



Τεχνολογικό  
Πανεπιστήμιο  
Κύπρου

Σχολή Επικοινωνίας και  
Μέσων Ενημέρωσης

Πτυχιακή εργασία

**ΕΞΟΡΥΞΗ ΓΝΩΣΗΣ ΣΕ ΣΧΟΛΙΑ ΤΟΥ FACEBOOK:  
ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ SAMSUNG**

Επιφάνειος Χατζηευαγγέλου

Λεμεσός, Μάιος 2018



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

Πτυχιακή εργασία

ΕΞΟΡΥΞΗ ΓΝΩΣΗΣ ΣΕ ΣΧΟΛΙΑ ΤΟΥ FACEBOOK: ΜΕΛΕΤΗ  
ΠΕΡΙΠΤΩΣΗΣ SAMSUNG

του

Επιφάνειου Χατζηευαγγέλου

Επιβλέπων Καθηγητής  
Δρ. Νικόλας Τσαπατσούλης

Λεμεσός, Μάιος 2018

## **Πνευματικά δικαιώματα**

Copyright © Επιφάνειος Χατζηευαγγέλου, 2018

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή μου κ. Νικόλα Τσαπατσούλη για την πολύτιμη βοήθεια του στη συγγραφή της πτυχιακής μου Εργασίας. Επίσης θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για τη στήριξη και τη συμπαράσταση τους καθ' όλη την πορεία της συγγραφής.

## ΠΕΡΙΛΗΨΗ

Τα σχόλια του Facebook αποτελούν ένα νέο τρόπο για την ανάλυση συναισθήματος. Το API Facebook μας επιτρέπει στο να μαζέψουμε όση ποσότητα σχολίων επιθυμούμε για να τα χρησιμοποιήσουμε για σκοπούς ανάλυσης συμπεριλαμβανομένων και της συναισθηματικής. Η ανάλυση μπορεί να γίνεται για σκοπούς θεμάτων όπως Μάρκετινγκ, εκλογές και γενικά στάσεις ανθρώπων για κάποιο συγκεκριμένο θέμα. Τα σχόλια στο Facebook έχουν κάποιες ιδιαιτερότητες και διαφορές από ένα συνηθισμένο κείμενο τα οποία μπορούν να χρησιμοποιηθούν για πρωτόγνωρες ταξινομήσεις κειμένων. Μερικές από αυτές τις ιδιαιτερότητες είναι τα emoticon στα οποία κάνουν πιο ευδιάκριτο το συναίσθημα του σχολιαστή. Επιπλέον τα σχόλια είναι ανεπίσημα και αδόμητα και δεν υπολογίζονται οι ορθογραφία, η γραμματική και η σύνταξη. Από όλα τα χαρακτηριστικά των σχολίων κανένα δεν κρίθηκε κατάλληλο και αποδοτικό για να δημιουργηθεί ένα μοντέλο ταξινόμησης. Στην παρούσα μελέτη θα γίνει ταξινόμηση σε σχόλια του Facebook που έχουν θέμα τις απόψεις των χρηστών για προϊόντα της Samsung και θα χρησιμοποιηθούν διάφορα χαρακτηριστικά για την ταξινόμηση τους από τους αλγόριθμους. Μερικά από αυτά τα χαρακτηριστικά που θα χρησιμοποιηθούν για την ταξινόμηση των σχολίων είναι το λεξικό συναισθημάτων, η διανυσματική αναπαράσταση λέξεων και τα n-grams. Για όλα τα χαρακτηριστικά που θα χρησιμοποιηθούν θα γίνει έλεγχος των επιδόσεων των ταξινομητών.

**Λέξεις – Κλειδιά:** Ταξινόμηση κειμένου, σχολιασμός χαρακτηριστικών σχολίων, ανάλυση συναισθήματος, μηχανική μάθηση.

## ABSTRACT

Facebook's comments are a new way to sentiment analysis. The Facebook API allows us to collect the amount of feedback we want to use for analysis, including emotional. The analysis can be done for purposes such as Marketing, elections and general attitudes of people on a particular subject. Comments on Facebook have some peculiarities and differences from a common text that can be used for unprecedented text classifications. Some of these peculiarities are the emoticons that make the commentator's feel more distinct. In addition, comments are informal and unstructured, and spelling, grammar and syntax are not counted. From all of the commentary features, it was not considered appropriate and efficient to create a classification model. This study will be categorized Facebook users' views about Samsung products and will use different attributes to classify them by algorithms. Some of these properties that will be used to classify comments are the emotion dictionary, the vector representation of words and n-grams. For all the features to be used, the performance of the classifiers will be checked.

**Keywords:** Text classification, commentation of featured comments, sentiment analysis, machine learning.