

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ



## Πτυχιακή εργασία

TWITTER SENTIMENT ANALYSIS FOR FX

Στέφανος Μανώλη

Λεμεσός 2016

ΤΕΧΝΟΛΟΓΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ  
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ  
ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

## **Πτυχιακή εργασία**

**TWITTER SENTIMENT ANALYSIS FOR FX**

Στέφανος Μανώλη

Σύμβουλος καθηγητής  
Δρ. Σωτήριος Χατζής

Λεμεσός 2016

# Πνευματικά δικαιώματα

Copyright © Στέφανος Μανώλη, 2016

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της πτυχιακής εργασίας από το Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής του Τεχνολογικού Πανεπιστημίου Κύπρου δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέπων καθηγητή μου Δρ. Σωτήριο Χατζή κυρίως για την εμπιστοσύνη που μου έδειξε, και την υπομονή που έκανε κατά τη διάρκεια υλοποίησης της πτυχιακής εργασίας. Όπως επίσης και για την πολύτιμη βοήθεια και καθοδήγηση της, για την επίλυση διάφορων θεμάτων.

Θα ήθελα επίσης να απευθύνω τις ευχαριστίες μου στους γονείς μου, οι οποίοι στήριξαν τις σπουδές μου με διάφορους τρόπους, φροντίζοντας για την καλύτερη δυνατή μόρφωση μου.

## ΠΕΡΙΛΗΨΗ

Η συγκεκριμένη διπλωματική εργασία θα ασχοληθεί αρχικά με την εξόρυξη δεδομένων μέσα από κοινωνικά μέσα δικτύωσης που στο επίκεντρο τους έχουν τον δείκτη FOREX. Ο δείκτης FOREX ή αλλιώς FX είναι μια αγορά, στην οποία πραγματοποιούνται οι συναλλαγές των διαφόρων διεθνών συναλλαγμάτων. Ως αποτέλεσμα του τεράστιου όγκου της και της ρευστότητάς της, η αγορά FX εξελίχθηκε στη μεγαλύτερη και σημαντικότερη χρηματοπιστωτική αγορά του κόσμου. Πιο συγκεκριμένα, όπως αναφέρεται και στον τίτλο της διπλωματικής εργασίας, θα συλλεχτεί αρχικά, με την χρήση ενός API που θα δημιουργηθεί, μεγάλος όγκος δεδομένων από χρήστες του Twitter οι οποίοι σχετίζονται έμμεσα ή άμεσα με τον δείκτη FOREX. Στην συνέχεια θα επεξεργαστούν και θα ελεγχτούν ώστε να μην περιέχουν ελλιπή ή αχρείαστα δεδομένα. Με αυτό τον τρόπο θα μείνουν μόνο τα χρήσιμα δεδομένα τα οποία θα αποτελούν την σωστή πληροφορία που θα χρειαστεί για την εκπόνηση της διπλωματικής εργασίας. Αργότερα αφού συλλεχθούν και επεξεργαστούν τα δεδομένα θα πρέπει να διαμορφωθεί ανάλογα ένας υλοποιημένος αλγόριθμος ο οποίος με την κατάλληλη εκπαίδευση θα αποτελεί ένα σύστημα. Το σύστημα αυτό θα είναι σε θέση να παίρνει ως είσοδο τα επεξεργασμένα δεδομένα και ως έξοδο να κατηγοριοποιεί αυτά τα δεδομένα. Πιο συγκεκριμένα θα ασχοληθούμε με την ανάλυση συναισθημάτων μέσω των δεδομένων από το Twitter όσο αφορά τον οικονομικό δείκτη FOREX. Για την εκπόνηση της παρούσας διπλωματικής εργασίας θα χρησιμοποιηθεί η γλώσσα προγραμματισμού Python και πιο συγκεκριμένα οι δύο αλγόριθμοι ομαδοποίησης, Latent Dirichlet Allocation (LDA) και Non-negative Matrix Factorization (NMF).

# ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ.....	iv
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ .....	v
ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ .....	vii
ΑΠΟΔΟΣΗ ΟΡΩΝ.....	ix
1. ΕΙΣΑΓΩΓΗ .....	1
1.1. Διατύπωση ερευνητικού προβλήματος .....	1
1.2. Αναγκαιότητα έρευνας .....	2
2. Βιβλιογραφική ανασκόπηση .....	3
2.1. Μέσο κοινωνικής δικτύωσης Twitter.....	3
2.2. Αρχή Λειτουργίας του Twitter.....	3
2.3. Αγορά συναλλάγματος Forex.....	4
3. Θεωρητικό υπόβαθρο.....	5
3.1.1. Γλώσσα προγραμματισμού Python .....	5
3.1.2. Λόγος επιλογής Python .....	5
3.1.3. Χαρακτηριστικά Python .....	6
3.2. Μοντέλα ομαδοποίησης .....	6
3.2.1. NMF .....	6
3.2.2. LDA.....	6
3.3. Ομαδοποίηση.....	7
3.4. Τεχνικές Ομαδοποίησης.....	7
3.4.1. Τεχνικές με επίβλεψη .....	7
3.4.1.1.Πλεονεκτήματα .....	7

3.4.1.2.Μειονέκτηματα.....	7
3.4.2. Τεχνικές χωρίς επίβλεψη .....	8
3.4.2.1.Πλεονεκτήματα .....	8
3.4.2.2.Μειονεκτήματα.....	8
4. Μεθοδολογία.....	9
4.1. Γενικές δυσκολίες.....	9
4.1.1. Δημιουργία API.....	9
4.1.2. Επεξεργασία Δεδομένων .....	9
4.2. Συλλογή Δεδομένων .....	10
4.2.1. Μέσο συλλογής δεδομένων.....	10
4.2.2. Τρόπος συλλογής δεδομένων .....	11
4.2.3. Επεξεργασία Δεδομένων .....	12
4.3. Εξαγωγή Topics.....	12
4.3.1. Ορίσματα του παρόν συστήματος .....	13
5. Αποτελέσματα.....	14
5.1. Γραφικές παράστασεις με αποτελέσματα.....	15
6. Συμπεράσματα .....	18
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	19
ΠΑΡΑΡΤΗΜΑΤΑ.....	20

## ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1 - Twitter Application.....	11
Εικόνα 2 - Μορφοποίηση Δεδομένων.....	12
Εικόνα 3 - Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 100 features .....	15
Εικόνα 4 - Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 1000 features .....	16
Εικόνα 5 - Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 2000 features .....	16
Εικόνα 6 - Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 5000 features .....	17
Εικόνα 7 - Σύγκριση δύο μοντέλων με βάσει τους μεγαλύτερους μέσους όρους.....	18



## ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

ΤΕΠΑΚ:	Τεχνολογικό Πανεπιστήμιο Κύπρου
API	Application Programming Interface
NMF	Non-negative Matrix Factorization
LDA	Latent Dirichlet allocation
FOREX	Foreign Exchange
FX	FOREX

## ΑΠΟΔΟΣΗ ΟΡΩΝ

Topic	Θέμα
Twitter	Μέσο κοινωνικής δικτύωσης
Tweets	Δημοσιεύσεις
Retweet	Κοινοποίηση δημοσίευσης
Python	Γλώσσα προγραμματισμού
Facebook	Μέσο κοινωνικής δικτύωσης
Follow	«Ακολουθώ» ένα λογαριασμό στο Twitter
Hashtags	Ετικέτα
Reply	Απάντηση σε μια δημοσίευση στο Twitter
Cluster	Ομάδα από θέματα
Supervised	Με επίβλεψη
Unsupervised	Χωρίς επίβλεψη
Modules	Ενότητα
Features	Χαρακτηριστικά
FOREX	Ξένα συναλλάγματα

# 1. ΕΙΣΑΓΩΓΗ

## 1.1. Διατύπωση ερευνητικού προβλήματος

Η παρούσα διπλωματική εργασία θα ασχοληθεί με την ανάλυση δεδομένων που θα συλλεχτούν μέσα από αναρτήσεις συγκεκριμένων προσωπικοτήτων και εταιριών στα μέσα κοινωνικής δικτύωσης. Πιο συγκεκριμένα, τα δεδομένα θα συλλεχτούν μέσα από το κοινωνικό δίκτυο Twitter. Θα μελετηθεί η δυνατότητα ανάλυσης συναισθημάτων όσο αφορά τον οικονομικό δείκτη FOREX μέσω των δεδομένων που θα συλλεχτούν.

Αναλυτικότερα, η παρούσα πτυχιακή εργασία θέτει τους παρακάτω στόχους:

- Δημιουργία ενός API το οποίο θα διαβάζει δεδομένα από συγκεκριμένους χρήστες του Twitter και θα τα αποθηκεύει σε αρχεία τύπου csv.
- Εύρεση μεθόδου με την οποία θα επεξεργαστούν τα δεδομένα και θα απαλλαγθούν από τυχόν αχρείαστα δεδομένα και θόρυβο.
- Ανάλογη διαμόρφωση των υλοποιημένων αλγόριθμων LDA και NMF σε γλώσσα προγραμματισμού python
- Ανάλυση των αποτελεσμάτων
- Εξαγωγή συμπερασμάτων
- Μελέτη προοπτικής για εφαρμογή του παρόντος συστήματος και σε άλλους τομείς

## **1.2.Αναγκαιότητα έρευνας**

Τα μέσα κοινωνικής δικτύωσης όπως είναι το Twitter, το Facebook και άλλα, τα τελευταία χρόνια είναι στο επίκεντρο της προσοχής. Τα συγκεκριμένα κοινωνικά δίκτυα χρησιμοποιούνται από εκατομμύρια χρήστες καθημερινά οι οποίοι τα χρησιμοποιούν κυρίως για επικοινωνία, διαφήμιση και ψυχαγωγία. Πλέον βάσει της χρήσης που τυγχάνουν καθημερινά διαθέτουν μεγάλο όγκο δεδομένων που είναι διαθέσιμα προς όλους τους χρήστες. Αυτά τα δεδομένα δίνουν την μοναδική ευκαιρία σε ανθρώπους που ασχολούνται με την επεξεργασία φυσικής γλώσσας να τα εκμεταλλευτούν σε συστήματα όπως το συγκεκριμένο (Xiang, Zhou, 2014). Με λίγα λόγια αυτή την ευκαιρία μπορεί να την εκμεταλλευτεί οποιοσδήποτε ακόμα και εταιρίες μικρές και μεγάλες για εμπορικούς λόγους όπως είναι η διαφήμιση. Ένα σχετικό παράδειγμα μπορεί αποτελέσει μία εταιρεία η οποία κυκλοφορεί στην αγορά ένα πρωτότυπο προϊόν. Η εν λόγω εταιρεία μπορεί μέσω της ανάλυσης των δεδομένων που θα έχει στην διάθεση της από τα μέσα κοινωνικής δικτύωσης να εισπράξει το συναίσθημα του κόσμου για αυτό το προϊόν και ανάλογα μετά θα γνωρίζει που και πώς να το προωθήσει.

## **2. Βιβλιογραφική ανασκόπηση**

### **2.1. Μέσο κοινωνικής δικτύωσης Twitter**

Το twitter είναι ένα από τα δημοφιλέστερα κοινωνικά δίκτυα παγκόσμια. Πιο συγκεκριμένα είναι μια υπηρεσία η οποία δίνει την δυνατότητα στους χρήστες της να κρατούν επαφή, να ενημερώνονται για τις καθημερινές δραστηριότητες των φίλων τους και των ανθρώπων που τους ενδιαφέρει. Επίσης όπως και άλλα κοινωνικά μέσα δικτύωσης δίνει την δυνατότητα σε χρήστες να το χρησιμοποιούν για εμπορικούς σκοπούς. Αυτό που ξεχωρίζει το συγκεκριμένο κοινωνικό δίκτυο από τα υπόλοιπα, είναι το γεγονός ότι οι δημοσιεύσεις (tweets) περιορίζονται στους 140 χαρακτήρες. Γλώσσες προγραμματισμού στο Android

### **2.2. Αρχή Λειτουργίας του Twitter**

Στο Twitter κάθε χρήστης δημιουργεί δωρεάν το δικό του προφίλ, το οποίο αποτελείται από εικόνες και διάφορες πληροφορίες, όπως μίνι βιογραφικό, περιοχή διαμονής, username κλπ. Το ψευδώνυμο, έχει πολύ μεγάλη σημασία καθώς χρησιμοποιείται περισσότερο και από το όνομα του χρήστη. Οι περισσότεροι είναι γνωστοί από το ψευδώνυμο τους, καθώς ένας χρήστης για να αναφέρει έναν άλλον χρησιμοποιεί το username. Κάθε προφίλ μπορεί να δημοσιεύσει μηνύματα με μέγιστο όριο τους 140 χαρακτήρες, τα λεγόμενα και ως tweets. Από προεπιλογή, κάθε λογαριασμός δημοσιεύει δημόσια, με οποιονδήποτε χρήστη του διαδικτύου να μπορεί να προβάλει τα μηνύματα του. Βέβαια, υπάρχει και επιλογή για ιδιωτικά tweets, αν και δε συνίσταται για την καλύτερη εμπειρία χρήσης της συγκεκριμένης εφαρμογής. Στο Twitter δε συναντάει κανείς φίλιες, αλλά ακόλουθους, ή αλλιώς followers. Δεν είναι απαραίτητο για δύο χρήστες να ακολουθεί ο ένας τον άλλον. Κάθε χρήστης ακολουθεί όποιους θέλει και ακολουθείται από οποιονδήποτε. Αυτό σημαίνει πως ο καθένας διαμορφώνει την αρχική του σελίδα με περιεχόμενο το οποίο επιθυμεί. Μία από τις σημαντικότερες λειτουργίες των Κοινωνικών Δικτύων, τα hashtags, καθιερώθηκαν από το Twitter και χρησιμοποιούνται σε μεγάλο βαθμό για συγκέντρωση όλων των tweets γύρω από μια συζήτηση. Οι κυριότερες λειτουργίες του Twitter είναι το retweet το οποίο χρησιμοποιείται για κοινοποίηση ενός tweet, το favorite το οποίο χρησιμοποιείται περισσότερο ως το like του Facebook, και το reply το οποίο χρησιμοποιείται για απάντηση σε ένα tweet.

### **2.3. Αγορά συναλλάγματος Forex**

Η αγορά του Forex είναι μία παγκόσμια χρηματαγορά συναλλάγματος. Συχνά αναφέρεται και με την συντομογραφία FX. Η αγορά Forex αποτελεί την μεγαλύτερη χρηματοοικονομική αγορά στον κόσμο όπου πραγματοποιούνται αγοραπωλησίες όλων των εθνικών νομισμάτων ανάμεσα σε τράπεζες, οργανισμούς και μεγάλες εταιρίες, επενδυτές, εισαγωγείς και εξαγωγείς, κερδοσκόπους και μεσίτες συναλλάγματος. Οι πιο πάνω συναλλαγές εκτελούνται μόνο μέσω διαδικτύου 24 ώρες το εικοσιτετράωρο και 5 μέρες την βδομάδα ανά τον κόσμο. Βασικός στόχος των επενδυτών είναι το κέρδος που μπορεί να τους επιφέρουν οι αγοραπωλησίες εθνικών νομισμάτων με βάση στις διακυμάνσεις των ισοτιμιών.

### **3. Θεωρητικό υπόβαθρο**

#### **3.1.1. Γλώσσα προγραμματισμού Python**

Η γλώσσα προγραμματισμού Python είναι μια δυνατή γλώσσα προγραμματισμού, πολύ υψηλού επιπέδου και ανήκει στην κατηγορία των γλωσσών προγραμματισμού που είναι τόσο εύκολες και παράλληλα τόσο ισχυρές. Αποτελεί μία από τις πιο δημοφιλή γλώσσες προγραμματισμού στο τομέα αυτό. Τέτοιες γλώσσες εκτός από την Python είναι κυρίως η Java και η C++ οι οποίες αποτελούν και τα δυνατά όπλα στο κόσμο του προγραμματισμού. Συνήθως όμως οι προγραμματιστές προτιμούν την Python από τις άλλες γλώσσες προγραμματισμού σε περιπτώσεις όπου θέλουν να αναπτύξουν μια μικρή εφαρμογή που να κάνει συγκεκριμένες λειτουργίες.

#### **3.1.2. Λόγος επιλογής Python**

Επιλέχθηκε η γλώσσα προγραμματισμού Python για αρκετούς λόγους. Καταρχήν η Python είναι αρκετά εύκολη γλώσσα όσο αφορά την εκμάθησή της και ταυτόχρονα έχει πολλές δυνατότητες. Επίσης με την συγκεκριμένη γλώσσα προγραμματισμού ο προγραμματιστής δεν χρειάζεται να μεταγλωττίζει τον κώδικά του, απλά τον αποθηκεύει και τον τρέχει. Παρέχει το πλεονέκτημα να δημιουργούνται modules που είναι επαναχρησιμοποιήσιμα για άλλες εφαρμογές γραμμένες σε Python. Αξιοσημείωτο είναι ότι κυρίως η Python χρησιμοποιείται για την ανάπτυξη μικρότερων εφαρμογών σε σύγκριση με άλλες αντικειμενοστραφείς γλώσσες προγραμματισμού όπως οι C++, C# και Java.

### **3.1.3. Χαρακτηριστικά Python**

Η γλώσσα προγραμματισμού Python χαρακτηρίζεται ως μια απλή γλώσσα αφού μοιάζει αρκετά με ψευδοκώδικα. Είναι εύκολη όσο αφορά την εκμάθηση της και εύκολη στην σύνταξη της. Αξίζει να σημειωθεί ότι η Python αποτελεί ένα παράδειγμα λογισμικού ανοικτού κώδικα. Για αυτό και αποτελεί μια πολύ καλή γλώσσα αφού δημιουργήθηκε και ακόμα βελτιώνεται από μια ομάδα προγραμματιστών έτσι ώστε να παράγουν μελλοντικές εκδοχές της Python που θα την καθιστούν ακόμα καλύτερη γλώσσα προγραμματισμού στο τομέα της. Λόγω του ότι είναι μία γλώσσα ανοικτού κώδικα εύκολα τρέχει σε όλες τις πλατφόρμες και σε όλα τα λογισμικά περιβάλλοντα. Είναι μια αντικειμενοστραφής γλώσσα υψηλού επιπέδου που δεν χρειάζεται μεταγλώττιση ο κώδικας της για να τρέξει.

## **3.2. Μοντέλα ομαδοποίησης**

### **3.2.1. Non-negative Matrix Factorization**

Ο αλγόριθμος παραγοντοποίησης μη αρνητικών τιμών (Non-negative Matrix Factorization) είναι μια πολύ ενδιαφέρουσα μέθοδος ομαδοποίησης κειμένων και συνδέετε στενά με τον αλγόριθμο K-means. Η συγκεκριμένη μέθοδος παίρνει σαν είσοδο ένα πίνακα  $X$  όπου η στήλη αντιστοιχεί στην συχνότητα εμφάνισης ενός εγγράφου. Εφαρμόζοντας την μέθοδο στο συγκεκριμένο πίνακα  $X$  τότε παίρνουμε σαν έξοδο δύο πίνακες  $U$  και  $V$  όπου  $U \cdot V = X$ . Οι δύο πίνακες  $U$  και  $V$  περιέχουν μόνο μη αρνητικές τιμές που χαρακτηρίζουν στην ουσία τα δεδομένα.

### **3.2.2. Latent Dirichlet Allocation**

Ο αλγόριθμος Λανθάνουσας Κατανομής Dirichlet είναι ένα παραγωγικό πιθανοκρατικό μοντέλο για συλλογές από διακριτά δεδομένα όπως μια συλλογή κειμένων (David M. Blei, 2003). Ο συγκεκριμένος αλγόριθμος έχει την ικανότητα, μέσα από ένα σύνολο δεδομένων, να ανακαλύπτει ομάδες που δεν είχαν παρατηρηθεί στο παρελθόν μεταξύ των δεδομένων. Δηλαδή μπορεί να βρίσκει ομοιότητες μεταξύ των δεδομένων δεν είχαν ή δεν μπορούσαν να βρεθούν προηγουμένως.



### **3.3. Ομαδοποίηση**

Ο όρος ομαδοποίηση αναφέρεται στη διαδικασία χωρισμού ενός συνόλου δεδομένων σε άλλα πιο μικρά υποσύνολα (clusters). Πιο συγκεκριμένα οργανώνει μία συλλογή από δεδομένα σε ομάδες με βάση κάποιο μέτρο ομοιότητας. Οι τεχνικές ομαδοποίησης κατατάσσονται στις τεχνικές χωρίς επίβλεψη, με επίβλεψη και με ημί-επίβλεψη.

### **3.4. Τεχνικές Ομαδοποίησης**

#### **3.4.1. Τεχνικές με επίβλεψη**

Στις τεχνικές με επίβλεψη, στην φάση της εκπαίδευσης ο χρήστης ξέρει ποια είναι η είσοδος που δίνει και ποια θα είναι η επιθυμητή έξοδος.

##### **3.4.1.1. Πλεονεκτήματα**

- Με βάση τα δεδομένα που δίνει ο χρήστης ως είσοδο, τα αποτελέσματα δεν αλλάζουν από εκτέλεση σε εκτέλεση.
- Σε περίπτωση όπου τα αποτελέσματα δεν κρίνονται ικανοποιητικά, ο χρήστης έχει την δυνατότητα να δώσει διαφορετική είσοδο (περισσότερους σπόρους, διαφορετική θέση) έτσι ώστε να πάρει το καλύτερο αποτέλεσμα.

##### **3.4.1.2. Μειονέκτηματα**

- Ο αριθμός και η θέση των αρχικών σπόρων παίζουν σημαντικό ρόλο στο τελικό αποτέλεσμα. Κάθε χρήστης μπορεί να θεωρεί διαφορετικά σημεία σπόρους σε αντίθεση με άλλους χρήστες.
- Τις περισσότερες φορές τα δεδομένα που δίνουν οι χρήστες είναι μικρής ποσότητας και δεν μπορούν να βρεθούν εύκολα. Επίσης με βάση το γεγονός ότι η αξιοπιστία των δεδομένων έχει ρόλο στο τελικό αποτέλεσμα, υπάρχει κίνδυνος να υπάρχουν μεγάλες αποκλίσεις λόγω λάθος δεδομένων.

### **3.4.2. Τεχνικές χωρίς επίβλεψη**

Στις τεχνικές χωρίς επίβλεψη, ο χρήστης γνωρίζει μόνο τα δεδομένα που δίνει σαν είσοδο για την εκπαίδευση.

#### **3.4.2.1. Πλεονεκτήματα**

- Δεν χρειάζεται να οριστούν τα σημεία σπόρους έτσι δεν υπάρχει κίνδυνος να επιδράσει το αποτέλεσμα. Χρειάζεται μόνο ο αριθμός των ομάδων (clusters) στις οποίες θα χωριστούν τα δεδομένα.

#### **3.4.2.2. Μειονεκτήματα**

- Η απόδοση του τελικού αποτελέσματος εξαρτάται από το βαθμό της αναπαράστασης που επιλέγουμε.
- Ο μέσος χρόνος εκτέλεσης είναι μεγαλύτερος σε σύγκριση με τις τεχνικές με επίβλεψη λόγω του ότι στις τεχνικές με επίβλεψη ο χρήστης δίνει επιπλέον πληροφορίες.

## **4. Μεθοδολογία**

### **4.1. Γενικές δυσκολίες**

#### **4.1.1. Δημιουργία API**

Όσο αφορά την δημιουργία του API με την βοήθεια της γλώσσας προγραμματισμού python αρχικά έγινε προσπάθεια να δημιουργηθεί κάτι νέο με πληροφορίες από το διαδίκτυο. Έγιναν πολλές προσπάθειες αλλά δεν είχαν κανένα αποτέλεσμα. Τελικά λύθηκε το πρόβλημα αυτό αφού για να μπορεί κάποιος χρήστης να πάρει μαζικά δεδομένα από το κοινωνικό δίκτυο Twitter έπρεπε να έχει τις απαραίτητες άδειες δημιουργώντας ένα Twitter\_App με το λογαριασμό του και χρησιμοποιώντας επίσης την βιβλιοθήκη tweepy.

#### **4.1.2. Επεξεργασία Δεδομένων**

Αφού συλλέχτηκαν και αποθηκεύτηκαν τα δεδομένα που θα χρησιμοποιηθούν στο σύστημα ως είσοδο ήταν αναγκαίο να μεταφερθούν σε αρχεία τύπου text (txt) και να τύχουν της ανάλογης επεξεργασίας. Η δυσκολία που αντιμετωπίστηκε τελικά με την απλή μέθοδο «Find and Replace» ήταν ότι τα δεδομένα περιείχαν αρκετούς άσχετους χαρακτήρες οι οποίοι δεν άφηναν τα δεδομένα να φορτωθούν σωστά στο σύστημα. Επίσης με την ίδια μέθοδο αντιμετωπίστηκε και το γεγονός ότι τα δεδομένα περιείχαν πολλούς αριθμούς λόγω κυρίως της ημερομηνίας και ώρας της κάθε δημοσίευσης. Λόγω των πολλών αριθμών που περιείχαν τα δεδομένα υπήρχε αλλοίωση στην ομαδοποίηση των δεδομένων και στο αποτέλεσμα.

## **4.2. Συλλογή Δεδομένων**

Για την ολοκλήρωση της παρούσας πτυχιακής εργασίας χρειαζόταν να συλλεχτεί μεγάλος όγκος δεδομένων, τα οποία δεδομένα θα σχετίζονται με τον οικονομικό δείκτη forex. Τα δεδομένα αυτά χρειάζονται για να δοθούν ως είσοδος στο σύστημα από το οποίο θα παίρνουμε σαν γενικό αποτέλεσμα το συναίσθημα της αγοράς όσο αφορά τις οικονομικές αλλαγές και συγκεκριμένα τον δείκτη forex.

### **4.2.1. Μέσο συλλογής δεδομένων**

Τα δεδομένα τα οποία χρησιμοποιήθηκαν για να επιτευχθεί ο στόχος της παρούσας πτυχιακής εργασίας συλλέχτηκαν αποκλειστικά μέσα από το κοινωνικό δίκτυο Twitter. Συγκεκριμένα δημιουργήθηκε ένας νέος προσωπικός λογαριασμός στο Twitter αποκλειστικά για την συλλογή των δεδομένων και στην συνέχεια μέσω αυτού του λογαριασμού ακολουθήθηκαν (follow) συγκεκριμένοι άλλοι χρήστες του Twitter οι οποίοι είναι φυσικά πρόσωπα ή εταιρίες. Τα συγκεκριμένα πρόσωπα και εταιρίες έχουν άμεση ή έμμεση σχέση με τον οικονομικό δείκτη Forex και τις οικονομικές συναλλαγές. Κυρίως είναι άνθρωποι και εταιρίες που ασχολούνται με οικονομικά και εμπορικά θέματα μεγάλης εμβέλειας στο εξωτερικό. Τα δεδομένα αποτελούν όλες οι δημοσιεύσεις (tweets) τον πιο πάνω λογαριασμών οι οποίες όμως δημοσιεύσεις περιέχουν δύο συγκεκριμένες φράσεις, «#FOREX» ή «#FX». Δεν υπάρχει διάκριση κεφαλαίων και πεζών γραμμάτων όσο αφορά τις δύο πιο πάνω φράσεις.

## 4.2.2. Τρόπος συλλογής δεδομένων

Για να συλλεχτούν τα πιο πάνω δεδομένα χρειάστηκε να δημιουργηθεί αρχικά ένα API με την βοήθεια της γλώσσας προγραμματισμού Python. Συγκεκριμένα μέσω του API έχουν συλλεχτεί τα δεδομένα μας και έχουν αποθηκευτεί σε αρχεία τύπου comma separated values (csv). Για την δημιουργία του συγκεκριμένου API πάρθηκαν πληροφορίες από την ίδια την ιστοσελίδα του Twitter το οποίο παρέχει συγκεκριμένη βοηθητική βιβλιοθήκη για την python, την βιβλιοθήκη tweepy. Η συγκεκριμένη βιβλιοθήκη περιέχει ένα τυποποιημένο API το οποίο για να μπορέσει να τρέξει και να λειτουργήσει σωστά έτσι ώστε να κατεβάσει και να αποθηκεύσει τα δεδομένα, χρειάζεται τέσσερα βασικά κλειδιά ασφαλείας τα οποία παρέχονται δωρεάν από το Twitter. Για να τα παρέχει όμως το Twitter δωρεάν σε κάποιο χρήστη αυτά τα κλειδιά χρειάζεται αρχικά να δημιουργήσει ένα Twitter\_App με τον λογαριασμό του. Αφού δημιουργήσει την συγκεκριμένη εφαρμογή μπορεί μέσα από αυτή να πάρει αυτά τα τέσσερα κλειδιά που δείχνουν την αυθεντικότητα του λογαριασμού του χρήστη. Τα τέσσερα κλειδιά είναι το Consumer Key, το Consumer Secret, το Access Token και το Access Token Secret.

The image shows a screenshot of the Twitter application settings and access token page. It is divided into two main sections: 'Application Settings' and 'Your Access Token'.

**Application Settings**  
*Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.*

Consumer Key (API Key)	YEiNO85Ech5VH07RHrh9v1cq
Consumer Secret (API Secret)	BUHbMf0XS5P5xFytqDKDKH2nzODCjUBjDG4H45aXHB8Mab4Fca
Access Level	Read and write (modify app permissions)
Owner	Stefanos_Manoli
Owner ID	3301449071

**Application Actions**

Regenerate Consumer Key and Secret    Change App Permissions

**Your Access Token**  
*This access token can be used to make API requests on your own account's behalf. Do not share your access token se...*

Access Token	3301449071- HhunUQdaJPUGi7e7eyWo9OFJB3fQvblIBAM95QU
Access Token Secret	vZ7pvLHmlwYHnflrZoJ69Fr9LdfjlL79w1OM4AV7zgs2x

Εικόνα 1 - Twitter Application

### 4.2.3. Επεξεργασία Δεδομένων

Αφού συλλέχθηκαν όλα τα δεδομένα μέσω του API που δημιουργήθηκε, έπρεπε να τύχουν επεξεργασίας για το λόγο ότι μέσα στα δεδομένα υπήρχε πολλή άχρηστη πληροφορία που αποτελούσε θόρυβο και υπήρχε αλλοίωση στο αποτέλεσμα της ομαδοποίησης των δεδομένων. Έγινε επεξεργασία των δεδομένων με στόχο την αφαίρεση όλων των άχρηστων πληροφοριών όπως ημερομηνίες, ώρες και σύμβολα με στόχο να παραμείνει η σωστή πληροφορία ώστε να μην υπάρχει μεγάλη αλλοίωση των αποτελεσμάτων. Τα τελικά δεδομένα αποθηκεύτηκαν σε αρχεία τύπου text (txt) και κάθε γραμμή περιείχε δεδομένα από μία δημοσίευση (tweet) ενός από τους καθορισμένους χρήστες του twitter. Το αρχείο που περιέχει τα τελικά δεδομένα έχει την μορφή όπως φαίνεται στην εικόνα 2.

```
,The best way to increase the PROFITABILITY of any #forex system ==> https://t.co/LdQi9DYRwT #fx https://t.co/MMivc4kFXw  
,Fed Beige Book preview - Nomura https://t.co/0WssYoPvy1 #FX #Trading  
,Over 0 #pips with easy-to-follow buy and sell #forex signals ==> https://t.co/kDNOAjvPM1 #forextrading #fx  
,FREE Webinar: Pro #FX #Trader & Trainer Josh Taylor Presents.. Create A #Forex #Trading CASH MONEY MACHINE. https://t.co/y58jgyZ9MM  
,New GBP/CHF Forex signal #Forex #fx #gbpchf https://t.co/1qcXp5z9Ep
```

Εικόνα 2 - Μορφοποίηση Δεδομένων

### 4.3. Εξαγωγή Topics

Αφού αποθηκεύτηκαν και επεξεργάστηκαν όλα τα δεδομένα χρειάστηκε να δημιουργηθεί το τελικό σύστημα, αναπτύσσοντας το με την γλώσσα προγραμματισμού Python, το οποίο θα παίρνει σαν είσοδο τα τελικά δεδομένα. Το παρόν σύστημα που δημιουργήθηκε εξάγει μια λίστα από topics τα οποία αντιστοιχούν σε μια λίστα από όρους. Το σύστημα αυτό περιέχει δύο βασικά μοντέλα ώστε να εξάγει topics, το μοντέλο NMF και το μοντέλο LDA. Επίσης παίρνει σαν είσοδο τρία ορίσματα τα οποία ορίζει ο χρήστης.

### **4.3.1. Ορίσματα του παρόν συστήματος**

- ❖ `n_features` : Είναι ο αριθμός των χαρακτηριστικών που ορίζει ο χρήστης
- ❖ `n_topics` : Είναι ο αριθμός των topics που ορίζει ο χρήστης να εξάγει κάθε μοντέλο
- ❖ `n_samples` : Είναι ο αριθμός των δειγμάτων που ορίζει ο χρήστης

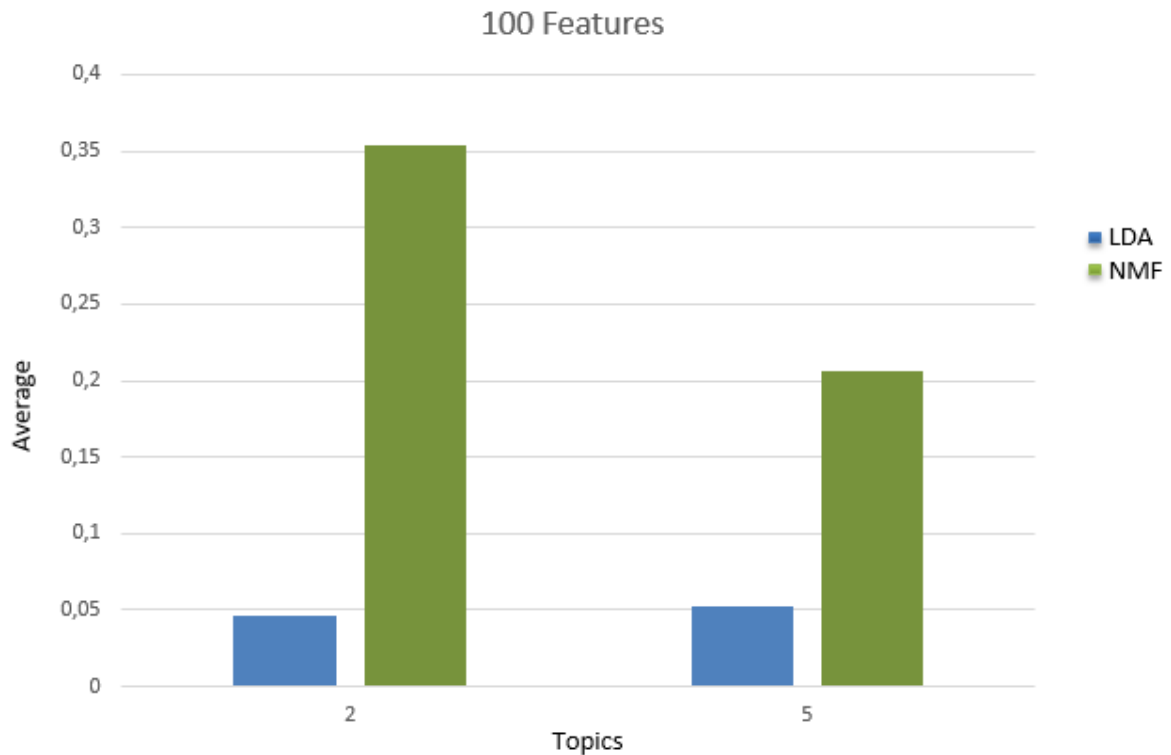
## 5. Αποτελέσματα

Η παρούσα πτυχιακή εργασία έχει ως κύριο στόχο της να εξάγει όσο καλύτερα αποτελέσματα γινόταν έτσι ώστε να εξάγει το σωστό κλίμα που επικρατεί για κάθε μέρα όσο αφορά τον οικονομικό δείκτη FX. Για να καταλήξει στο καλύτερο δυνατό αποτέλεσμα προσθέσαμε στο σύστημα την μέθοδο `silhouette_score` η οποία μας δίνει για κάθε συνδυασμό μεταξύ αριθμών features και αριθμό topics ένα αριθμό μεταξύ του 0 και του 1 για κάθε μοντέλο (NMF, LDA) ξεχωριστά. Με αυτό τον τρόπο βλέπουμε πιο μοντέλο εξάγει το καλύτερο δυνατό αποτέλεσμα για κάθε συνδυασμό. Η καλύτερη τιμή είναι το ένα και η χειρότερη το -1. Σε περίπτωση αρνητικών τιμών αυτό κατά μεγάλη πιθανότητα ευθύνεται στο γεγονός ότι κάποια δείγματα αποδόθηκαν σε λάθος ομάδα. Σε περιπτώσεις όπου αριθμός τείνει κοντά στο 0 τότε αυτό ευθύνεται στο γεγονός ότι υπάρχουν επικαλυπτόμενες ομάδες (clusters). Στην συνέχεια αφού εφαρμόστηκε η πιο πάνω μέθοδος τότε για κάθε συνδυασμό έτρεξε 100 φορές ο κώδικας γιατί κάθε φορά το αποτέλεσμα αλλάζει. Στόχος ήταν να βγει ο μέσος όρος και η διασπορά για κάθε συνδυασμό και κάθε μοντέλο ξεχωριστά και για τα 100 αποτελέσματα έτσι ώστε να καταλήξουμε στο καλύτερο συνδυασμό features και topics.

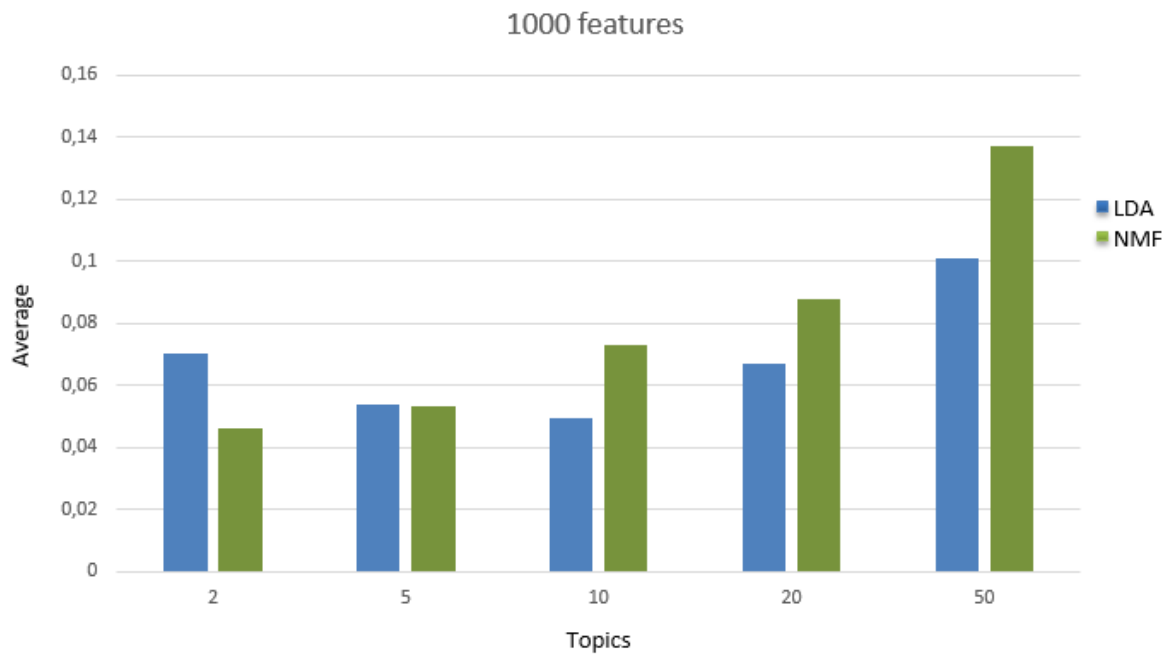


## 5.1. Γραφικές παραστάσεις με αποτελέσματα

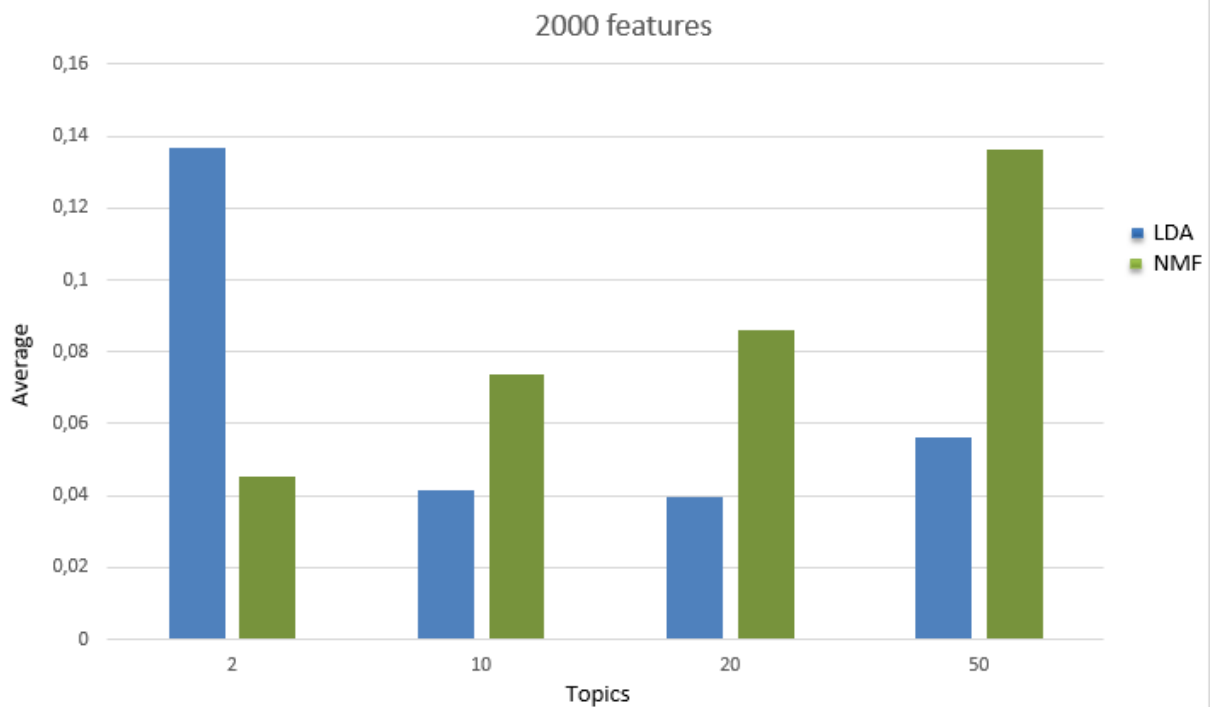
Με βάση τα αποτελέσματα που δίνονται στα παραρτήματα σε μορφή πινάκων για κάθε συνδυασμό ξεχωριστά δημιουργήθηκαν οι εξής γραφικές παραστάσεις για καλύτερη κατανόηση των αποτελεσμάτων.



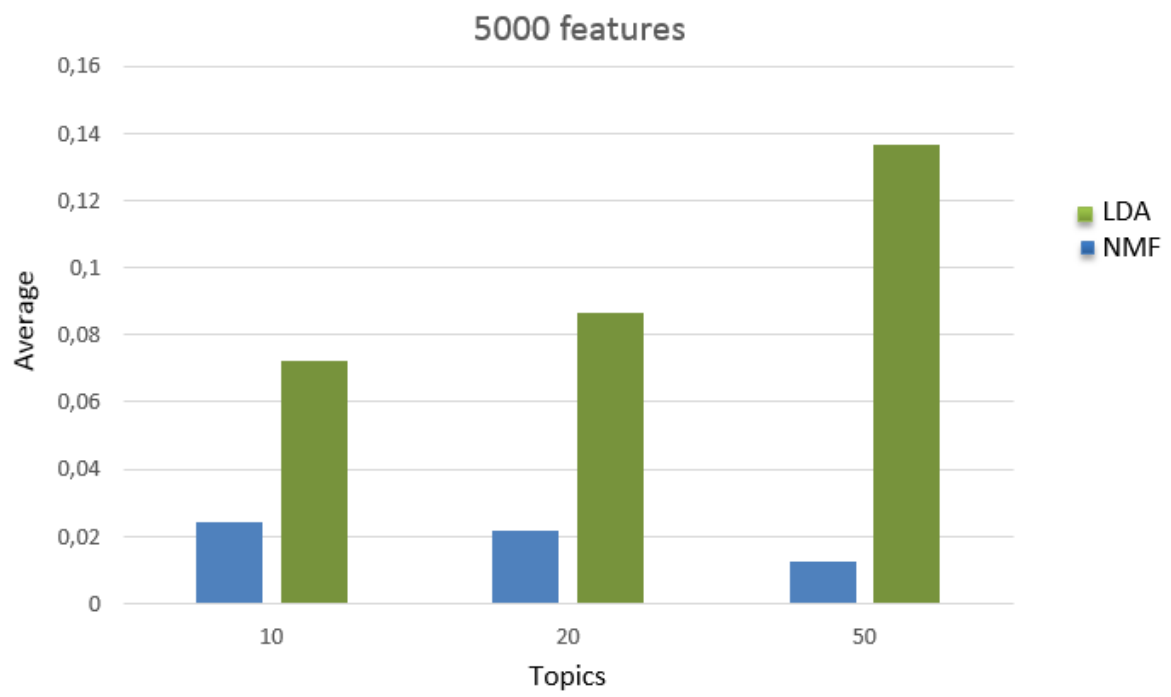
Εικόνα 3 - Σύγκριση μέσου όρου των δύο μοντέλων με βάση τα 100 features



Εικόνα 4 - Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 1000 features



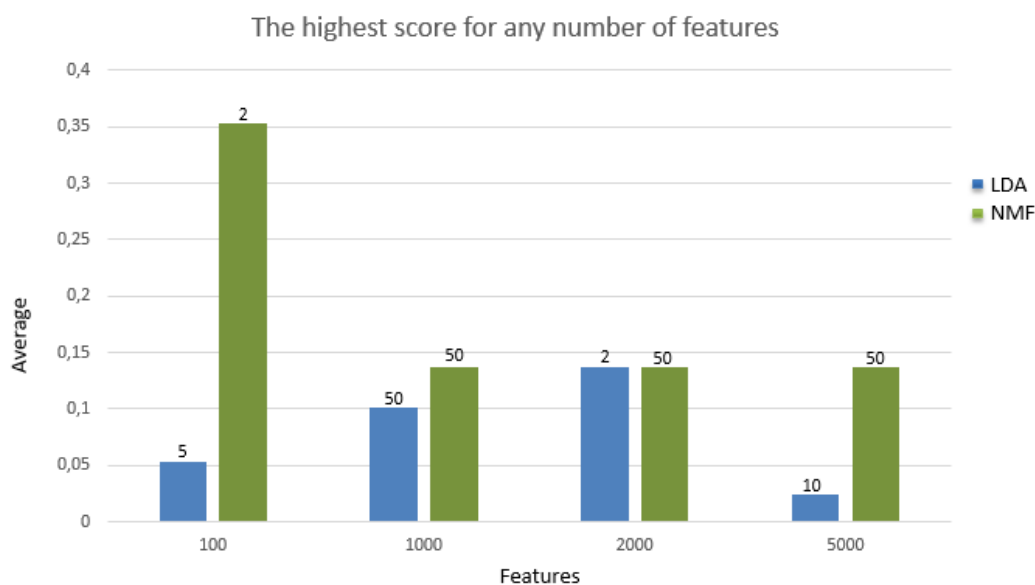
Εικόνα 5 - Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 2000 features



**Εικόνα 6 Σύγκριση μέσου όρου των δύο μοντέλων με βάσει τα 5000 features**

## 6. Συμπεράσματα

Για την εξαγωγή συμπερασμάτων συγκεντρώθηκαν, όπως φαίνεται στην εικόνα 7 πιο κάτω, οι μεγαλύτεροι μέσοι όροι από κάθε μοντέλο, για όλους τους αριθμούς χαρακτηριστικών (features). Από την μια πλευρά βάσει της γραφικής παράστασης πιο κάτω παρατηρείτε ότι όσο αφορά την ομαδοποίηση με το μοντέλο Latent Dirichlet Allocation ο καλύτερος συνδυασμός μεταξύ των αριθμών των topics και των features είναι 2 topics και 2000 features. Με τον συγκεκριμένο συνδυασμό ο μέσος όρος για τα 100 αποτελέσματα του silhouette\_score είναι ο υψηλότερος και ανέρχεται κοντά στο 0,15. Από την άλλη πλευρά όσο αφορά την ομαδοποίηση με το μοντέλο Non-negative Matrix Factorization παρατηρείται βάσει της γραφικής παράστασης ότι ο καλύτερος συνδυασμός μεταξύ των αριθμών των topics και των features είναι 2 topics και 100 features. Με τον συγκεκριμένο συνδυασμό ο μέσος όρος για τα 100 αποτελέσματα του silhouette\_score είναι ο υψηλότερος και ανέρχεται κοντά στο 0,35. Με βάσει τις πιο πάνω παρατηρήσεις καταλήγουμε στο συμπέρασμα ότι και για τα δύο μοντέλα γίνεται καλύτερη ομαδοποίηση όταν τα ομαδοποιούμε σε λίγα topics. Πιο συγκεκριμένα όταν δώσουμε στο σύστημα ο αριθμός των topics να είναι 2 και για τις δύο περιπτώσεις μοντέλων. Όσο αφορά τον αριθμό των χαρακτηριστικών με βάσει τις παρατηρήσεις πιο πάνω καταλήγουμε στο συμπέρασμα ότι όσο αφορά και τα δύο μοντέλα ομαδοποίησης NMF και LDA ο αριθμός των χαρακτηριστικών που πρέπει να δώσει ο χρήστης στο σύστημα ως όρισμα ώστε να γίνει καλύτερη ομαδοποίηση πρέπει να είναι ανάλογος του αριθμού των δεδομένων. Πιο συγκεκριμένα όσο αφορά το μοντέλο NMF όσο πιο λίγο αριθμό χαρακτηριστικών ορίσουμε στο σύστημα τόσο καλύτερη γίνεται η ομαδοποίηση των δεδομένων σε topics.



Εικόνα 7 - Σύγκριση δύο μοντέλων με βάσει τους μεγαλύτερους μέσους όρους

## BIBΛΙΟΓΡΑΦΙΑ

- [1] "2.3. Clustering — Scikit-Learn 0.17.1 Documentation". Scikit-learn.org
- [2] sklearn, silhouette. "Silhouette Coefficient In Python With Sklearn". Stackoverflow.com
- [3] "Sklearn.Decomposition.Latentdirichletallocation — Scikit-Learn Documentation". Scikit-learn.org
- [4] "Stanford Large Network Dataset Collection". Snap.stanford.edu
- [5] "Tweepy/Tweepy". GitHub
- [6] "Twitter". Twitter.com
- [7] "Twitter Application Management". Apps.twitter.com
- [8] "Twitter Developers". Dev.twitter.com
- [9] D. M. Blei, "Introduction to Probabilistic Topic Models," pp. 1–16.
- [10] N. W. Road, H. I. Storage, and R. Information, "Document Clustering Based On Non-negative Matrix Factorization," pp. 267–273.
- [11] K. Stevens, P. Kegelmeyer, D. Andrzejewski, and D. Buttler, "Exploring Topic Coherence over many models and many topics," no. July, pp. 952–961, 2012.
- [12] B. Xiang, "Improving Twitter Sentiment Analysis with Topic-Based Mixture Modeling and Semi-Supervised Training," pp. 434–439, 2014.
- [13] J. Yang, "Patterns of Temporal Variation in Online Media Categories and Subject Descriptors."

# ΠΑΡΑΡΤΗΜΑΤΑ

## API

```
import tweepy

import csv

consumer_key = 'YEiNO85Ech5VHo7RHrIh9v1cq'
consumer_secret = 'BUHbMf0XS5P5xFytqDKDKH2nzODCjUBjDG4H45aXHB8Mab4Fca'
access_token = '3301449071-HhunUQdaJPUGi7e7eyWo9OFJB3fQvblIBAM95QU'
access_token_secret = 'vZ7pvLHmIwYHnfIrZoJ69Fr9LdfjL79w1OM4AV7zgs2x'

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth)

# Open/Create a file to append data
csvFile = open('tweets.csv', 'a')

#Use csv Writer
csvWriter = csv.writer(csvFile)

for tweet in tweepy.Cursor(api.search,q="#forex",count=100,\
                             lang="en").items():
    print tweet.text
    csvWriter.writerow([tweet.text.encode('utf-8')])
```

## Μορφοποιημένος κώδικας συστήματος

```
from __future__ import print_function

from time import time

from sklearn.feature_extraction.text import TfidfVectorizer, CountVectorizer

from sklearn.decomposition import NMF, LatentDirichletAllocation

import numpy as np

#####

from sklearn import metrics

#####

#import text file with tweets

#####

x1a = np.genfromtxt('C:\Users\Stefanos\Desktop\test2.txt',delimiter=';',dtype=None)

#####

n_samples = 2000

n_features = 100

n_topics = 5

n_top_words = 20

#####

def print_top_words(model, feature_names, n_top_words):

    for topic_idx, topic in enumerate(model.components_):

        print("Topic #%d:" % topic_idx)

        print(" ".join([feature_names[i]

                        for i in topic.argsort()[:-n_top_words - 1:-1]]))

    print()
```

```

#####
print("Loading dataset...")
t0 = time()
data_samples = x1a
print("done in %0.3fs." % (time() - t0))
#####

#NMF MODEL

# Use tf-idf features for NMF.

#####

print("Extracting tf-idf features for NMF...")
tfidf_vectorizer=TfidfVectorizer(max_df=0.95, min_df=2,#max_features=n_features,
                                stop_words='english')

t0 = time()
tfidf = tfidf_vectorizer.fit_transform(data_samples)
print("done in %0.3fs." % (time() - t0))
#####

# Fit the NMF model

#####

print("Fitting the NMF model with tf-idf features,"
      "n_samples=%d and n_features=%d..."
      % (n_samples, n_features))

t0 = time()
nmf=NMF(n_components=n_topics,random_state=1, alpha=.1, l1_ratio=.5)
nmf2 = nmf.fit(tfidf)

#exit()

nmf3=nmf2.transform(tfidf)

print("done in %0.3fs." % (time() - t0))

```



```

print_top_
words(lda, tf_feature_names, n_top_words)
label2=np.argmax(tf2,axis=1)
#####
print("Silhouette Coefficient for LDA model: %0.3f"
      % metrics.silhouette_score(tf, label2, sample_size=2000))
a = 0
while a <99:
    a = a + 1
    print("%0.3f"
          % metrics.silhouette_score(tf, label2, sample_size=2000))
#####
print("Silhouette Coefficient for NMF model: %0.3f"
      % metrics.silhouette_score(tfidf, label1, sample_size=2000))
b = 0
while b < 99:
    b = b + 1
    print("%0.3f"
          % metrics.silhouette_score(tfidf, label1, sample_size=2000))

**END OF CODE**

```

## Στιγμιότυπο μορφής δεδομένων

"2016-04-1223:59:15,""Upcomingeventin30min[00:30GMT]-AustralianConsumerSentiment(Index,cons:n/a,prev:99.1)#forex#fx#finance""

2016-04-1223:58:20,Bought#GBPCHF1.36295SL1.34095|Auto-copy#tradeFREEwith#Forexviahttps://t.co/dCcc1zhyKw#forex#news#fx#ff#rt

2016-04-1223:58:09,Bought#GBPCHF1.36295SL1.34095|Auto-copy#tradeFREEwith#Forexviahttps://t.co/p87N54HZnH#forex#news#fx#ff#rt

2016-04-1223:56:43,Asscheeksareperfecthttps://t.co/csGJITd3Y#FX#Trading

2016-04-1223:40:23,Whyaren'tyouwatching??#tv#television#fxhttps://t.co/wTW8iTpgSQhttps://t.co/6cvRQLWHKT

"2016-04-1223:39:16,""Upcomingeventin10min[23:50GMT]-JapaneseM2MoneySupply(YearonYear,cons:3.1%,prev:3.1%)#forex#fx#finance""

2016-04-1223:39:12,TheULTIMATEGUIDEonhowtotradelessandmakemore:https://t.co/lx8XYsFwZf#forex#fx#forextradinghttps://t.co/mCmj1cwiJ

2016-04-1223:38:46,TOPSECRET#forex#scalpingstrategyREVEALED==&gthttps://t.co/MCLsFL1Kag#fx#forextradinghttps://t.co/Viitr3nDEQ

2016-04-1223:37:43,RT@soompi:#fx'sLunaChosenasNewMCoFCableBeautyShowhttps://t.co/l9zXXaOjDIhttps://t.co/Lsuj5vSUaA

2016-04-1223:37:33,#FXOilandCADBreakout:SignsofanoiloutputfreezepushedoilandtheCanadiandollartomulti-monthhighson...https://t.co/NHPbeJbmER

2016-04-1223:34:02,#GBPJPYhttps://t.co/5sZSzdqBmPot.cont.UptrendINTRAWEEK60m-15m-4mcombochart#forex#fxhttps://t.co/oswlvjIybX

2016-04-1223:33:36,Laseraccurateandveryfast#forexsignalsdirectlyonyourchart==&gthttps://t.co/kDNOAfvPMI#forextrading#fxhttps://t.co/lrp5jW3SZt

"2016-04-1223:32:57,""RT@OfficialMwave:TipsforPerfectSkinLikeK-PopGoddesses#SNSD's#Yoona,#fx's#Krystaland#AOA's#Seolhyunhttps://t.co/GS3m9qAFSel...""

2016-04-1223:31:59,#forex#fxFedBeigeBookpreview-Nomura:AnalystsatNomuraofferedapreviewofthe...https://t.co/35A3kd77qYhttps://t.co/eoLU2fet2C

2016-04-1223:31:55,FedBeigeBookpreview-Nomurahttps://t.co/m4H2HnaBnT#forex#fx

2016-04-1223:31:55,EUR/JPY:126.00onthecards?-FXStreethttps://t.co/1T2mohmQPC#forex#fx

2016-04-1223:31:55,Fed'sLacker:remainshawkishfor2016https://t.co/M3kjt8YKYG#forex#fx

2016-04-1223:31:55,OilstabilizingthesupplyafterbigAPIbuildhttps://t.co/kDHRMxi80x#forex#fx

2016-04-1223:31:06,#FXForexTrendy:ReduceunnecessarylossesandincreaseyourodsofwinningwithourForextrendscanner.https://t.co/9GjrL0jJ4f

2016-04-1223:30:21,MostMisunderstoodStatistictodayhttps://t.co/WKsiUyHNII#systems#soundcloud#trading#invest#money#forex#fxhttps://t.co/BkBIbCsiAo

2016-04-1223:30:21,HowtoBuildaMulti-Billion\$Firmhttps://t.co/s5bxmpdQTO#daytrading#business#soundcloud#markets#trading#invest#money#forex#fx

2016-04-1223:30:21,WhybeingSystematicWinshttps://t.co/k7MNJc3hbO#daytrading#business#soundcloud#markets#trading#invest#money#forex#fx#win#silver

2016-04-1223:30:21,NewBusiness=SmallTreehttps://t.co/CAI0136LkQ#listen#fx#markets#market#startup#investing#moneyhttps://t.co/Eh3iYXGmYj

2016-04-1223:30:21,HowtomeasureREALRISKhttps://t.co/IMuozHJA93#daytrading#cash#funds#markets#rt#fx#listen#winhttps://t.co/mN88cdHsuL

2016-04-1223:30:21,TradelikeaLegendhttps://t.co/sRXAT8xc5s#listen#podcast#stocks#market#money#trading#fx#businesshttps://t.co/QQU5jboh0s

2016-04-1223:30:20,ReturnOFCapitalismoreimportantforsome!https://t.co/njjBW7ELVQ#investing#success#money#trading#fxhttps://t.co/aeXdMv8nH9

2016-04-1223:29:36,LongandIntermediateTrendsCrushingCommercials...https://t.co/Dnggaoke4R#Forex#FX

2016-04-1223:27:48,NewEUR/JPYForexsignal#forex#fx#eurjpyhttps://t.co/UIQ1uGaG4J

2016-04-1223:27:08,TheeasiestwaytoREDUCEunnecessary#forexLOSSES==&gthttps://t.co/K1idcKmuJB#fx#currency

## Αποτελέσματα silhouette\_score για όλους τους συνδυασμούς

LDA 10-1000	NMF 10-1000	LDA 20-1000	NMF 20-1000	LDA 50-1000	NMF 50-1000	LDA 10-2000	NMF 10-2000	LDA 20-2000	NMF 20-2000
0,057	0,07	0,061	0,081	0,081	0,133	0,038	0,071	0,024	0,094
0,054	0,08	0,064	0,094	0,082	0,125	0,043	0,073	0,034	0,091
0,051	0,065	0,07	0,093	0,093	0,148	0,039	0,076	0,047	0,082
0,038	0,063	0,071	0,098	0,104	0,13	0,039	0,065	0,047	0,078
0,052	0,069	0,077	0,088	0,111	0,152	0,035	0,075	0,035	0,079
0,053	0,082	0,057	0,099	0,103	0,145	0,051	0,079	0,044	0,081
0,027	0,071	0,073	0,094	0,087	0,149	0,053	0,071	0,033	0,085
0,049	0,069	0,057	0,087	0,106	0,133	0,027	0,066	0,036	0,081
0,048	0,077	0,086	0,084	0,092	0,144	0,041	0,074	0,053	0,071
0,04	0,081	0,079	0,086	0,098	0,132	0,046	0,071	0,055	0,094
0,048	0,064	0,062	0,076	0,102	0,131	0,051	0,07	0,04	0,086
0,046	0,069	0,064	0,088	0,114	0,145	0,037	0,075	0,048	0,084
0,047	0,078	0,074	0,085	0,104	0,14	0,047	0,063	0,027	0,1
0,042	0,07	0,055	0,084	0,121	0,14	0,041	0,069	0,046	0,087
0,048	0,061	0,06	0,084	0,098	0,138	0,047	0,085	0,04	0,084
0,046	0,067	0,066	0,078	0,086	0,14	0,043	0,073	0,028	0,092
0,052	0,075	0,049	0,093	0,093	0,14	0,045	0,072	0,037	0,085
0,064	0,068	0,047	0,103	0,095	0,131	0,04	0,084	0,038	0,081
0,037	0,071	0,065	0,085	0,081	0,135	0,042	0,07	0,033	0,084
0,05	0,089	0,064	0,095	0,114	0,127	0,033	0,072	0,044	0,077
0,046	0,083	0,066	0,101	0,09	0,136	0,06	0,076	0,037	0,077
0,05	0,082	0,067	0,088	0,102	0,128	0,042	0,073	0,072	0,081
0,059	0,082	0,052	0,081	0,098	0,154	0,037	0,083	0,048	0,088
0,054	0,08	0,078	0,094	0,106	0,125	0,044	0,075	0,05	0,103
0,062	0,07	0,062	0,086	0,084	0,129	0,044	0,069	0,046	0,079
0,056	0,066	0,063	0,081	0,088	0,132	0,023	0,073	0,046	0,092
0,036	0,071	0,069	0,09	0,081	0,123	0,041	0,066	0,054	0,09
0,062	0,07	0,067	0,093	0,098	0,138	0,031	0,087	0,022	0,088
0,055	0,072	0,071	0,084	0,11	0,127	0,04	0,087	0,02	0,076
0,039	0,069	0,08	0,097	0,094	0,145	0,037	0,086	0,021	0,094
0,044	0,072	0,058	0,094	0,103	0,132	0,039	0,062	0,035	0,083
0,044	0,076	0,069	0,082	0,083	0,124	0,034	0,071	0,049	0,094
0,053	0,075	0,079	0,094	0,129	0,135	0,037	0,088	0,048	0,075
0,052	0,071	0,063	0,08	0,084	0,137	0,053	0,075	0,053	0,086
0,05	0,072	0,076	0,072	0,115	0,153	0,051	0,073	0,046	0,089
0,052	0,072	0,054	0,075	0,113	0,15	0,04	0,074	0,037	0,096
0,055	0,081	0,064	0,086	0,105	0,132	0,047	0,07	0,043	0,068
0,048	0,063	0,07	0,094	0,101	0,12	0,041	0,073	0,036	0,09
0,059	0,079	0,05	0,094	0,092	0,138	0,053	0,067	0,016	0,079
0,048	0,073	0,062	0,095	0,104	0,132	0,044	0,083	0,038	0,095
0,053	0,076	0,063	0,083	0,112	0,143	0,036	0,079	0,039	0,084
0,048	0,072	0,068	0,077	0,103	0,155	0,035	0,061	0,029	0,09
0,046	0,084	0,08	0,082	0,106	0,162	0,045	0,066	0,036	0,093
0,034	0,059	0,075	0,08	0,103	0,141	0,04	0,068	0,058	0,086
0,054	0,075	0,066	0,09	0,098	0,151	0,041	0,08	0,03	0,087
0,048	0,069	0,057	0,077	0,115	0,147	0,042	0,079	0,058	0,083
0,059	0,079	0,073	0,087	0,105	0,153	0,041	0,069	0,04	0,084
0,052	0,079	0,092	0,083	0,094	0,132	0,025	0,074	0,031	0,083
0,05	0,077	0,068	0,079	0,1	0,145	0,039	0,064	0,024	0,092
0,049	0,078	0,072	0,092	0,11	0,123	0,04	0,065	0,033	0,089

0,054	0,07	0,068	0,083	0,098	0,134	0,047	0,082	0,04	0,093
0,045	0,077	0,056	0,08	0,112	0,142	0,043	0,071	0,044	0,092
0,056	0,071	0,042	0,087	0,08	0,125	0,043	0,072	0,044	0,075
0,039	0,076	0,053	0,083	0,11	0,126	0,042	0,071	0,031	0,095
0,06	0,074	0,058	0,102	0,089	0,139	0,059	0,087	0,048	0,098
0,052	0,076	0,068	0,086	0,085	0,114	0,04	0,07	0,035	0,089
0,059	0,071	0,072	0,086	0,097	0,125	0,048	0,069	0,033	0,08
0,047	0,082	0,047	0,092	0,091	0,143	0,037	0,086	0,057	0,089
0,048	0,078	0,069	0,079	0,104	0,136	0,051	0,07	0,04	0,094
0,056	0,075	0,076	0,103	0,109	0,125	0,052	0,081	0,046	0,087
0,077	0,067	0,075	0,083	0,103	0,146	0,035	0,074	0,042	0,087
0,058	0,078	0,081	0,097	0,109	0,134	0,037	0,085	0,039	0,089
0,053	0,079	0,077	0,087	0,101	0,136	0,049	0,062	0,036	0,078
0,059	0,067	0,075	0,082	0,121	0,129	0,041	0,075	0,018	0,077
0,047	0,076	0,069	0,094	0,112	0,152	0,036	0,07	0,008	0,092
0,048	0,063	0,061	0,081	0,09	0,133	0,043	0,081	0,037	0,079
0,053	0,067	0,06	0,088	0,087	0,147	0,038	0,061	0,015	0,081
0,034	0,067	0,077	0,085	0,095	0,144	0,032	0,075	0,055	0,086
0,05	0,076	0,067	0,097	0,103	0,133	0,05	0,078	0,056	0,082
0,054	0,067	0,078	0,099	0,11	0,124	0,034	0,074	0,043	0,088
0,057	0,07	0,064	0,077	0,106	0,14	0,049	0,084	0,041	0,077
0,038	0,065	0,047	0,086	0,112	0,145	0,031	0,076	0,028	0,086
0,061	0,07	0,072	0,084	0,121	0,149	0,048	0,076	0,054	0,09
0,047	0,067	0,078	0,094	0,084	0,133	0,043	0,06	0,039	0,079
0,037	0,075	0,07	0,076	0,093	0,125	0,047	0,074	0,037	0,093
0,064	0,082	0,078	0,086	0,097	0,147	0,049	0,074	0,056	0,087
0,068	0,076	0,074	0,096	0,106	0,131	0,044	0,077	0,036	0,086
0,057	0,061	0,061	0,09	0,106	0,138	0,036	0,079	0,03	0,093
0,045	0,071	0,07	0,093	0,117	0,139	0,04	0,076	0,055	0,08
0,046	0,073	0,064	0,094	0,114	0,133	0,024	0,081	0,035	0,086
0,05	0,065	0,064	0,087	0,112	0,115	0,052	0,07	0,043	0,077
0,052	0,08	0,048	0,083	0,11	0,14	0,039	0,075	0,051	0,082
0,03	0,077	0,064	0,087	0,091	0,15	0,057	0,071	0,049	0,079
0,041	0,072	0,064	0,075	0,098	0,145	0,036	0,061	0,044	0,094
0,055	0,078	0,081	0,073	0,106	0,135	0,035	0,088	0,024	0,086
0,039	0,065	0,065	0,093	0,087	0,145	0,031	0,075	0,03	0,107
0,047	0,066	0,065	0,102	0,095	0,131	0,042	0,068	0,055	0,089
0,042	0,076	0,075	0,088	0,118	0,14	0,042	0,073	0,043	0,079
0,041	0,072	0,078	0,079	0,111	0,13	0,041	0,065	0,024	0,083
0,05	0,072	0,075	0,096	0,096	0,129	0,032	0,074	0,054	0,087
0,063	0,067	0,095	0,083	0,094	0,142	0,046	0,076	0,047	0,079
0,045	0,077	0,081	0,09	0,121	0,141	0,029	0,096	0,055	0,083
0,037	0,081	0,057	0,086	0,104	0,134	0,049	0,079	0,021	0,085
0,056	0,066	0,061	0,101	0,088	0,146	0,023	0,078	0,025	0,084
0,042	0,07	0,059	0,086	0,098	0,144	0,029	0,069	0,051	0,086
0,036	0,076	0,054	0,079	0,108	0,139	0,05	0,072	0,031	0,095
0,045	0,076	0,058	0,083	0,102	0,12	0,038	0,068	0,034	0,086
0,037	0,08	0,07	0,091	0,108	0,128	0,045	0,071	0,041	0,085
0,038	0,075	0,061	0,092	0,109	0,15	0,041	0,07	0,029	0,097
0,04	0,072	0,056	0,08	0,092	0,138	0,05	0,074	0,041	0,085
0,04921	0,07296	0,06673	0,0875	0,10101	0,13709	0,04135	0,07394	0,03954	0,08606
7,24259E-05	3,50384E-05	9,85371E-05	0,00005327	0,00011757	8,9922E-05	5,639E-05	4,8716E-05	0,0001286	4,58964E-05

LDA 50-2000	NMF 50-2000	LDA 10-5000	NMF 10-5000	LDA 20-5000	NMF 20-5000	LDA 50-5000	NMF 50-5000	LDA 5-1000	NMF 5-1000
0,042	0,179	0,024	0,076	0,036	0,086	0,025	0,132	0,054	0,049
0,062	0,134	0,015	0,08	0,015	0,093	-0,01	0,13	0,052	0,039
0,057	0,135	0,019	0,077	0,028	0,067	0,011	0,123	0,049	0,059
0,062	0,141	0,025	0,063	0,016	0,083	0,013	0,129	0,054	0,055
0,048	0,123	0,026	0,075	0,029	0,085	-0,019	0,149	0,054	0,045
0,055	0,151	0,031	0,067	0,017	0,102	0,016	0,146	0,047	0,051
0,054	0,115	0,02	0,076	0,03	0,084	-0,091	0,142	0,06	0,049
0,064	0,133	0,016	0,068	0,025	0,086	0,018	0,135	0,043	0,049
0,067	0,138	0,017	0,064	0,021	0,085	0,026	0,137	0,054	0,05
0,045	0,125	0,023	0,07	0,017	0,094	0,032	0,137	0,046	0,055
0,056	0,13	0,031	0,072	0,016	0,091	-0,001	0,141	0,063	0,062
0,04	0,132	0,024	0,081	0,04	0,089	0,019	0,129	0,051	0,047
0,062	0,144	0,016	0,076	0,024	0,091	-0,005	0,118	0,049	0,049
0,056	0,141	0,02	0,078	0,02	0,078	-0,001	0,139	0,055	0,049
0,053	0,131	0,022	0,07	0,023	0,098	-0,084	0,123	0,063	0,043
0,044	0,137	0,025	0,077	0,034	0,079	0,005	0,14	0,05	0,052
0,063	0,137	0,014	0,081	0,023	0,083	-0,007	0,142	0,055	0,048
0,074	0,139	0,027	0,071	0,043	0,09	-0,007	0,143	0,056	0,063
0,061	0,144	0,03	0,064	0,014	0,095	0,016	0,142	0,053	0,054
0,068	0,146	0,022	0,071	0,016	0,082	0,011	0,134	0,061	0,062
0,048	0,128	0,024	0,07	0,033	0,072	0,059	0,141	0,047	0,048
0,065	0,129	0,027	0,076	0,025	0,1	0,018	0,131	0,052	0,057
0,06	0,148	0,023	0,073	0,027	0,093	0,021	0,14	0,061	0,059
0,079	0,148	0,016	0,083	0,028	0,08	0,05	0,148	0,049	0,051
0,07	0,134	0,024	0,065	0,015	0,081	0,032	0,132	0,05	0,059
0,046	0,142	0,032	0,074	0,013	0,096	0,036	0,125	0,035	0,048
0,055	0,123	0,026	0,07	0,035	0,078	0,018	0,15	0,057	0,049
0,049	0,124	0,019	0,072	0,02	0,082	0,04	0,132	0,057	0,062
0,06	0,138	0,02	0,073	0,024	0,091	0,027	0,133	0,067	0,053
0,05	0,135	0,017	0,064	0,032	0,099	-0,008	0,135	0,06	0,037
0,071	0,127	0,021	0,081	0,038	0,084	0,005	0,138	0,067	0,055
0,07	0,132	0,032	0,078	0,046	0,072	-0,078	0,135	0,048	0,062
0,044	0,131	0,026	0,067	0,022	0,082	0,024	0,14	0,056	0,057
0,054	0,149	0,018	0,071	0,029	0,08	0,019	0,134	0,053	0,056
0,073	0,132	0,027	0,06	0,026	0,092	0,021	0,128	0,052	0,058
0,068	0,12	0,026	0,083	0,016	0,089	-0,007	0,139	0,061	0,052
0,034	0,142	0,031	0,07	0,033	0,089	-0,006	0,15	0,042	0,051
0,046	0,131	0,03	0,079	0,025	0,093	0,036	0,144	0,045	0,055
0,052	0,141	0,027	0,079	0,035	0,089	0,037	0,124	0,049	0,046
0,069	0,114	0,02	0,077	0,017	0,094	0,044	0,151	0,061	0,054
0,055	0,143	0,011	0,078	0,016	0,08	0,041	0,14	0,039	0,06
0,057	0,138	0,027	0,074	0,01	0,082	0,018	0,14	0,059	0,052
0,054	0,141	0,03	0,075	0,029	0,097	0,017	0,143	0,057	0,042
0,045	0,147	0,012	0,07	0,013	0,087	-0,039	0,133	0,063	0,054
0,052	0,133	0,027	0,066	0,003	0,088	0,038	0,136	0,049	0,055
0,06	0,142	0,023	0,075	0,01	0,084	0,016	0,146	0,062	0,051
0,06	0,132	0,027	0,073	0,031	0,076	0,033	0,136	0,041	0,056
0,057	0,145	0,025	0,074	0,019	0,084	-0,006	0,136	0,055	0,055
0,062	0,134	0,03	0,071	0,004	0,089	0,02	0,131	0,052	0,059

0,031	0,146	0,025	0,072	0,015	0,092	0,05	0,141	0,056	0,053
0,048	0,143	0,021	0,085	0,022	0,084	0,029	0,138	0,067	0,063
0,047	0,127	0,04	0,074	0,026	0,097	0,039	0,157	0,044	0,052
0,055	0,146	0,031	0,073	0,015	0,092	-0,003	0,133	0,041	0,052
0,06	0,151	0,012	0,079	0,03	0,094	-0,009	0,136	0,05	0,058
0,068	0,134	0,015	0,064	0,023	0,067	-0,021	0,129	0,035	0,059
0,049	0,142	0,018	0,07	0,036	0,107	-0,007	0,14	0,055	0,054
0,062	0,141	0,02	0,064	0,03	0,081	0,026	0,139	0,049	0,06
0,082	0,131	0,03	0,069	0,02	0,082	0,002	0,145	0,045	0,059
0,06	0,138	0,018	0,08	0,008	0,104	0,035	0,141	0,065	0,061
0,056	0,122	0,023	0,065	0,019	0,095	0,033	0,139	0,067	0,055
0,072	0,134	0,026	0,071	0,023	0,092	-0,009	0,117	0,058	0,065
0,053	0,148	0,01	0,076	0,014	0,095	0,029	0,138	0,045	0,049
0,038	0,141	0,028	0,083	0,012	0,08	0,047	0,139	0,065	0,053
0,05	0,137	0,023	0,065	0,009	0,086	0,03	0,132	0,056	0,046
0,059	0,142	0,026	0,069	0,018	0,09	0,011	0,156	0,037	0,054
0,04	0,135	0,034	0,066	0,019	0,082	0,001	0,127	0,047	0,051
0,044	0,153	0,016	0,079	0,027	0,098	0,023	0,132	0,05	0,047
0,071	0,142	0,031	0,079	0,005	0,082	0,03	0,133	0,037	0,042
0,078	0,14	0,021	0,073	0,034	0,095	0,05	0,131	0,033	0,053
0,054	0,145	0,04	0,073	0,008	0,085	0,044	0,137	0,059	0,061
0,07	0,129	0,029	0,074	0,014	0,082	0,032	0,136	0,046	0,048
0,055	0,137	0,01	0,07	0,028	0,09	-0,008	0,137	0,053	0,05
0,048	0,138	0,025	0,073	0,028	0,079	0,021	0,142	0,055	0,048
0,071	0,13	0,019	0,059	-0,001	0,085	-0,036	0,134	0,054	0,044
0,064	0,151	0,03	0,065	0,012	0,09	0,044	0,128	0,055	0,055
0,059	0,131	0,032	0,077	-0,003	0,089	0,026	0,126	0,047	0,063
0,063	0,125	0,028	0,074	0,027	0,074	-0,069	0,139	0,057	0,054
0,043	0,128	0,036	0,066	0,025	0,083	-0,008	0,144	0,064	0,055
0,053	0,137	0,022	0,077	0,006	0,098	0	0,13	0,059	0,054
0,049	0,136	0,018	0,067	0,014	0,069	0,03	0,127	0,049	0,055
0,056	0,117	0,019	0,065	0,021	0,082	0,018	0,13	0,045	0,056
0,084	0,15	0,038	0,067	0,018	0,083	0,018	0,127	0,059	0,058
0,054	0,129	0,024	0,061	0,013	0,076	0,037	0,152	0,047	0,046
0,064	0,135	0,02	0,067	0,019	0,084	0,033	0,13	0,056	0,052
0,034	0,143	0,025	0,071	0,028	0,078	0,023	0,123	0,055	0,056
0,069	0,117	0,014	0,062	0,023	0,069	0,026	0,118	0,071	0,071
0,038	0,14	0,019	0,071	0,016	0,076	0,03	0,139	0,056	0,044
0,054	0,13	0,03	0,062	0,016	0,086	0,033	0,131	0,057	0,054
0,066	0,134	0,033	0,076	0,014	0,091	0,04	0,136	0,058	0,044
0,05	0,139	0,025	0,073	0,03	0,075	0,03	0,137	0,062	0,051
0,072	0,139	0,031	0,072	0,01	0,101	-0,007	0,141	0,056	0,05
0,062	0,132	0,024	0,077	0,016	0,083	-0,006	0,136	0,059	0,054
0,053	0,138	0,034	0,078	0,032	0,078	0,017	0,134	0,06	0,05
0,052	0,143	0,024	0,074	0,018	0,107	-0,03	0,134	0,065	0,061
0,046	0,126	0,035	0,075	0,016	0,088	-0,029	0,148	0,056	0,064
0,037	0,13	0,022	0,068	0,033	0,081	0,045	0,142	0,05	0,053
0,054	0,136	0,034	0,071	0,024	0,097	0,036	0,135	0,074	0,066
0,051	0,146	0,031	0,077	0,037	0,086	-0,063	0,143	0,058	0,048
0,038	0,122	0,031	0,072	0,008	0,077	0,028	0,15	0,056	0,05
0,0564	0,13636	0,02434	0,07209	0,02141	0,08665	0,01257	0,13647	0,05371	0,0534
0,0001208	9,02504E-05	4,33244E-05	3,3582E-05	9,23E-05	7,3908E-05	0,0008421	6,1649E-05	6,445E-05	0,0000373

LDA 2-1000	NMF 2-1000	LDA 2-2000	NMF 2-2000	LDA 2-100	NMF 2-100	LDA 2-50	NMF 2-50	LDA 5-100	NMF 5-100
0,072	0,045	0,146	0,046	0,052	0,355	0,04	0,216	0,042	0,205
0,063	0,049	0,135	0,046	0,047	0,351	0,048	0,204	0,057	0,182
0,067	0,043	0,125	0,047	0,045	0,343	0,047	0,208	0,053	0,217
0,081	0,053	0,133	0,038	0,044	0,378	0,041	0,212	0,06	0,188
0,059	0,046	0,14	0,047	0,052	0,355	0,045	0,218	0,048	0,202
0,076	0,047	0,128	0,048	0,043	0,352	0,056	0,215	0,054	0,211
0,073	0,057	0,147	0,047	0,046	0,349	0,038	0,212	0,038	0,196
0,069	0,039	0,137	0,047	0,043	0,328	0,044	0,192	0,06	0,222
0,069	0,047	0,138	0,039	0,05	0,348	0,045	0,215	0,048	0,222
0,068	0,046	0,126	0,049	0,052	0,342	0,047	0,224	0,059	0,199
0,067	0,038	0,131	0,044	0,042	0,366	0,033	0,201	0,05	0,204
0,07	0,047	0,136	0,055	0,04	0,365	0,047	0,206	0,056	0,214
0,067	0,049	0,149	0,05	0,044	0,361	0,047	0,222	0,062	0,204
0,064	0,055	0,147	0,039	0,056	0,368	0,049	0,205	0,058	0,208
0,068	0,053	0,128	0,039	0,045	0,346	0,04	0,203	0,05	0,223
0,071	0,04	0,135	0,045	0,046	0,33	0,049	0,225	0,053	0,183
0,079	0,049	0,147	0,045	0,05	0,354	0,04	0,229	0,06	0,182
0,077	0,049	0,134	0,035	0,046	0,358	0,05	0,213	0,061	0,205
0,072	0,051	0,143	0,046	0,049	0,349	0,052	0,219	0,05	0,209
0,074	0,043	0,143	0,043	0,046	0,347	0,042	0,21	0,058	0,215
0,071	0,041	0,132	0,043	0,048	0,344	0,045	0,224	0,046	0,225
0,066	0,051	0,139	0,042	0,034	0,358	0,038	0,207	0,048	0,222
0,067	0,056	0,142	0,05	0,046	0,354	0,053	0,208	0,057	0,204
0,084	0,049	0,152	0,042	0,04	0,363	0,052	0,204	0,058	0,215
0,062	0,048	0,144	0,044	0,052	0,335	0,042	0,196	0,054	0,214
0,076	0,041	0,132	0,049	0,052	0,361	0,045	0,217	0,048	0,213
0,065	0,044	0,166	0,043	0,05	0,362	0,042	0,224	0,051	0,22
0,065	0,051	0,143	0,034	0,038	0,379	0,043	0,222	0,053	0,205
0,071	0,048	0,13	0,044	0,051	0,351	0,046	0,218	0,055	0,203
0,068	0,045	0,143	0,043	0,055	0,353	0,05	0,219	0,048	0,206
0,072	0,034	0,132	0,04	0,04	0,356	0,046	0,217	0,059	0,221
0,076	0,055	0,142	0,044	0,053	0,343	0,049	0,212	0,044	0,2
0,066	0,046	0,143	0,054	0,045	0,362	0,049	0,208	0,049	0,18
0,064	0,032	0,137	0,046	0,047	0,362	0,049	0,223	0,052	0,213
0,067	0,044	0,121	0,048	0,039	0,35	0,047	0,217	0,057	0,199
0,078	0,047	0,14	0,053	0,056	0,353	0,048	0,212	0,049	0,193
0,071	0,039	0,131	0,043	0,046	0,363	0,048	0,205	0,055	0,196
0,076	0,053	0,134	0,055	0,044	0,349	0,046	0,215	0,048	0,19
0,078	0,053	0,131	0,043	0,042	0,336	0,045	0,225	0,055	0,176
0,077	0,037	0,136	0,048	0,044	0,335	0,043	0,21	0,055	0,202
0,075	0,054	0,134	0,04	0,052	0,371	0,037	0,227	0,061	0,211
0,068	0,045	0,147	0,044	0,046	0,36	0,049	0,207	0,05	0,223
0,081	0,042	0,136	0,049	0,048	0,356	0,035	0,21	0,062	0,208
0,059	0,036	0,156	0,047	0,05	0,339	0,047	0,224	0,057	0,222
0,076	0,052	0,134	0,045	0,043	0,336	0,036	0,216	0,058	0,213
0,068	0,045	0,131	0,053	0,042	0,345	0,04	0,227	0,06	0,192
0,063	0,043	0,137	0,047	0,047	0,36	0,04	0,212	0,059	0,222
0,063	0,046	0,142	0,053	0,049	0,349	0,046	0,213	0,051	0,22
0,067	0,044	0,123	0,038	0,048	0,346	0,052	0,215	0,054	0,215

0,061	0,051	0,125	0,059	0,044	0,375	0,045	0,217	0,05	0,207
0,074	0,053	0,124	0,054	0,048	0,359	0,045	0,21	0,055	0,201
0,069	0,044	0,142	0,04	0,045	0,333	0,049	0,221	0,054	0,216
0,084	0,04	0,144	0,045	0,045	0,359	0,044	0,218	0,055	0,201
0,07	0,046	0,137	0,04	0,046	0,351	0,059	0,233	0,055	0,201
0,067	0,048	0,159	0,051	0,054	0,344	0,042	0,225	0,052	0,207
0,066	0,049	0,139	0,044	0,042	0,359	0,05	0,22	0,067	0,206
0,072	0,057	0,143	0,036	0,049	0,364	0,046	0,217	0,051	0,195
0,067	0,042	0,131	0,043	0,047	0,359	0,047	0,228	0,065	0,197
0,067	0,042	0,134	0,044	0,048	0,361	0,057	0,217	0,047	0,205
0,069	0,043	0,143	0,036	0,037	0,348	0,047	0,208	0,055	0,223
0,068	0,05	0,134	0,059	0,045	0,339	0,04	0,221	0,046	0,214
0,076	0,054	0,131	0,045	0,048	0,351	0,053	0,214	0,047	0,195
0,071	0,046	0,124	0,045	0,04	0,363	0,048	0,193	0,055	0,222
0,071	0,043	0,128	0,043	0,045	0,344	0,044	0,212	0,056	0,213
0,069	0,042	0,137	0,043	0,051	0,367	0,049	0,211	0,055	0,201
0,057	0,047	0,133	0,047	0,042	0,353	0,053	0,224	0,043	0,22
0,086	0,042	0,137	0,035	0,043	0,355	0,052	0,216	0,056	0,215
0,069	0,049	0,127	0,048	0,042	0,364	0,047	0,215	0,052	0,226
0,072	0,055	0,126	0,034	0,041	0,328	0,049	0,205	0,053	0,201
0,071	0,046	0,118	0,039	0,063	0,371	0,055	0,227	0,048	0,21
0,069	0,053	0,151	0,039	0,053	0,33	0,044	0,229	0,055	0,195
0,069	0,049	0,14	0,036	0,046	0,345	0,05	0,206	0,056	0,196
0,074	0,047	0,131	0,044	0,048	0,373	0,046	0,209	0,056	0,19
0,082	0,047	0,127	0,042	0,043	0,339	0,037	0,205	0,053	0,201
0,075	0,052	0,144	0,04	0,038	0,377	0,05	0,206	0,055	0,214
0,07	0,04	0,137	0,053	0,038	0,377	0,044	0,206	0,047	0,213
0,076	0,043	0,12	0,048	0,05	0,35	0,048	0,22	0,047	0,202
0,072	0,045	0,134	0,047	0,046	0,363	0,052	0,218	0,048	0,211
0,073	0,044	0,147	0,04	0,041	0,35	0,044	0,215	0,048	0,213
0,058	0,044	0,135	0,044	0,046	0,374	0,045	0,217	0,05	0,209
0,071	0,05	0,147	0,056	0,051	0,354	0,033	0,22	0,05	0,19
0,066	0,035	0,132	0,036	0,063	0,345	0,042	0,209	0,046	0,207
0,081	0,05	0,135	0,052	0,046	0,35	0,042	0,224	0,043	0,21
0,083	0,052	0,121	0,044	0,043	0,343	0,052	0,217	0,047	0,183
0,077	0,046	0,132	0,052	0,052	0,341	0,051	0,222	0,053	0,22
0,064	0,041	0,148	0,047	0,047	0,377	0,061	0,219	0,044	0,205
0,07	0,033	0,117	0,04	0,058	0,345	0,045	0,219	0,049	0,205
0,069	0,044	0,126	0,056	0,046	0,333	0,052	0,202	0,058	0,183
0,073	0,046	0,142	0,038	0,047	0,363	0,042	0,224	0,058	0,217
0,071	0,043	0,137	0,045	0,04	0,35	0,045	0,228	0,051	0,202
0,068	0,05	0,146	0,049	0,05	0,345	0,043	0,216	0,054	0,211
0,056	0,053	0,115	0,046	0,047	0,35	0,036	0,21	0,048	0,2
0,06	0,038	0,148	0,046	0,042	0,345	0,045	0,226	0,05	0,225
0,077	0,051	0,142	0,051	0,046	0,361	0,043	0,205	0,052	0,211
0,068	0,047	0,129	0,046	0,05	0,378	0,051	0,225	0,053	0,205
0,068	0,051	0,151	0,055	0,04	0,369	0,05	0,211	0,058	0,207
0,075	0,043	0,134	0,053	0,046	0,34	0,046	0,219	0,054	0,205
0,067	0,045	0,139	0,039	0,051	0,38	0,056	0,215	0,054	0,209
0,072	0,045	0,144	0,048	0,036	0,355	0,049	0,217	0,045	0,199

0,07044	0,04627	0,13663	0,04532	0,04654	0,35363	0,04617	0,2152	0,05295	0,20625
3,65E-05	2,8737E-05	8,36E-05	3,2E-05	2,75E-05	0,000157	2,83E-05	0,0000677	2,93E-05	0,00013



