΚΟΜΜΑΤΟΣ-ΧΡΗΣΤΗ - ΜΕΡΟΣ Γ



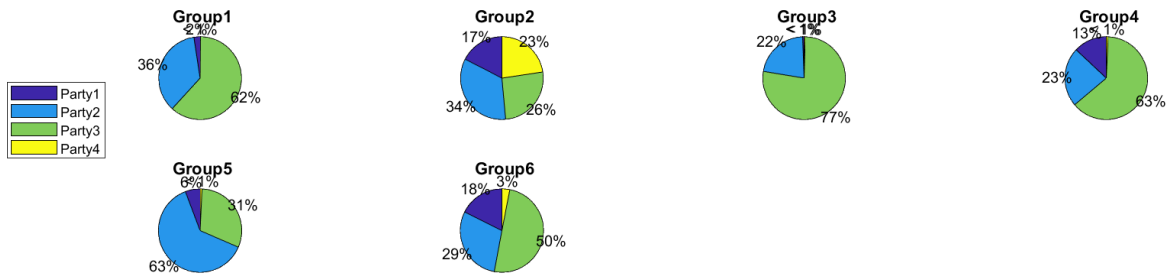
Σχήμα VII.5: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ιταλία – 24 ομάδες



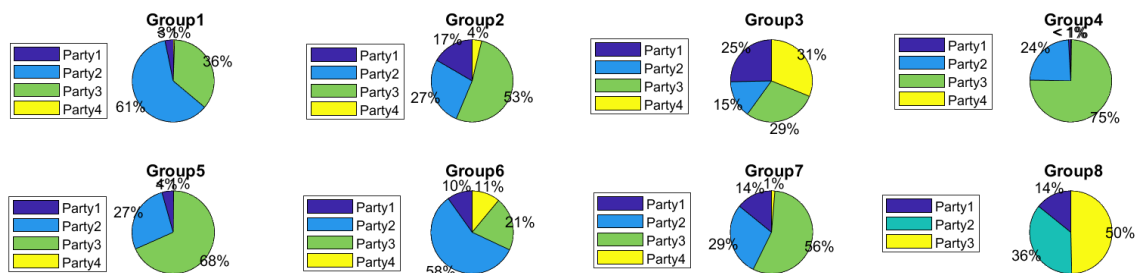
Σχήμα VII.6: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Λιθουανία (α) 4 ομάδες (β) 6 ομάδες (γ) 8 ομάδες (δ) 12 ομάδες



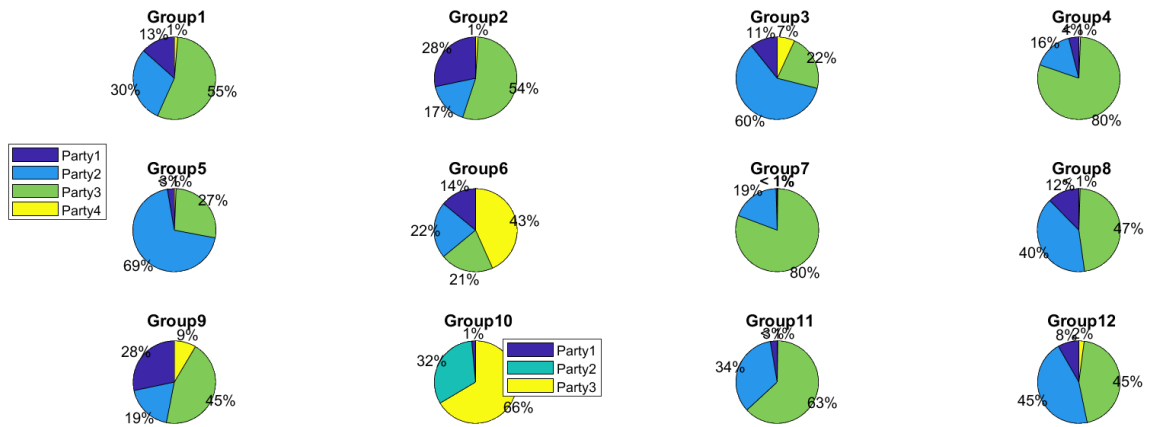
Σχήμα VII.7: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 4 ομάδες



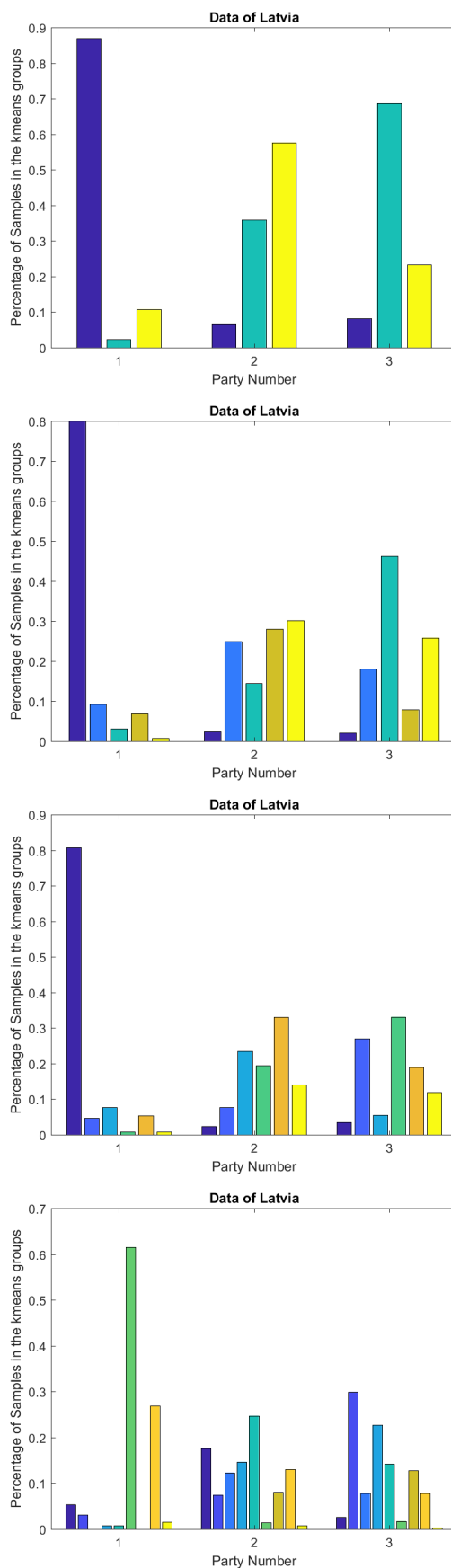
Σχήμα VII.8: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 6 ομάδες



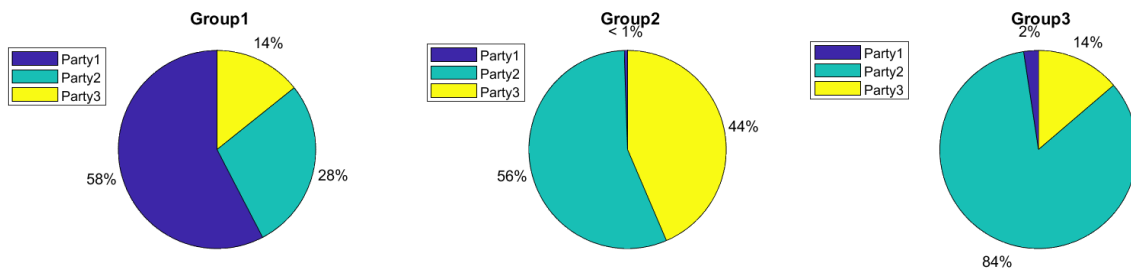
Σχήμα VII.9: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 8 ομάδες



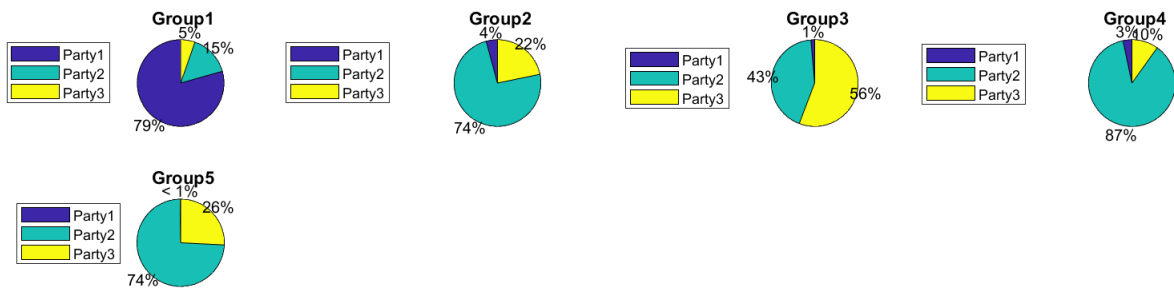
Σχήμα VII.10: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λιθουανία – 12 ομάδες



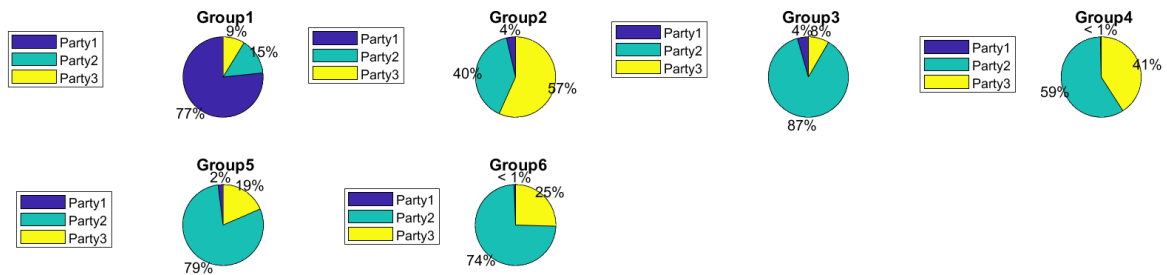
Σχήμα VII.11: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Λετονία (α) 3 ομάδες (β) 5 ομάδες (γ) 6 ομάδες (δ) 9 ομάδες



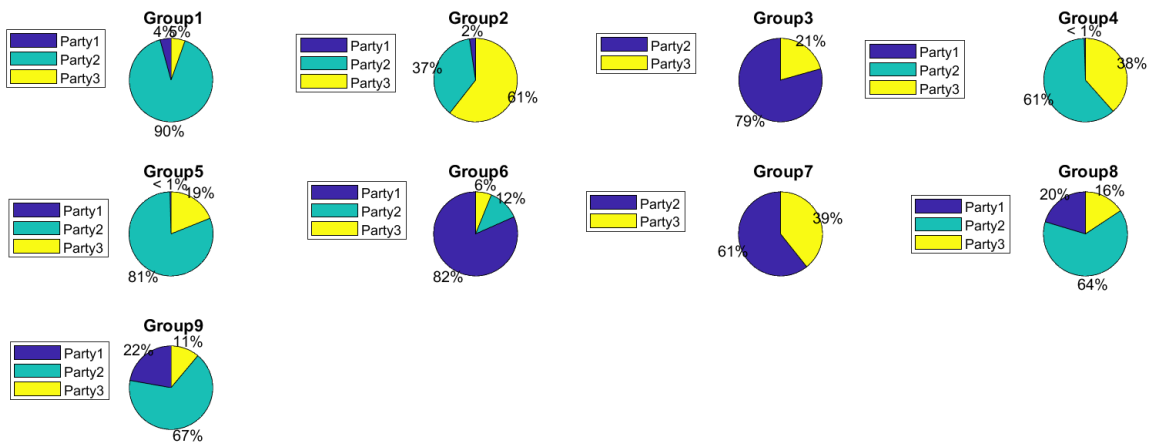
Σχήμα VII.12: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 3 ομάδες



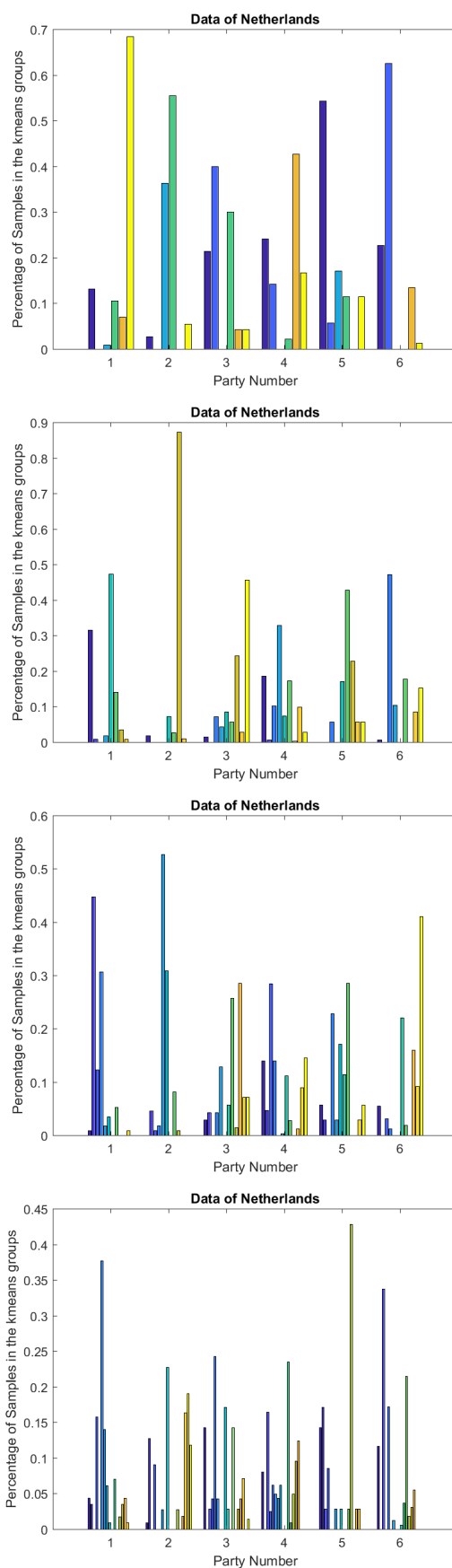
Σχήμα VII.13: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 5 ομάδες



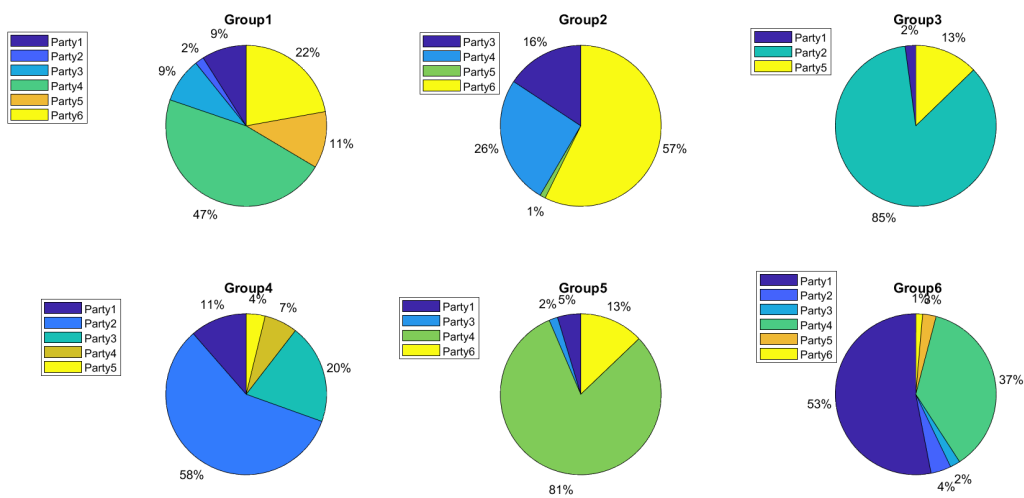
Σχήμα VII.14: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 6 ομάδες



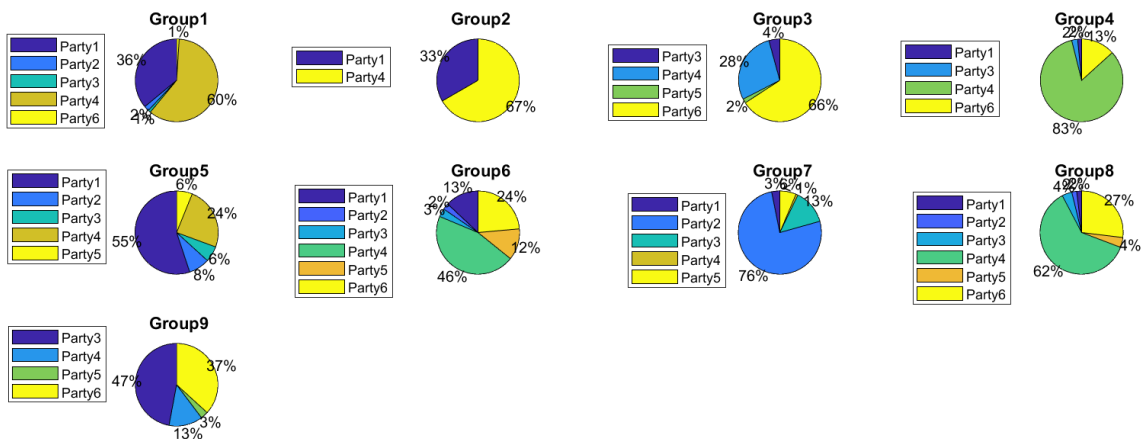
Σχήμα VII.15: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Λετονία – 9 ομάδες



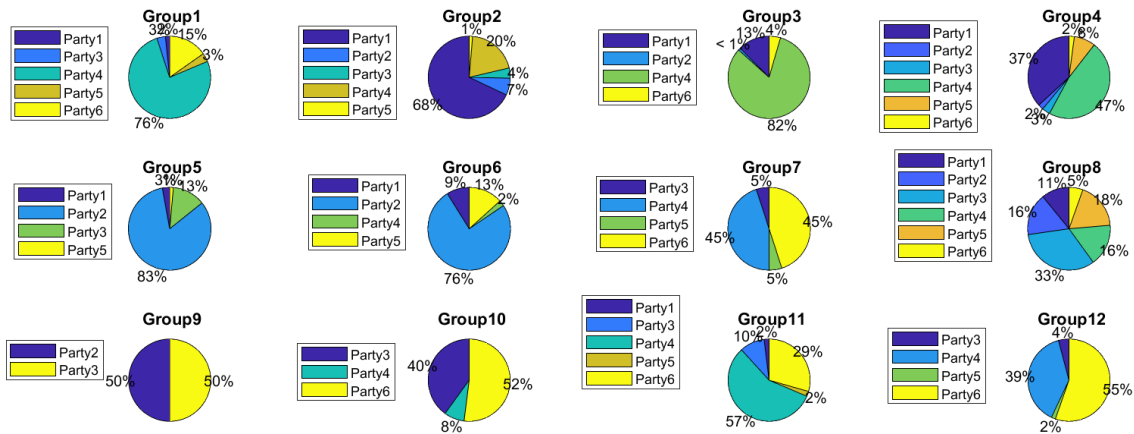
Σχήμα VII.16: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ολλανδία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες



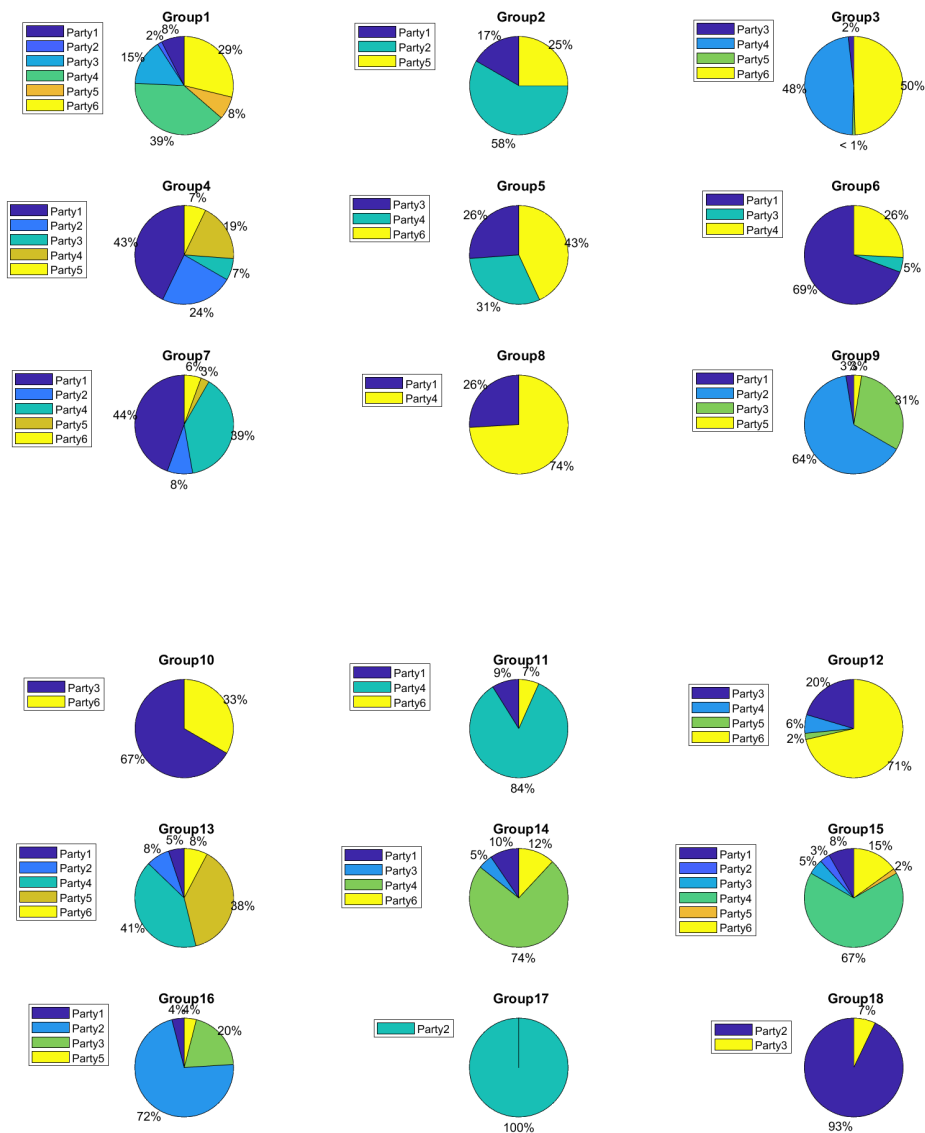
Σχήμα VII.17: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 6 ομάδες



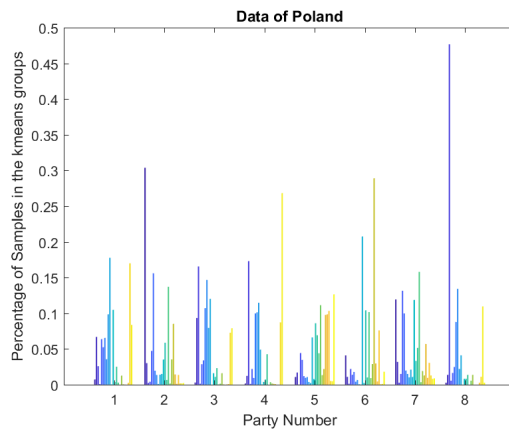
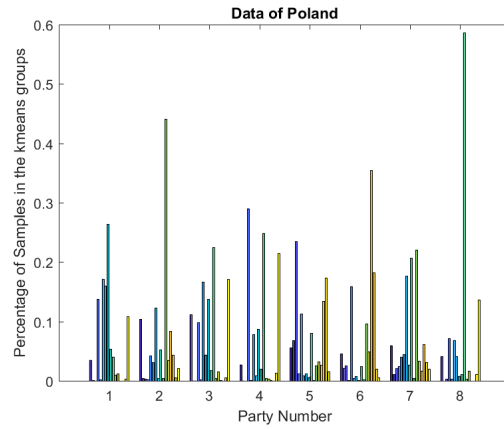
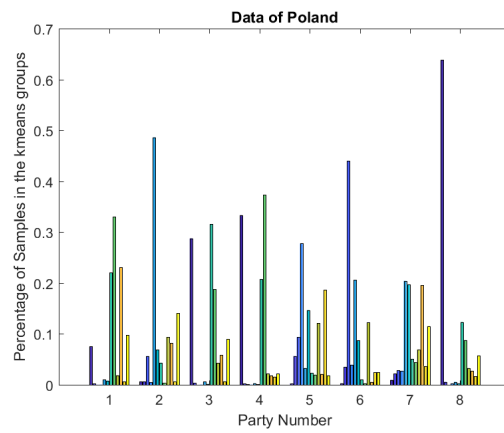
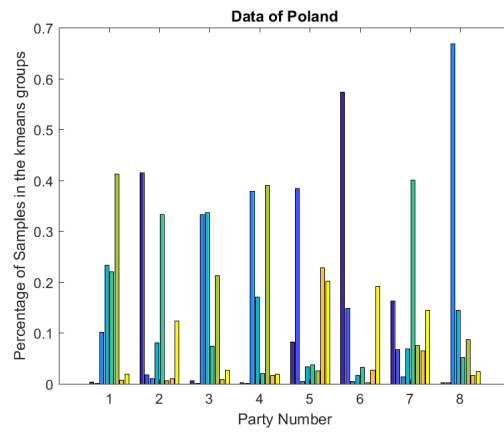
Σχήμα VII.18: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 9 ομάδες



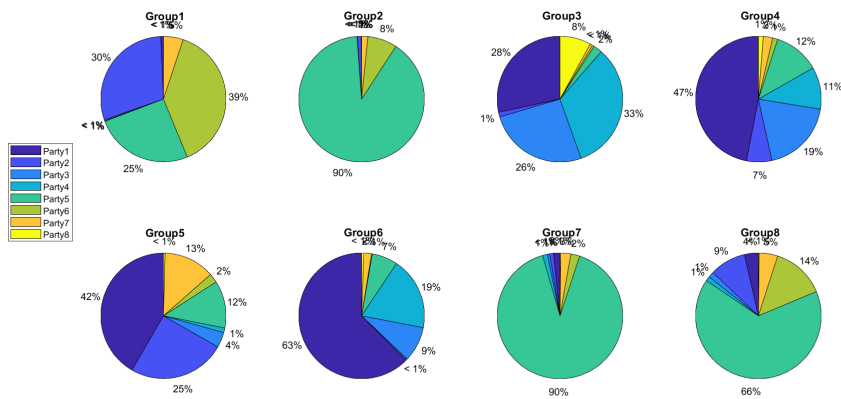
Σχήμα VII.19: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 12 ομάδες



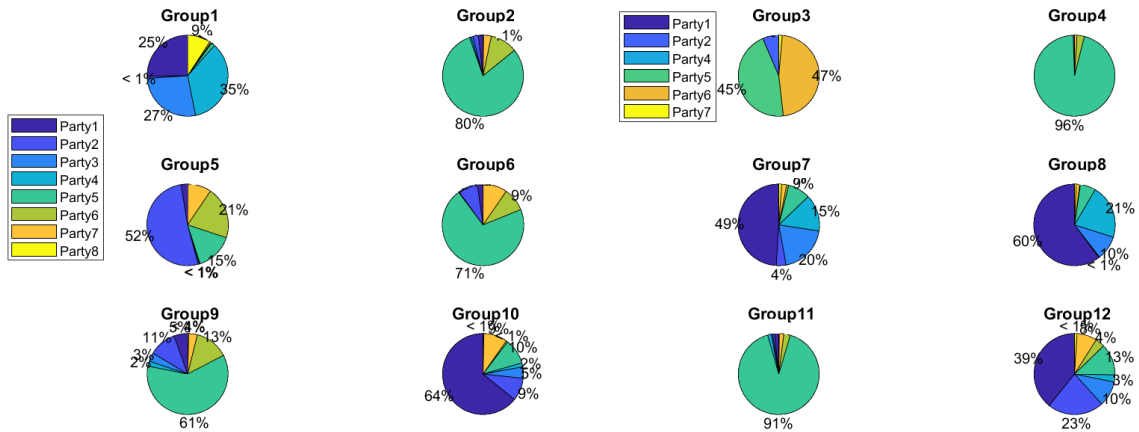
Σχήμα VII.20: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ολλανδία – 18 ομάδες



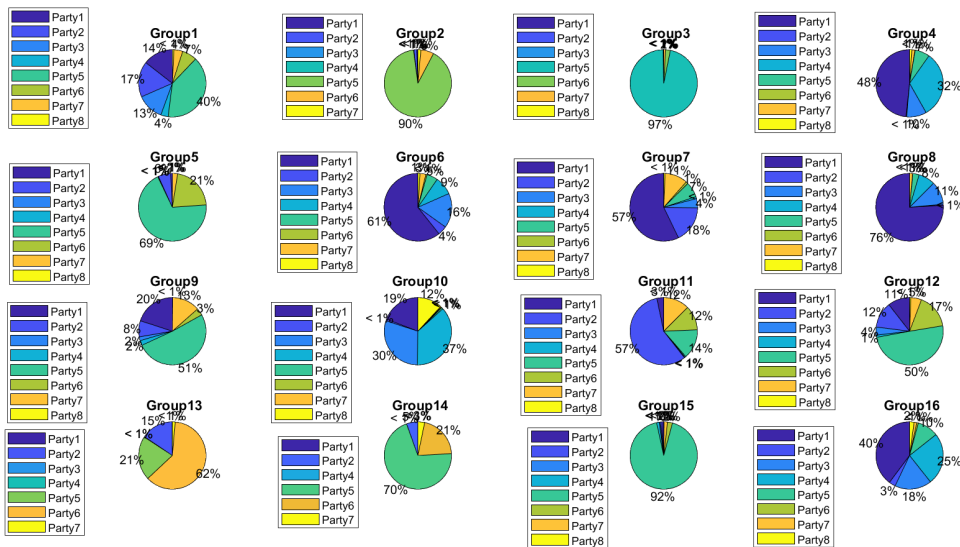
Σχήμα VII.21: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Πολωνία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες



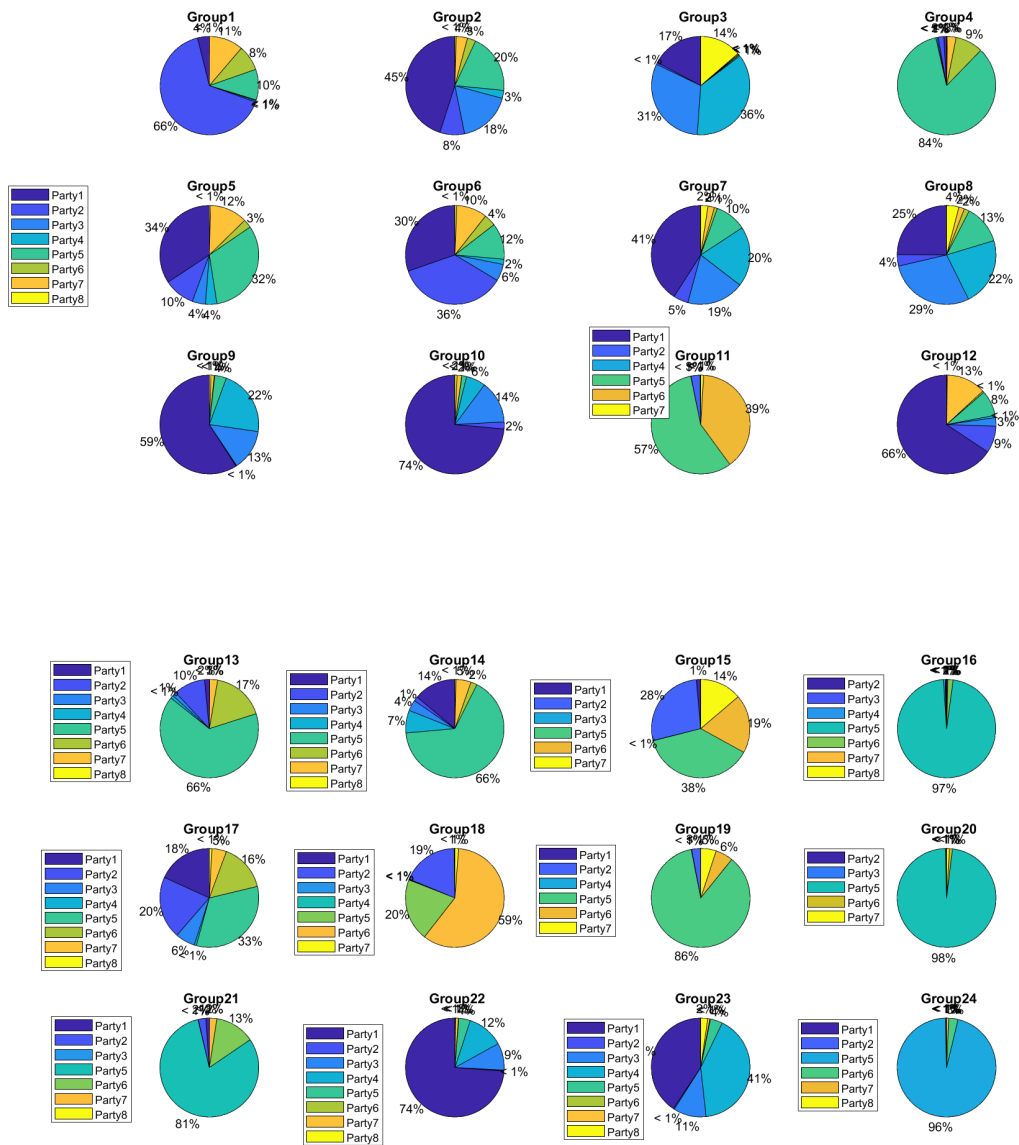
Σχήμα VII.22: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 8 ομάδες



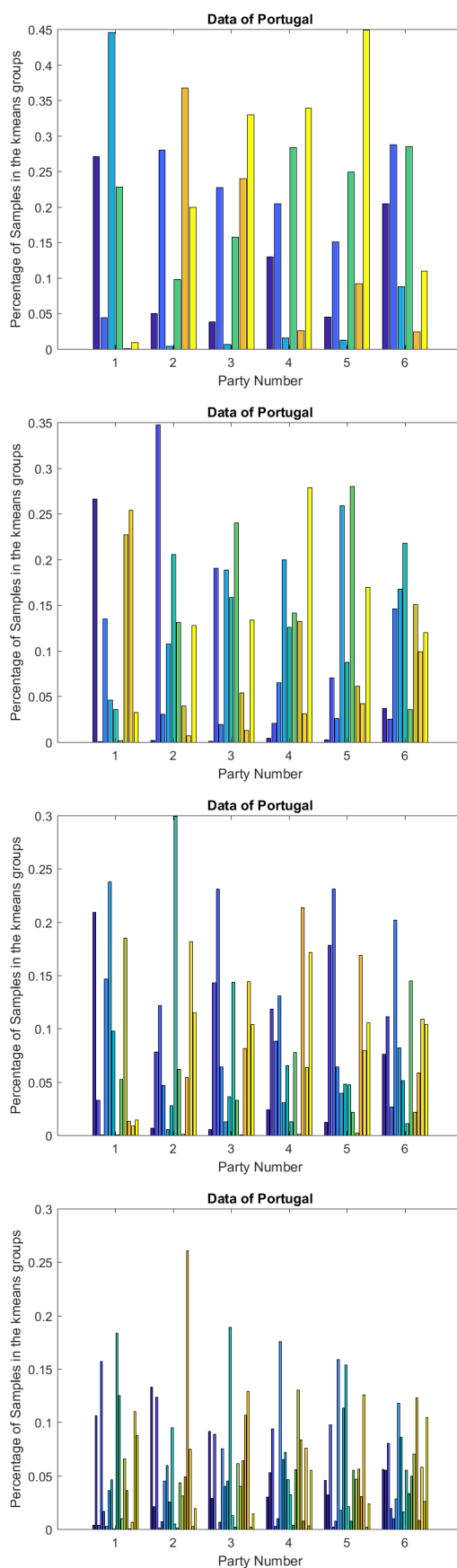
Σχήμα VII.23: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 12 ομάδες



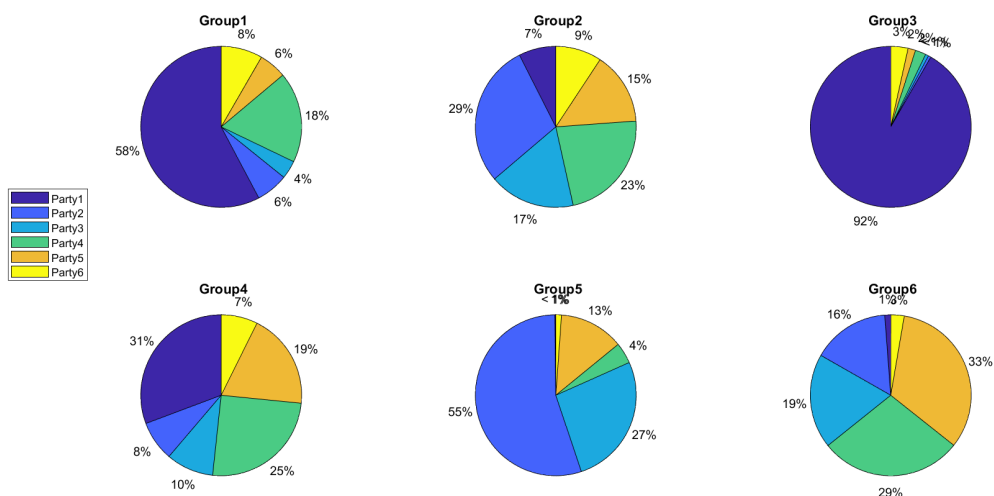
Σχήμα VII.24: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 16 ομάδες



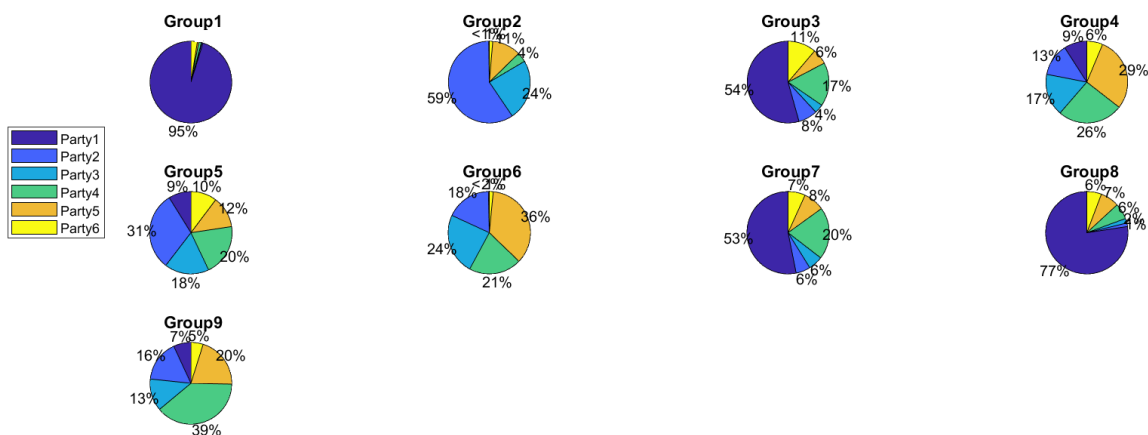
Σχήμα VII.25: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πολωνία – 24 ομάδες



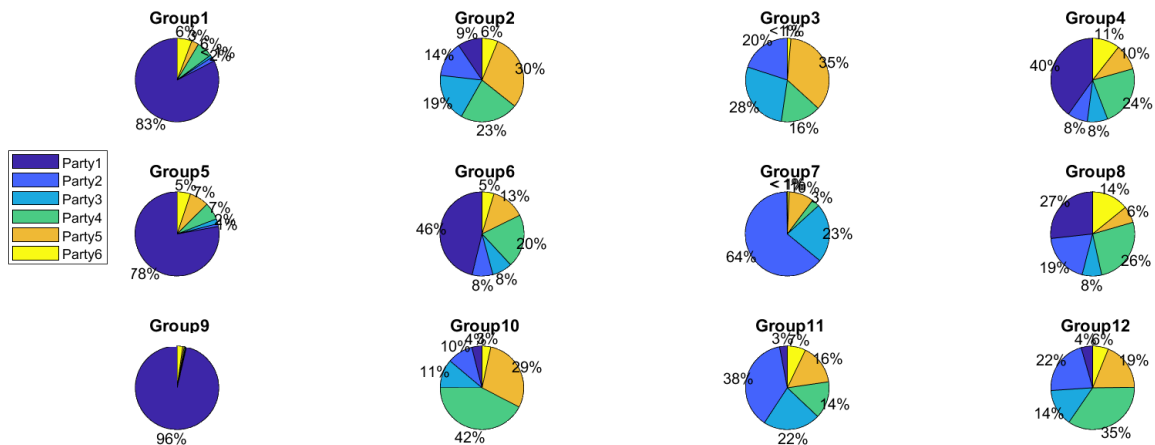
Σχήμα VII.26: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Πορτογαλία (α) 6 ομάδες (β) 9 ομάδες (γ) 12 ομάδες (δ) 18 ομάδες



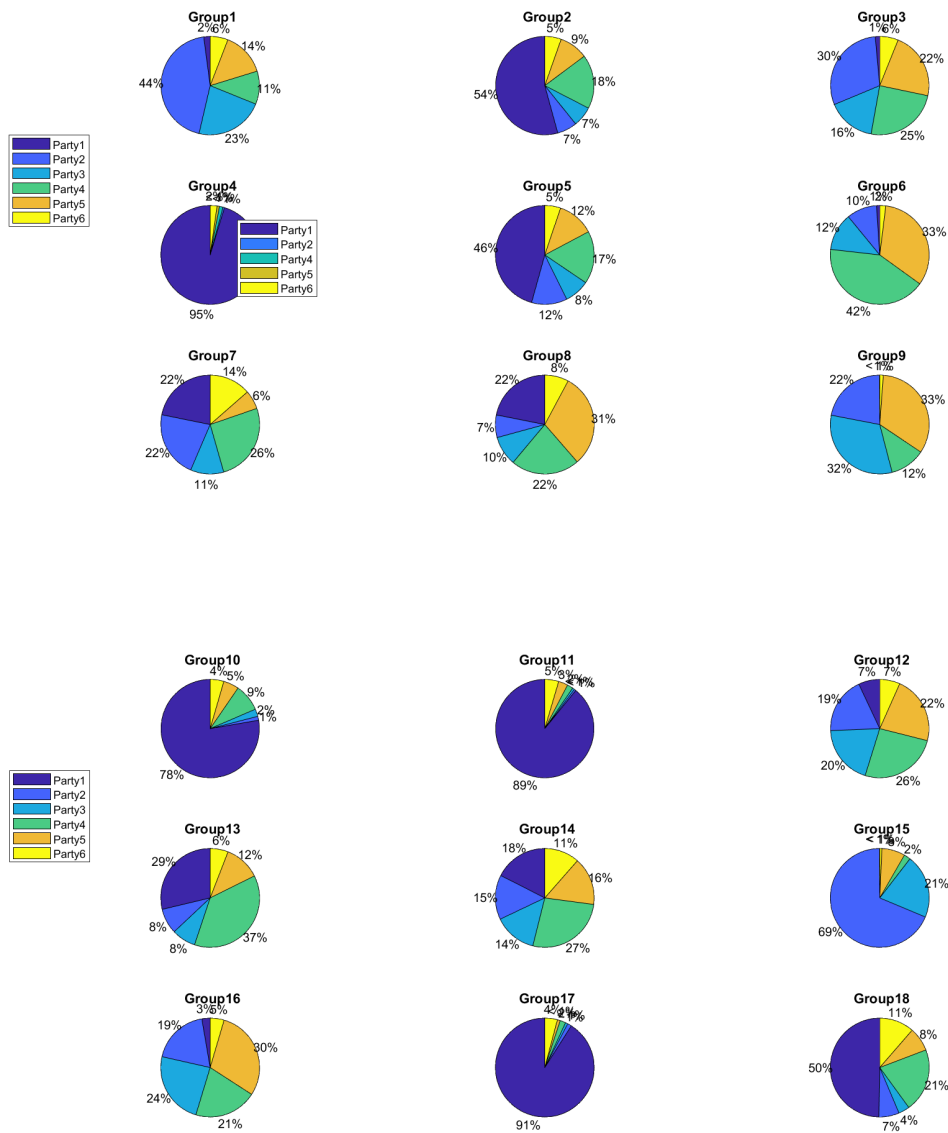
Σχήμα VII.27: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 6 ομάδες



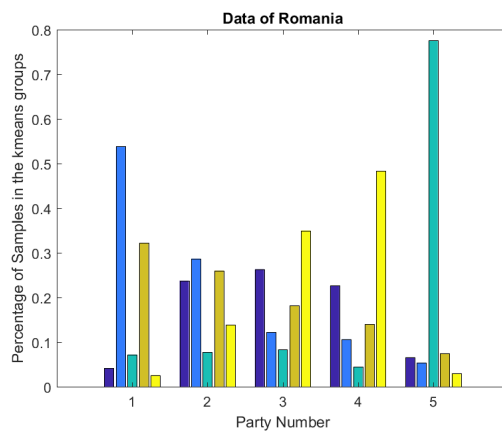
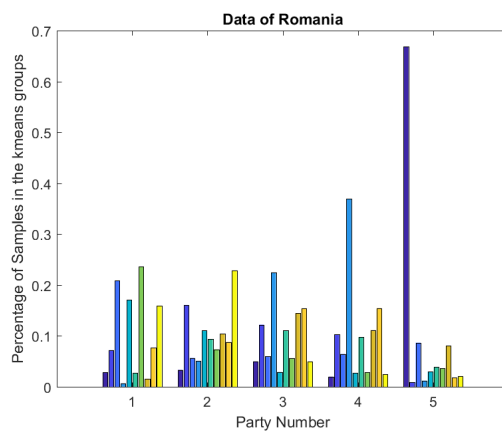
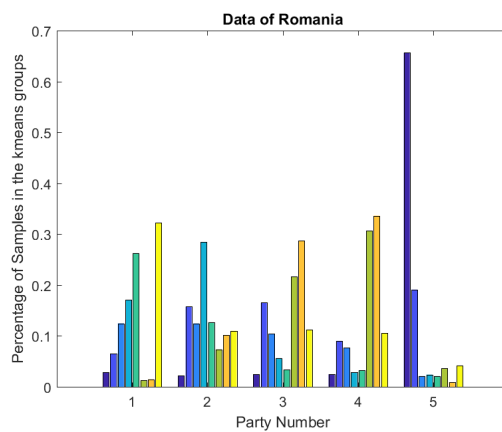
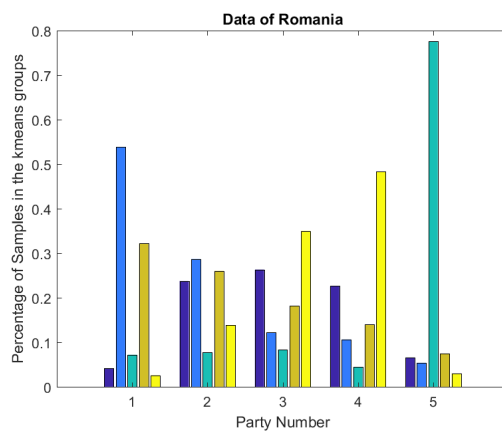
Σχήμα VII.28: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 9 ομάδες



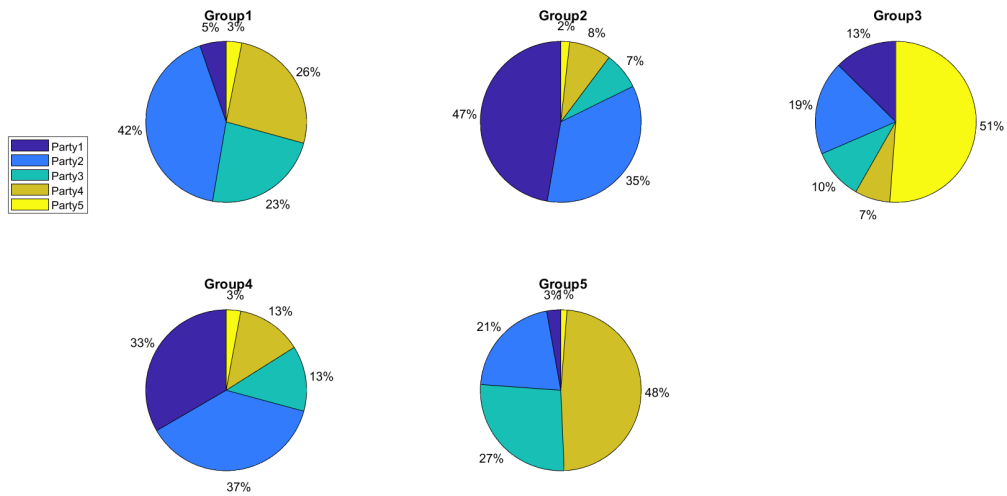
Σχήμα VII.29: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία – 12 ομάδες



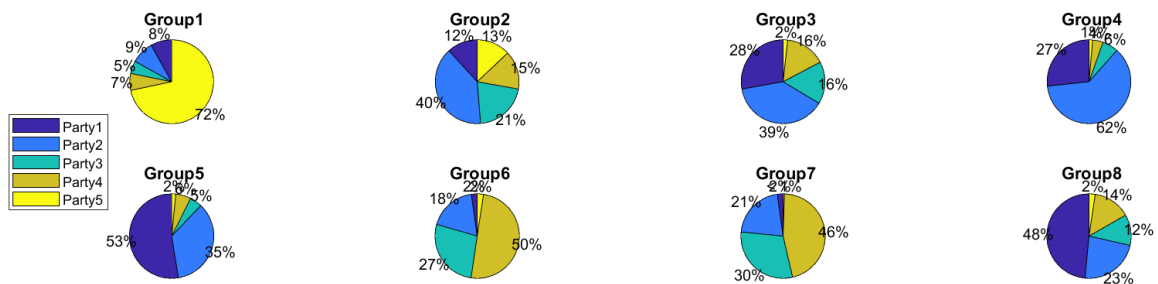
Σχήμα VII.30: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Πορτογαλία– 18 ομάδες



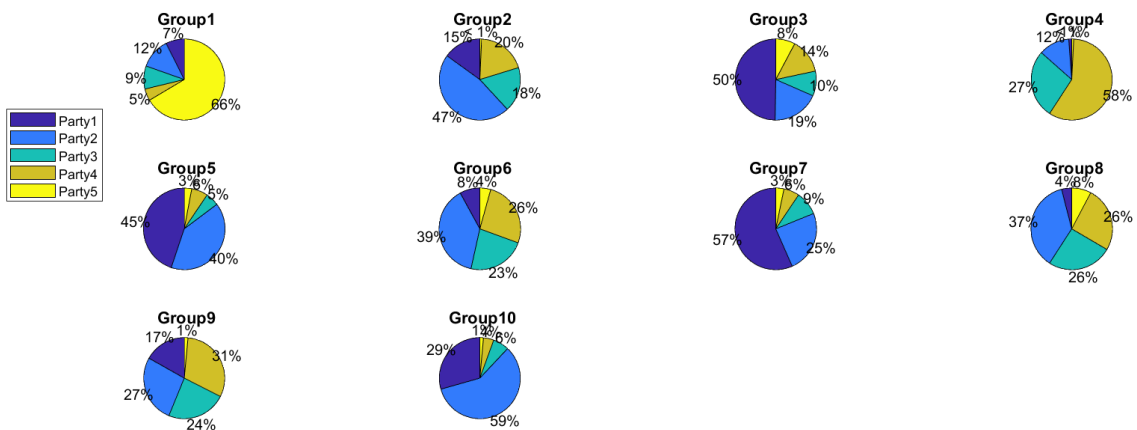
Σχήμα VII.31: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Ρουμανία (α) 5 ομάδες (β) 8 ομάδες (γ) 10 ομάδες (δ) 15 ομάδες



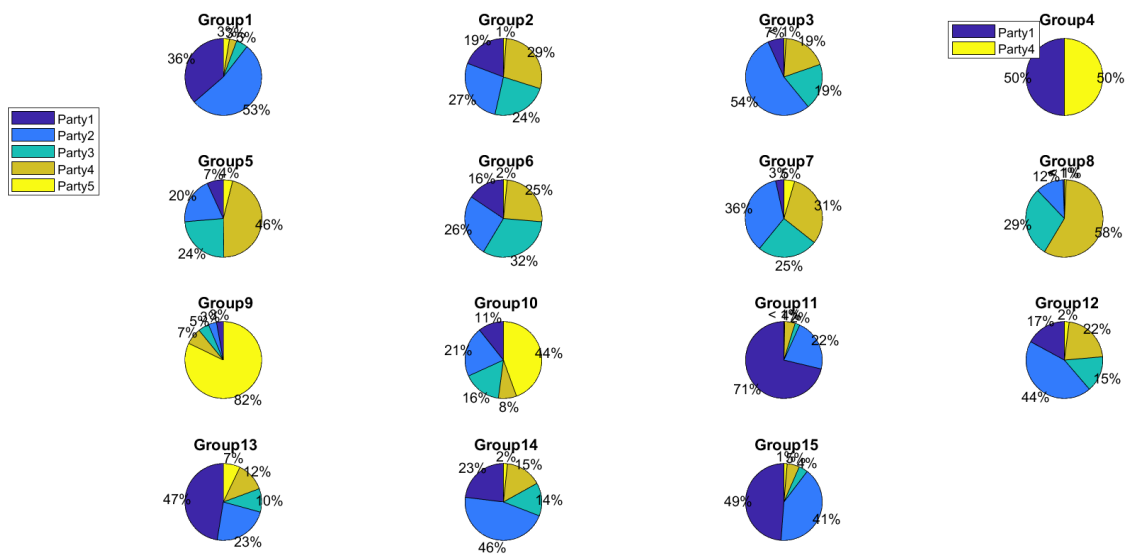
Σχήμα VII.32: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ρουμανία – 5 ομάδες



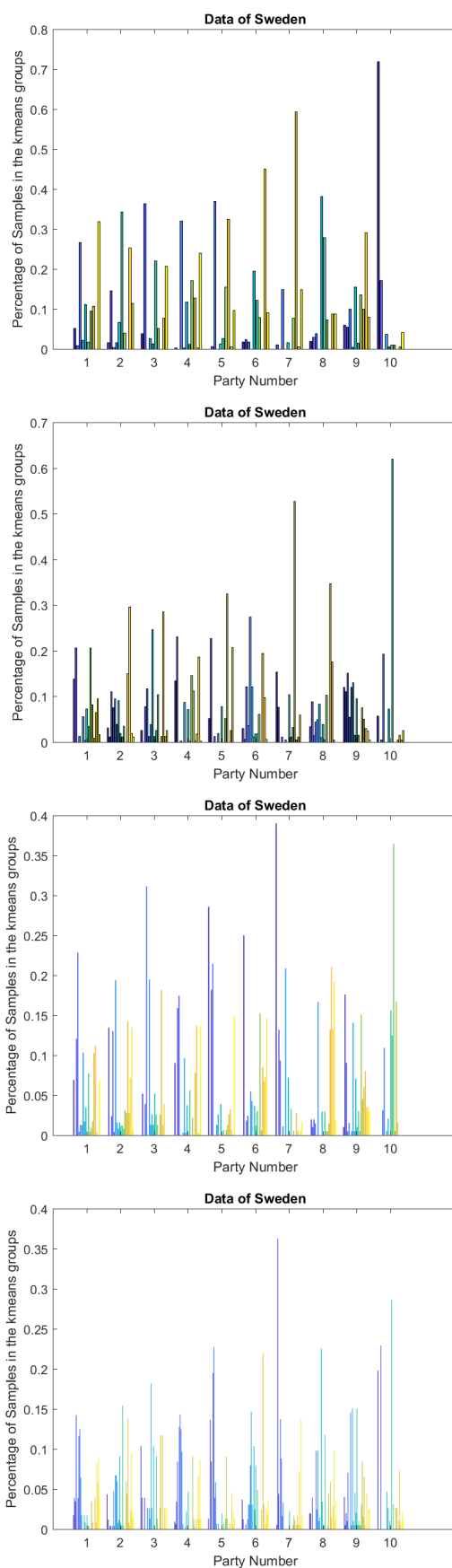
Σχήμα VII.33: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα –Ρουμανία – 8 ομάδες



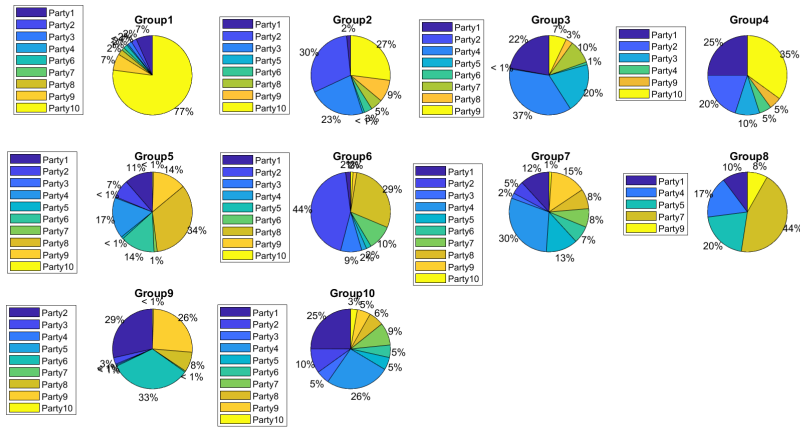
Σχήμα VII.34: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Ρουμανία – 10 ομάδες



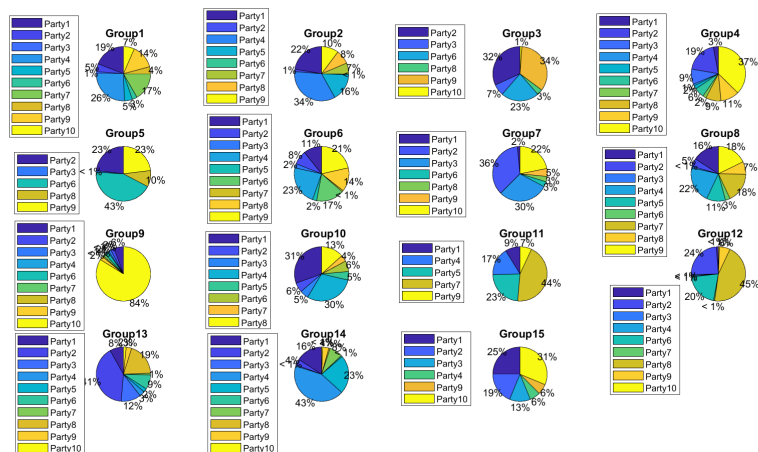
Σχήμα VII.35: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα –Ρουμανία – 15 ομάδες



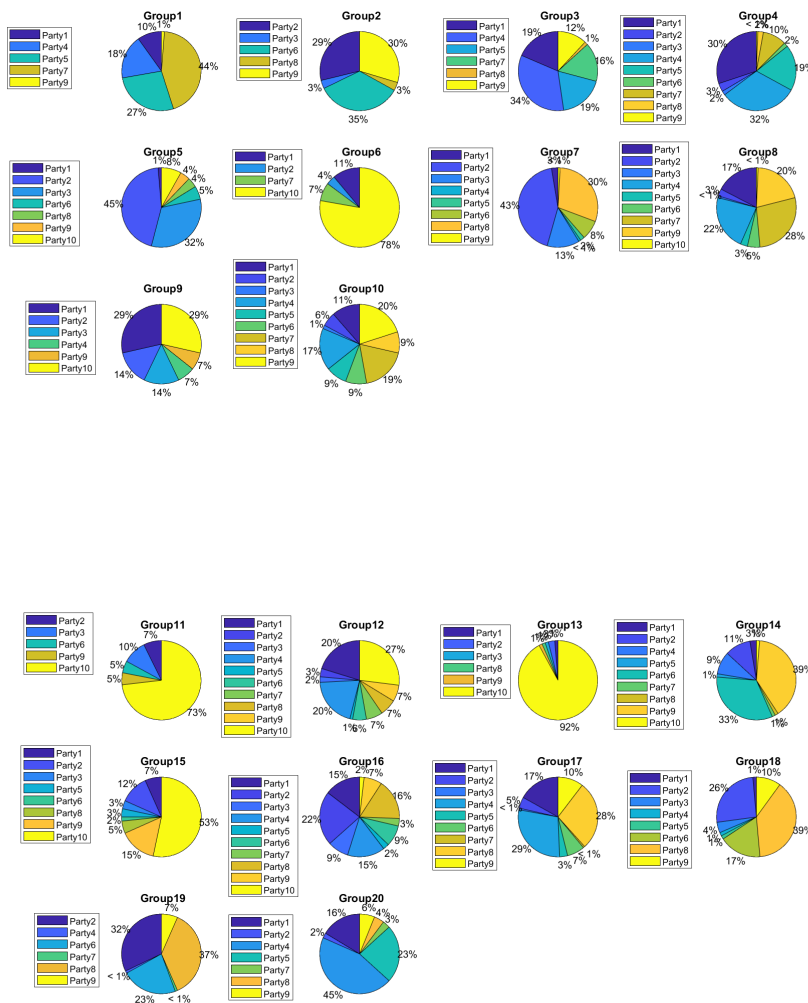
Σχήμα VII.36: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Σουηδία (α) 10 ομάδες (β) 15 ομάδες (γ) 20 ομάδες (δ) 30 ομάδες



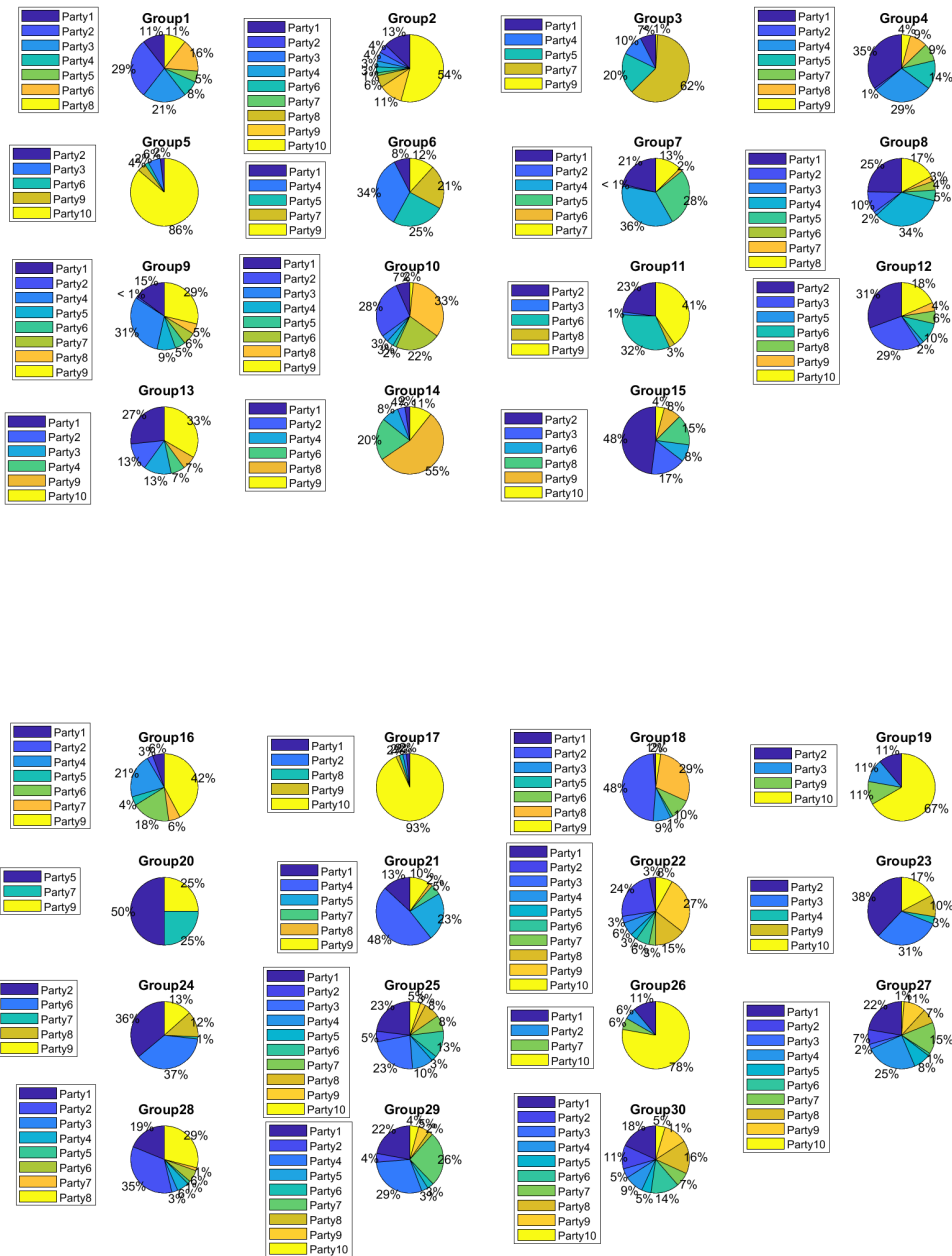
Σχήμα VII.37: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 10 ομάδες



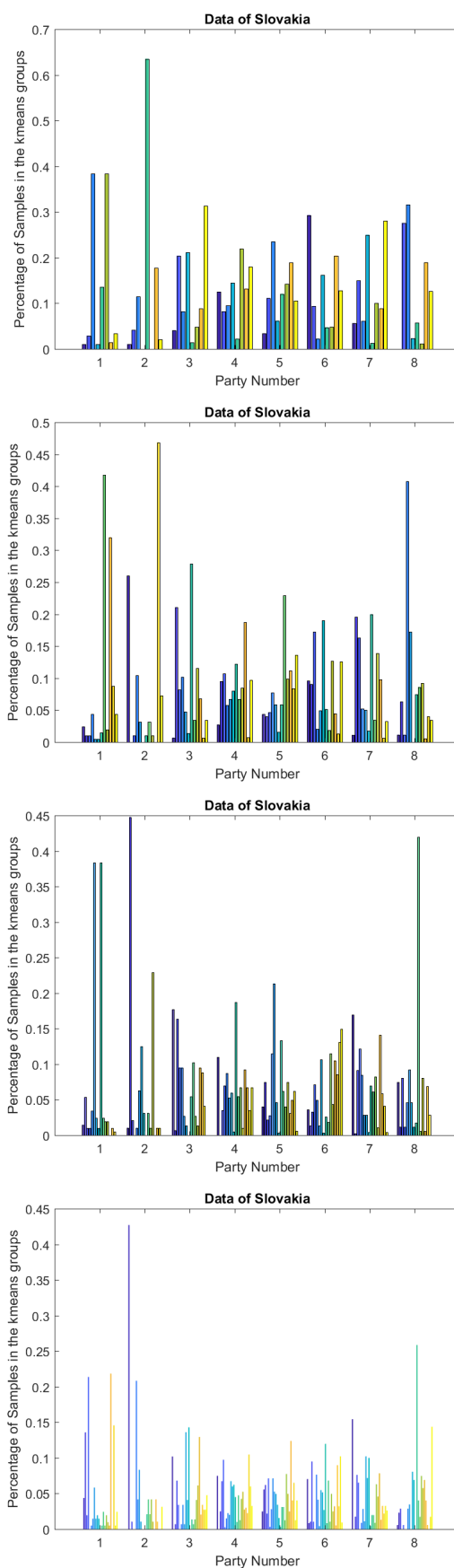
Σχήμα VII.38: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 15 ομάδες



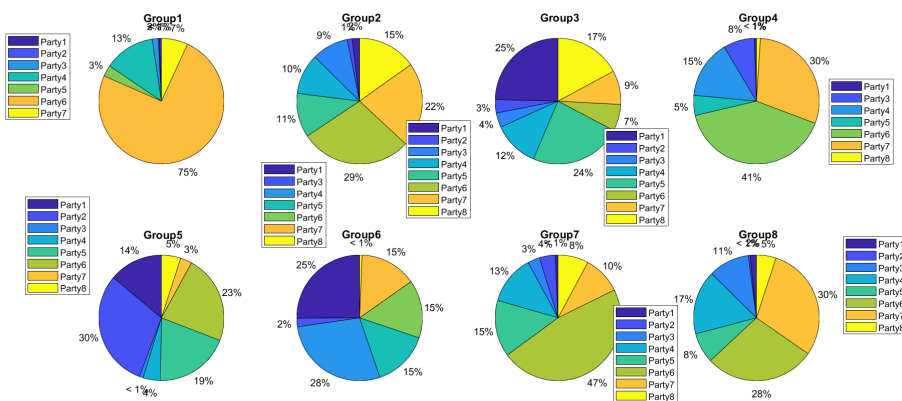
Σχήμα VII.39: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 20 ομάδες



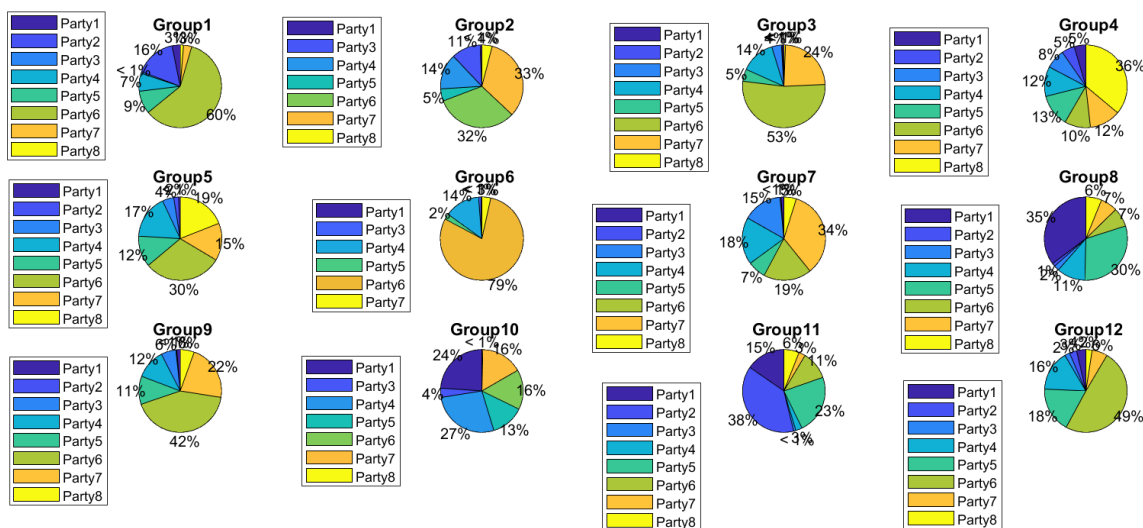
Σχήμα VII.40: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σουηδία – 30 ομάδες



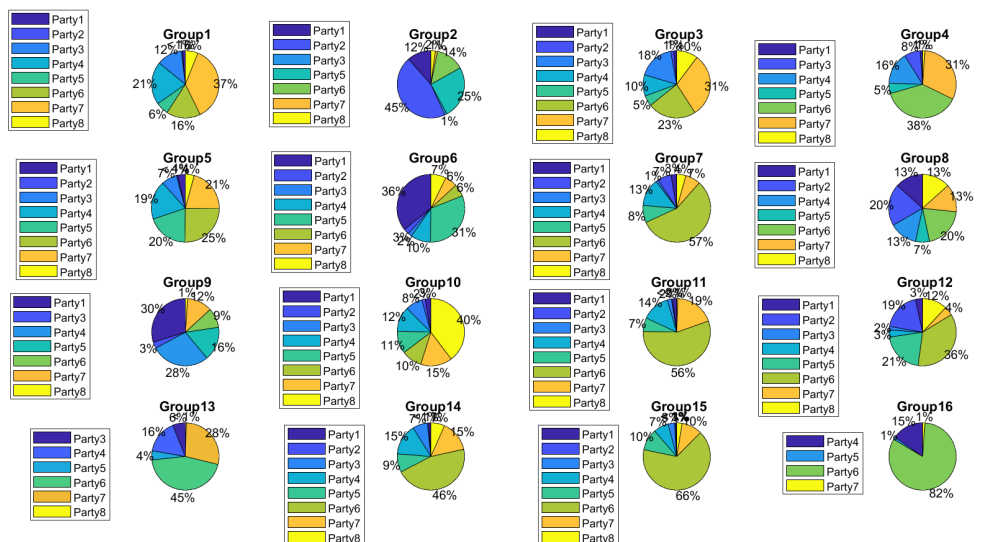
Σχήμα VII.41: Το ποσοστό των χρηστών ανά κόμμα που βρίσκεται σε κάθε ομάδα – Σλοβακία (α) 8 ομάδες (β) 12 ομάδες (γ) 16 ομάδες (δ) 24 ομάδες



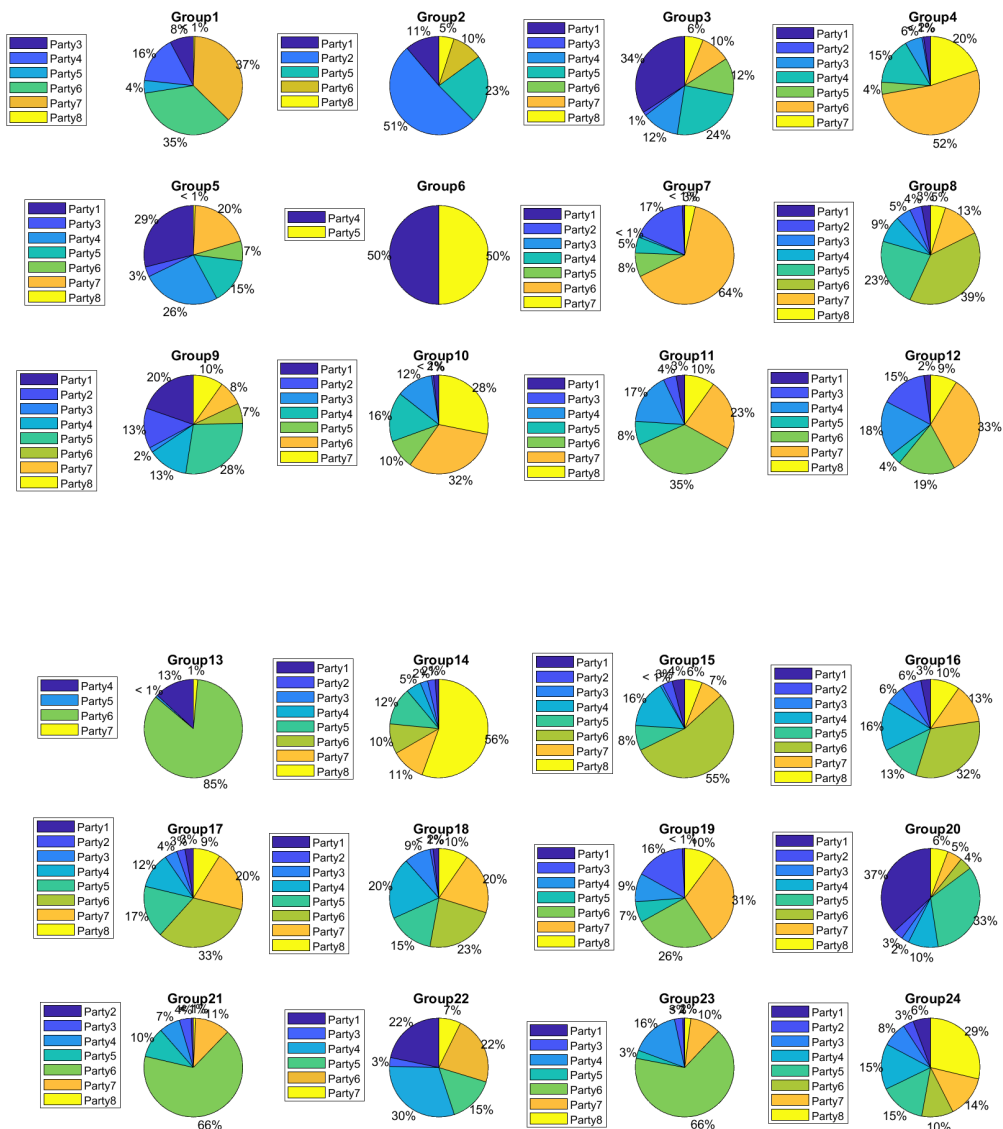
Σχήμα VII.42: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 8 ομάδες



Σχήμα VII.43: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 12 ομάδες



Σχήμα VII.44: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 16 ομάδες



Σχήμα VII.45: Το ποσοστό των κομμάτων που αποτελούν την κάθε ομάδα – Σλοβακία – 24 ομάδες

Ευρετήριο

k πλησιέστεροι γείτονες, 33

Αλληλεπίδραση Ανθρώπου Υπολογιστή, 7

Ανάκτηση Πληροφορίας, 7

Εξόρυξη Δεδομένων, 7

Ηλεκτρονικός Σύμβουλος Ψήφου, 1

Κρυμμένα Μοντέλα Markov, 104

Μηχανές Διανυσμάτων Υποστήριξης, 30

Μηχανική Μάθηση, 7

Συνεργατικό Φιλτράρισμα, 1, 6

αριθμητικός μέσος ακριβείας, 37

βασικό ερωτηματολόγιο, 15

δηλώσεις πολιτικής, 15

διανύσματα υποστήριξης, 35

ηλεκτρονική δημοκρατία, 11

καθαρισμός δεδομένων, 25

μοντέλο διανυσματικού χώρου, 29

νευρωνικά δίκτυα πρόσθιας τροφοδότησης, 34

παραγοντική ανάλυση, 23

συμπληρωματικές ερωτήσεις, 16

συνεργατικό φιλτράρισμα, 8

συσχέτιση Pearson, 43

φιλτράρισμα με βάση το περιεχόμενο, 8