



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ & ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ & ΜΕΣΩΝ ΜΑΖΙΚΗΣ ΕΝΗΜΕΡΩΣΗΣ**

**ΠΜΣ «ΨΗΦΙΑΚΑ ΜΕΣΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΠΕΡΙΒΑΛΛΟΝΤΑ
ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ»**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Εύρεση Κανόνων Συγγραφής Δημοσιογραφικών Άρθρων
στο Διαδίκτυο με Στόχο την Υψηλή Δημοτικότητα: Η Χρήση
Αλγορίθμων Μηχανικής Μάθησης**

Δημήτριος-Αλέξανδρος Κ. Σίνης

ΑΘΗΝΑ

ΦΕΒΡΟΥΑΡΙΟΣ 2020

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Εύρεση Κανόνων Συγγραφής Δημοσιογραφικών Άρθρων στο Διαδίκτυο με
Στόχο την Υψηλή Δημοτικότητα: Η Χρήση Αλγορίθμων Μηχανικής Μάθησης

Δημήτριος-Αλέξανδρος Κ. Σίνης

A.M.: 9983201838401

ΕΠΙΒΛΕΠΩΝ: **Κωνσταντίνος Μουρλάς, Αναπληρωτής Καθηγητής**

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία μελετήθηκε η δημοτικότητα των δημοσιογραφικών ιστοριών που δημοσιεύονται στο διαδίκτυο. Πιο συγκεκριμένα, κατασκευάστηκαν κανόνες παραγωγής δημοσιογραφικών άρθρων, τα οποία έχουν επανειλημμένη πιθανότητα να αποκτήσουν υψηλή δημοτικότητα στην διαδικτυακή πλατφόρμα που αναρτώνται.

Μια ευρεία γκάμα χαρακτηριστικών, που προέρχονται από τους τομείς της δημοσιογραφίας, του μάρκετινγκ, της επικοινωνίας καθώς και των μέσων κοινωνικής δικτύωσης, ερευνήθηκε με σκοπό την αποκάλυψη εκείνων που συμβάλλουν στην υψηλή δημοτικότητα των άρθρων καθώς και του τρόπου με τον οποίο, αυτά συνδυάζονται και συνθέτουν τους κανόνες επιτυχίας.

Η έρευνα πραγματοποιήθηκε με χρήση καινοτόμων πρακτικών εξόρυξης πληροφορίας από κείμενα, ανάλυσης φυσικής γλώσσας καθώς και αλγορίθμων Μηχανικής Μάθησης. Για την εφαρμογή αυτών των πρακτικών, χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python καθώς και μια πληθώρα βιβλιοθηκών που παρέχονται από τη γλώσσα αυτή.

Τα αποτελέσματα έδειξαν ότι υπάρχουν συγκεκριμένοι συνδυασμοί χαρακτηριστικών, τα οποία όταν λαμβάνουν συγκεκριμένες τιμές, είναι δυνατόν να προσδώσουν αυξημένη πιθανότητα συγκέντρωσης υψηλής δημοτικότητας στα άρθρα που τα διαθέτουν .

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Διαδικτυακή Δημοσιογραφία

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Μέσα Κοινωνικής Δικτύωσης, Πρόβλεψη Δημοτικότητας, Κανόνες Συγγραφής Άρθρων, Ποιότητα, Μηχανική Μάθηση

ABSTRACT

The present research studied the popularity of online news articles. More specifically, was created a set of rules, that they can produce news stories with high probability of gaining popularity on the online platform they have posted.

A wide range of news articles features, that came from journalism, marketing, communication and social media fields was investigated in order to uncover them that they lead in high popularity.

The research carried out with the help of innovative practices, such as text mining, natural language processing and Machine Learning algorithms. To implement these practices, was used the Python programming language as well as a multitude of libraries provided by this language.

The results have shown that there are specific combinations of attributes which when obtain specific values, they may give an increased chance of garnering high popularity to the articles that have them.

SUBJECT AREA: Online Journalism

KEYWORDS: Social Media, Popularity Prediction, Rules of Article Writing, Quality, Machine Learning

Στη μνήμη του πατέρα μου, Κωνσταντίνου.

ΕΥΧΑΡΙΣΤΙΕΣ

Για τη διεκπεραίωση της παρούσας Διπλωματικής Εργασίας, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα, κ. Κωνσταντίνο Μουρλά καθώς και την υποψήφια διδάκτορα κ. Αικατερίνη-Αλεξάνδρα Σωτηράκου, για τη συνεργασία και την πολύτιμη συμβολή τους στην ολοκλήρωσή της.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ.....	11
1. ΕΙΣΑΓΩΓΗ	12
2. ΔΙΑΔΙΚΤΥΑΚΗ ΔΗΜΟΣΙΟΓΡΑΦΙΑ - Η ΔΗΜΟΣΙΟΓΡΑΦΙΑ ΣΤΗΝ ΨΗΦΙΑΚΗ ΕΠΟΧΗ.....	14
2.1 Η δημοσιογραφία στα ψηφιακά μέσα επικοινωνίας.....	14
2.1.1 Ενημερωτικές Ιστοσελίδες – Ειδησεογραφικοί Ιστοχώροι	15
2.1.2 Ιστολόγια.....	16
2.1.3 Μέσα Κοινωνικής Δικτύωσης.....	16
2.1.4 Η πλατφόρμα του Medium.....	17
3. Η ΕΝΝΟΙΑ ΤΗΣ ΔΗΜΟΤΙΚΟΤΗΤΑΣ ΣΤΑ ΨΗΦΙΑΚΑ ΜΕΣΑ ΕΠΙΚΟΙΝΩΝΙΑΣ.....	19
3.1 Δημοτικότητα Χρηστών – Δημοτικότητα Περιεχομένου	19
3.2 Μέτρηση Δημοτικότητας	20
3.2.1 Μέτρηση δημοτικότητας στην πλατφόρμα του Medium	22
3.3 Πρόβλεψη δημοτικότητας	23
4. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΡΘΡΩΝ.....	25
4.1 Χρήστης	26
4.1.1 Δημοτικότητα του Χρήστη.....	26
4.1.2 Στυλ συγγραφής.....	29
4.2 Περιεχόμενο.....	33
4.2.1 Δομή του άρθρου	33
4.2.2 Συναισθηματική Ανάλυση	36
4.3 Γενικό Πλαίσιο	39
4.3.1 Είδος άρθρου.....	39
4.3.2 Χρόνος δημοσίευσης.....	39
4.3.3 Γεωγραφική εγγύτητα	40
5. ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ	41
5.1 Τι είναι η Μηχανική Μάθηση	41

5.2	Είδη Μηχανικής Μάθησης	42
6.	ΕΡΕΥΝΑ.....	43
6.1	Σκοπός της Έρευνας.....	43
6.2	Μεθοδολογία.....	43
7.	ΑΠΟΤΕΛΕΣΜΑΤΑ.....	52
7.1	Κανόνες για ειδησεογραφικά άρθρα	52
7.2	Κανόνες για τεχνολογικά άρθρα	52
7.3	Κανόνες για επιχειρηματικά άρθρα	53
7.4	Κανόνες για άρθρα υγείας.....	54
7.5	Κανόνες για αθλητικά άρθρα.....	54
7.6	Κανόνες για άρθρα τρόπου ζωής	55
8.	ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ	56
8.1	Ειδησεογραφικά Άρθρα	56
8.2	Τεχνολογικά Άρθρα	57
8.3	Επιχειρηματικά Άρθρα	58
8.4	Άρθρα Υγείας	59
8.5	Αθλητικά Άρθρα.....	60
8.6	Άρθρα Τρόπου Ζωής.....	60
9.	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ.....	62
	ΠΑΡΑΡΤΗΜΑ Ι.....	64
	ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ	67

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Μοντέλο χαρακτηριστικών άρθρων	25
Εικόνα 2: Αλγόριθμος Δένδρου Απόφασης	45
Εικόνα 3: Κόμβοι Δένδρου Απόφασης.....	46
Εικόνα 4: Πλήθος άρθρων ανά κόμβο (Οι 20 μεγαλύτεροι)	47
Εικόνα 5: Πιθανότητες Δημοτικότητας (Οι 20 χαμηλότερες)	48
Εικόνα 6: Πιθανότητες Δημοτικότητας (Οι 20 υψηλότερες).....	49
Εικόνα 7: Διαφορές Πιθανοτήτων Δημοτικότητας (Οι 10 μεγαλύτερες).....	49

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Στατιστικά στοιχεία ανά είδος άρθρου	45
---	----

ΠΡΟΛΟΓΟΣ

Η παρούσα εργασία διενεργήθηκε ως Διπλωματική εργασία για το μεταπτυχιακό πρόγραμμα σπουδών «Ψηφιακά Μέσα Επικοινωνίας και Περιβάλλοντα Αλληλεπίδρασης» του τμήματος Επικοινωνίας και Μέσων Μαζικής Ενημέρωσης του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών-ΕΚΠΑ.

1. ΕΙΣΑΓΩΓΗ

Η δημοσιογραφία στο διαδίκτυο έχει σημειώσει ραγδαία ανάπτυξη από την δεκαετία του '90 μέχρι και σήμερα. Η κυριαρχία των Μέσων Κοινωνικής Δικτύωσης, δε θα μπορούσε να αφήσει ανεπηρέαστο τον τομέα της δημοσιογραφίας. Ένα από τα χαρακτηριστικά των μέσων αυτών, είναι η δημοτικότητα, η οποία αποτελεί μάλιστα ένα στοιχείο, το οποίο εμφανίζεται δημόσια και μπορεί να γίνει ορατό από οποιονδήποτε χρήστη σε μία ψηφιακή πλατφόρμα κοινωνικής δικτύωσης. Υπάρχουν διάφορες ενδείξεις δημοτικότητας, όπως οι κοινοποιήσεις ή η δήλωση ικανοποίησης, που αφορούν τις αναρτήσεις που γίνονται καθώς και η ακολούθηση, η οποία δεν αφορά άμεσα τη δημοτικότητα του περιεχομένου, αλλά του χρήστη που το δημοσιεύει. Οι δημοσιογραφικοί οργανισμοί, αλλά και οι μεμονωμένοι χρήστες, οι οποίοι δημοσιοποιούν τα άρθρα που συντάσσουν διαδικτυακά, θέλουν οι αναρτήσεις που κάνουν να κερδίζουν τις θετικές αντιδράσεις του αναγνωστικού κοινού και για το λόγο αυτό θα ήταν χρήσιμο να έχουν υπόψη τους μια φόρμουλα η οποία θα αυξάνει την πιθανότητα να γίνει δημοφιλές διαδικτυακά, το περιεχόμενο που παράγουν. Το θέμα της παρούσας διπλωματικής εργασίας, λοιπόν, είναι η κατασκευή μιας τέτοιας φόρμουλας, η οποία θα έχει την μορφή κανόνων, που αν ληφθούν υπόψη κατά την παραγωγή των δημοσιογραφικών ιστοριών, μπορούν να βελτιώσουν την πιθανότητα απόκτησης υψηλής δημοτικότητας.

Για την παραγωγή αυτών των κανόνων θα χρειαστεί η προσοχή να στραφεί σε κάποια συγκεκριμένη ψηφιακή πλατφόρμα, που θα πληροί ορισμένες προϋποθέσεις. Στην περίπτωση αυτή, η έρευνα θα προσανατολιστεί στην πλατφόρμα του Medium, στην οποία αναρτώνται δημοσιογραφικές ιστορίες μιας ευρείας γκάμας ειδών και στην οποία υπάρχει η δυνατότητα ακολούθησης των χρηστών καθώς και δήλωσης αρεσκείας των ιστοριών μέσω της θετικής αντίδρασης των χειροκροτημάτων. Η παραγωγή των κανόνων θα πρέπει να στηριχτεί σε ένα μοντέλο χαρακτηριστικών, τα οποία προκύπτουν από τους τομείς της δημοσιογραφίας, της επικοινωνίας, του μάρκετινγκ και των Μέσων Κοινωνικής Δικτύωσης και αφορούν το ίδιο το άρθρο, τον συγγραφέα καθώς και το χωροχρονικό περιβάλλον του. Το μοντέλο των χαρακτηριστικών, το οποίο θα χρησιμοποιηθεί στη συγκεκριμένη διπλωματική εργασία, θα βασιστεί σε μεγάλο βαθμό στο μοντέλο που ανέπτυξαν για την έρευνα τους για τα χαρακτηριστικά που συμβάλλουν στη δημοτικότητα των άρθρων στην πλατφόρμα του Medium, οι Sotirakou, Trilling, Germanakos και Mourlas (2019). Τα χαρακτηριστικά των άρθρων, τα οποία θα ληφθούν υπόψη στην συγκεκριμένη έρευνα είναι ποσοτικά δεδομένα που προκύπτουν

εξ' αρχής από ένα άρθρο, όπως για παράδειγμα το πλήθος των λέξεων ή ο αριθμός των εικόνων, αλλά και δεδομένα τα οποία δεν είναι εξ' αρχής αριθμητικά, ωστόσο με την κατάλληλη επεξεργασία ποσοτικοποιούνται. Τέτοια, παραδείγματος χάρη, είναι τα δεδομένα που προκύπτουν από τη συναισθηματική ανάλυση του κειμένου, μέσω της οποίας αποδίδεται μια αριθμητική τιμή σε κάθε συναίσθημα, ανάλογα με την ύπαρξη αντίστοιχων συναισθηματικά φορτισμένων λέξεων ή φράσεων μέσα στο κείμενο.

Στη συνέχεια, θα χρησιμοποιηθεί ένας αλγόριθμος Μηχανικής Μάθησης, και συγκεκριμένα αυτός του δέντρου απόφασης (Decision Tree). Ο αλγόριθμος αυτός, θα αξιοποιήσει ένα μεγάλο μέρος από τα άρθρα του medium.com που έχουν ληφθεί για την εκπόνηση της παρούσας έρευνας, ώστε να αποκτήσει «γνώση» και να προβλέπει ποια χαρακτηριστικά συμβάλλουν στην υψηλή δημοτικότητα τους. Η εφαρμογή αυτού του αλγορίθμου θα οδηγήσει σε ένα δενδρικό σχηματισμό, του οποίου οι κόμβοι – φύλλα θα περιέχουν ένα υποσύνολο των αρχικών άρθρων.

Ωστόσο, η έρευνα δε θα προσανατολιστεί στο κατά πόσο ικανός είναι ο αλγόριθμος να προβλέπει ποια άρθρα έχουν υψηλή δημοτικότητα και ποια όχι, όπως συμβαίνει σε αντίστοιχα προβλήματα πρόβλεψης. Αντίθετα, θα στραφεί στον ίδιο τον δενδρικό σχηματισμό που έχει παραχθεί, ώστε να αποκαλύψει τους κανόνες που συμβάλλουν στην συγγραφή δημοφιλών άρθρων.

Κάθε κανόνας θα παράγεται από το μονοπάτι που ορίζεται από τη ρίζα του δένδρου απόφασης μέχρι τον κόμβο – φύλλο. Κάθε κόμβος του δένδρου έχει τη μορφή μιας αριθμητικής έκφρασης στην οποία εμπλέκεται ένα χαρακτηριστικό του άρθρου (πχ. Αριθμός εικόνων ≥ 2.5). Άρα ο τελικός κανόνας για κάθε κόμβο – φύλλο θα περιέχει ένα σύνολο αριθμητικών εκφράσεων που συνδέονται με το λογικό «ΚΑΙ».

Κλείνοντας, αξίζει να σημειωθεί πως δεν είναι όλοι οι κόμβοι – φύλλα ισάξιοι, ισοπίθανοι και ισοπληθείς και για το λόγο αυτό δεν μπορούν να εγγυηθούν όλοι την παραγωγή έγκυρων κανόνων. Έτσι, χρειάζεται η ανάπτυξη ενός μοντέλου επιλογής κανόνων, ώστε να υπάρχει μια δικλείδα ασφαλείας που να εγγυάται την εγκυρότητα των αποτελεσμάτων της έρευνας. Επιπλέον, χρειάζεται για κάθε κανόνα να προσδιοριστεί η πιθανότητα παραγωγής δημοφιλών άρθρων, γιατί όπως προκύπτει και από τα παραπάνω, αυτή ενδέχεται να διαφέρει σε κάθε περίπτωση.

2. ΔΙΑΔΙΚΤΥΑΚΗ ΔΗΜΟΣΙΟΓΡΑΦΙΑ - Η ΔΗΜΟΣΙΟΓΡΑΦΙΑ ΣΤΗΝ ΨΗΦΙΑΚΗ ΕΠΟΧΗ

Ο όρος διαδικτυακή δημοσιογραφία χρησιμοποιείται για να περικλείσει εννοιολογικά το σύνολο των πρακτικών της μαζικής ενημέρωσης και επικοινωνίας που υλοποιούνται μέσω του διαδικτύου. Οι αλλαγές που έχουν συντελεστεί ,όμως, στο διαδίκτυο, ιδίως από τις αρχές της τρέχουσας χιλιετίας μέχρι και σήμερα είναι ραγδαίες και σε συνδυασμό με τις επίσης ραγδαίες εξελίξεις που παρατηρούνται σε κοινωνικό αλλά και οικονομικό επίπεδο, είναι προφανές πως και οι τομείς της επικοινωνίας και της δημοσιογραφίας δεν θα μπορούσαν να μείνουν ανεπηρέαστοι. (Πολίτης, 2008) Έτσι, τα παραδοσιακά μέσα επικοινωνίας και οι δημοσιογράφοι «αναγκάζονται» να προσαρμοστούν και να αναθεωρήσουν τους τρόπους με τους οποίους συλλέγουν το δημοσιογραφικό υλικό τους, τη συχνότητα που μεταδίδουν τις ειδήσεις, καθώς και τον τρόπο που τις παρουσιάζουν στο κοινό, λαμβάνοντας υπόψη τις ανάγκες του κόσμου για μια σύγχρονη μορφή επικοινωνίας καθώς και την απαίτησή τους για συνεχή προσαρμογή του εκάστοτε μέσου στις αλλαγές που λαμβάνουν χώρα στον τομέα της ηλεκτρονικής δημοσιογραφίας αλλά και του διαδικτύου γενικότερα.

2.1 Η δημοσιογραφία στα ψηφιακά μέσα επικοινωνίας

«Τα νέα μέσα επικοινωνίας είναι ψηφιακά μέσα μαζικής δημόσιας ή ιδιωτικής επικοινωνίας, που διαμοιράζουν πολυμεσικό περιεχόμενο κατ' απαίτηση του χρήστη τους, συγχρονικά ή και ασύγχρονα.» (Καϊμάκη, Καραμπάσης, Παναγιωτάκης, Τσενέ και Τσιμιτάκης, 2016) Σύμφωνα με την παραπάνω προσέγγιση αντικατοπτρίζονται δύο όψεις των νέων μέσων σε σχέση με τα παραδοσιακά. Η μία όψη θεωρεί τους χρήστες καταναλωτές ψηφιακού περιεχομένου, όπως και στα παραδοσιακά μέσα επικοινωνίας, με ορισμένες διαφορές σχετικές με την χρήση πολυμέσων αλλά και τη δυνατότητα σύγχρονης και ασύγχρονης δικτύωσης μεταξύ των χρηστών. Ωστόσο, αυτή είναι η μία όψη του νομίσματος. Τα νέα μέσα επικοινωνίας ξεχωρίζουν γιατί δίνουν τη δυνατότητα στους «καθημερινούς» ανθρώπους, να μην είναι απλώς καταναλωτές, αλλά και παραγωγοί περιεχομένου. Θα μπορούσε κανείς να θεωρήσει πως αυτά τα μέσα είναι περισσότερο δημοκρατικά, μιας και δίνουν ίσες ευκαιρίες σε όλους. Επιπλέον, οι συντελεστές του περιεχομένου που δημοσιεύεται δε χρειάζεται απαραίτητα να είναι επαγγελματίες ώστε να δημοσιοποιήσουν το υλικό που παράγουν, ούτε επίσης χρειάζεται να υπάρχει κάποια εκπαίδευση, κατάρτιση ή προϋπηρεσία στον τομέα των μέσων επικοινωνίας. Άνθρωποι που είχαν θεωρηθεί προηγουμένως μόνο ως κοινό των

μέσων ενημέρωσης, μπορούν να γίνουν πλέον παραγωγοί, αρκεί να κατανοήσουν ότι η μορφή των παραδοσιακών μέσων μαζικής επικοινωνίας, δεν είναι η μόνη αποδεκτή μορφή μέσων επικοινωνίας στη σημερινή εποχή. (Boler, 2010)

Οι τεχνολογικές εξελίξεις των τελευταίων δύο δεκαετιών έχουν αλλάξει σε πολύ σημαντικό βαθμό τη φύση και την λειτουργία των μέσων μαζικής ενημέρωσης και κατ' επέκταση της δημοσιογραφίας, η οποία γίνεται αντιληπτή υπό το πρίσμα της ραγδαίας ανάπτυξης του παγκοσμίου ιστού ως διαδικτυακή δημοσιογραφία και φέρει χαρακτηριστικά, καινούργια ή διαφοροποιημένα, τα οποία πολλές φορές ανατρέπουν τη φύση της παραδοσιακής μορφής της. Ένα από τα βασικά χαρακτηριστικά είναι το γεγονός ότι δίνεται μεγαλύτερη βαρύτητα στη διασύνδεση. Στη διάρκεια της τελευταίας εικοσαετίας, παρατηρείται ένας ολοένα και αυξανόμενος αριθμός ιστοτόπων δημοσιογραφικού χαρακτήρα, με μεγάλο μέρος αυτών να αρκούνται στην αναπαραγωγή άρθρων, ενώ παράλληλα εμφανίζονται και τα πρώτα ιστολόγια, «ως υπηρεσία εκλαΐκευσης της ψηφιακής μικρό-έκδοσης» (Καϊμάκη κ.ά., 2016) Τέλος, την μεγάλη επανάσταση στη διαδικτυακή δημοσιογραφία, έρχονται να φέρουν τα μέσα κοινωνικής δικτύωσης. Στις υποενότητες που ακολουθούν, θα παρουσιαστούν αναλυτικά τα διάφορα ψηφιακά μέσα επικοινωνίας, τα οποία υποστηρίζουν και ενδείκνυνται για δημοσίευση δημοσιογραφικών ιστοριών και οι χρήστες αυτών μπορούν να διαδραματίζουν είτε το ρόλο του δέκτη, είτε το ρόλο του παραγωγού.

2.1.1 Ενημερωτικές Ιστοσελίδες – Ειδησεογραφικοί Ιστοχώροι

Ο όρος ιστοσελίδα, έχει επικρατήσει να χρησιμοποιείται στην ελληνική γλώσσα, πολλές φορές λανθασμένα, για την περιγραφή ενός ιστοχώρου, δηλαδή ενός συνόλου ιστοσελίδων. Στην καθημερινή γλώσσα και όταν γίνεται αναφορά σε δημοσιογραφικές ιστοσελίδες, συνήθως πρόκειται για ιστοχώρους κυρίως ειδησεογραφικού χαρακτήρα, οι οποίοι έχουν δημιουργηθεί για χάρη κάποιου οργανισμού, σωματίου, εταιρίας ή μέσου ενημέρωσης. (Καϊμάκη κ.ά., 2016) Από τα παραπάνω, είναι προφανές ότι τα δημοσιογραφικά άρθρα που δημοσιεύονται σε τέτοια μέσα επικοινωνίας, φέρουν χαρακτηριστικά τα οποία εξαρτώνται τόσο από τον δημοσιογράφο, όσο και από το μέσο στο οποίο δημοσιεύονται. Οι ενημερωτικές «ιστοσελίδες», συνηθίζεται να έχουν μια συγκεκριμένη δομή και οργάνωση των θεμάτων τους σε κατηγορίες. Επιπλέον, σε έναν ειδησεογραφικό ιστοχώρο, τις περισσότερες φορές αρθρογραφούν παραπάνω από ένα άτομα, κυρίως έχοντας την ανάλογη εκπαίδευση ή προϋπηρεσία, ενώ ενδέχεται ο κάθε

δημοσιογράφος να γράφει και για εξειδικευμένα θέματα. Τα τελευταία έτη, ωστόσο, με την άνθιση της διαδικτυακής δημοσιογραφίας, έχουν κάνει την εμφάνισή τους, ιστότοποι ειδησεογραφικού χαρακτήρα, οι οποίοι κυρίως επικεντρώνονται στην αναδημοσίευση άρθρων.

2.1.2 Ιστολόγια

Αν ισχυριστεί κανείς ότι οι ιστοσελίδες είναι η δημοσιογραφία των οργανισμών, τότε τα ιστολόγια θα ήταν η δημοσιογραφία των πολιτών. (Καϊμάκη κ.ά., 2016) Πρόκειται ουσιαστικά για ιστοχώρους οι οποίοι δεν έχουν δημιουργηθεί για χάρη κάποιου οργανισμού, μέσου ενημέρωσης ή εταιρίας, αλλά από έναν άνθρωπο με σκοπό την έκφραση της προσωπικής του γνώμης σε κάποια θέματα. Τα ιστολόγια έχουν συνήθως απλή δομή, μιας και τα άρθρα εμφανίζονται με χρονολογική σειρά, από το πιο πρόσφατο στο λιγότερο, ενώ αν το επιθυμεί ο χρήστης μπορεί να οργανώσει το περιεχόμενό του σε κατηγορίες. Η κατασκευή ενός ιστολογίου είναι εύκολη, ειδικά σε σύγκριση με την κατασκευή μιας «ιστοσελίδας», αφού δεν απαιτούνται ειδικές γνώσεις και μπορεί να γίνει δωρεάν από οποιονδήποτε. Τα ιστολόγια γνώρισαν ιδιαίτερη άνθιση στη χώρα μας, στις αρχές της περασμένης δεκαετίας, λίγο πριν την καθιέρωση των μέσων κοινωνικής δικτύωσης. Τα ενημερωτικά blogs φαίνεται πως είχαν έρθει σε μια άτυπη αντιπαράθεση με τα παραδοσιακά μέσα ενημέρωσης, για την κυριαρχία στην πληροφόρηση του κοινού, ωστόσο τα τελευταία έτη έχουν περιοριστεί σε μεγάλο βαθμό. Αξίζει να σημειωθεί, ότι η θεματολογία της πλειονότητας των ιστολογίων έχει αποκλίνει από την ειδησεογραφία, ωστόσο σε πολλούς ειδησεογραφικούς ιστοχώρους μπορεί να συναντήσει κανείς ιστοσελίδες που λειτουργούν ως ιστολόγια. (Καϊμάκη κ.ά., 2016)

2.1.3 Μέσα Κοινωνικής Δικτύωσης

Τα μέσα κοινωνικής δικτύωσης αποτελούν ίσως μια από τις δημοφιλέστερες μορφές ψηφιακών μέσων επικοινωνίας, αλλά και μέσων επικοινωνίας γενικότερα. Η ευρεία χρήση τους και η ενσωμάτωσή τους στην καθημερινότητα των ανθρώπων, είναι ένα αδιαμφισβήτητο γεγονός και για το λόγο αυτό, δε θα μπορούσε να μην υπάρχει εμπλοκή τους και στον τομέα της ενημέρωσης. Η φύση τους τα καθιστά μάλιστα κατάλληλα για την μετάδοση ειδήσεων, μιας και η αμεσότητα είναι ένα από τα βασικότερα χαρακτηριστικά τους.

Το Twitter, το οποίο αποτελεί ένα από τα πιο δημοφιλή κοινωνικά μέσα δικτύωσης, έχει διαδραματίσει σημαντικό ρόλο στην ενημέρωση των πολιτών σε παγκόσμια εμβέλεια. Μάλιστα, το συγκεκριμένο μέσο είναι ιδανικό για το διαμοιρασμό πληροφορίας και ενημέρωση της κοινότητας των ατόμων που ακολουθούν έναν λογαριασμό. Το Twitter, χρησιμοποιείται ευρέως από πολιτικά πρόσωπα, αλλά και από δημοσιογραφικούς οργανισμούς, ενώ εκτός από τη δυνατότητα σύνταξης ενός σύντομου κειμένου, υπάρχει και η δυνατότητα κοινοποίησης άρθρων ή πολυμέσων. Η λογική του συγκεκριμένου μέσου έχει ορισμένα κοινά στοιχεία με τα ιστολόγια, μιας και υπάρχει μια συνεχής ροή «τιτιβισμάτων», δυνατότητα προβολής των δημοσιεύσεων με πρώτα τις πιο πρόσφατες, αλλά και κατηγοριοποίηση της επίκαιρης θεματολογίας μέσω των hash tags.

Από την άλλη πλευρά, το Facebook, αν και δεν ξεκίνησε με αυτό το σκοπό, και σε μία προσπάθεια κάλυψης ενός ευρέος φάσματος λειτουργιών, χρησιμοποιείται από τους περισσότερους δημοσιογραφικούς οργανισμούς για διαμοιρασμό του περιεχομένου τους. Η απήχηση του συγκεκριμένου μέσου, όπως και η δύναμή του, είναι πολύ μεγάλη και κατά το παρελθόν έχει διαδραματίσει σημαντικό ρόλο στην ενημέρωση, αλλά και τον συντονισμό διαφόρων κοινωνικών ενεργειών και διαμαρτυριών τόσο στην Ελλάδα, όσο και σε παγκόσμιο επίπεδο. (Καϊμάκη κ.ά., 2016)

2.1.4 Η πλατφόρμα του Medium

Όπως αναφέρθηκε και στην προηγούμενη ενότητα, τα μέσα κοινωνικής δικτύωσης έχουν αρχίσει να χρησιμοποιούνται κατά κόρον από δημοσιογραφικούς οργανισμούς, αλλά και δημοσιογράφους γενικότερα. Το Medium αναπτύχθηκε από τον Evans William, συνιδρυτή του Blogger και του Twitter, με σκοπό την δημοσίευση κειμένων, μεγαλύτερης έκτασης από τα σύντομα τιτιβίσματα που επιτρέπει το Twitter, αλλά και εγγράφων. Η λειτουργία της πλατφόρμας ξεκίνησε το 2012 και έχει χαρακτηριστικά που μπορεί να συναντήσει κανείς σε ενημερωτικές ιστοσελίδες, ιστολόγια αλλά και στα μέσα κοινωνικής δικτύωσης. Στο Medium, δημοσιεύουν τα κείμενά τους ερασιτέχνες αλλά και επαγγελματίες δημοσιογράφοι, ενώ η πλατφόρμα μπορεί να θεωρηθεί σαν ένας ιστοχώρος που «φιλοξενεί» ιστολόγια. (Medium (website),χ.χ.) Οι δημοσιογραφικές ιστορίες που συναντώνται στο medium.com, είναι γραμμένες στην αγγλική γλώσσα, ωστόσο, ορισμένες δημοσιεύσεις μπορεί να είναι γραμμένες και στα γαλλικά, τα ισπανικά ή άλλες γλώσσες. Στα πλαίσια της παρούσας διπλωματικής εργασίας, η

έρευνα για τους κανόνες συγγραφής άρθρων, που έχουν μεγάλη πιθανότητα να γίνουν δημοφιλή, θα προσανατολιστεί μόνο σε άρθρα, τα οποία είναι γραμμένα στα Αγγλικά.

3. Η ΕΝΝΟΙΑ ΤΗΣ ΔΗΜΟΤΙΚΟΤΗΤΑΣ ΣΤΑ ΨΗΦΙΑΚΑ ΜΕΣΑ ΕΠΙΚΟΙΝΩΝΙΑΣ

Οι άνθρωποι «καταναλώνουν» ψηφιακό περιεχόμενο μέσα από μια ποικιλία καναλιών. Το ίδιο ισχύει και στην περίπτωση των ειδήσεων, αλλά και των δημοσιογραφικών άρθρων που αναρτώνται στο διαδίκτυο γενικότερα. (Cheng, Wang, Zhang και Hu, 2016). Αν αναλογιστεί κανείς το μεγάλο πλήθος των ψηφιακών μέσων, αλλά και των μέσων κοινωνικής δικτύωσης καθώς και το γεγονός ότι για κάθε μέσο υπάρχουν ποικίλα μέτρα δημοτικότητας, είναι φανερό ότι υπάρχει η ανάγκη καθορισμού της έννοιας της δημοτικότητας καθώς και των τρόπων με τους οποίους αυτή μετράται. Επισκεψιμότητα ιστοχώρου, πλήθος ακολούθων, πλήθος θετικών αντιδράσεων, πλήθος κοινοποιήσεων είναι μόνο κάποια από τα βασικά μέτρα της δημοτικότητας στα ψηφιακά μέσα, ωστόσο είναι κρίσιμο να αξιολογηθεί ποια από αυτά μπορούν να φανούν χρήσιμα στον τομέα της διαδικτυακής δημοσιογραφίας και πιο συγκεκριμένα στην πρόβλεψη της πορείας των δημοσιογραφικών άρθρων που δημοσιεύονται διαδικτυακά.

3.1 Δημοτικότητα Χρηστών – Δημοτικότητα Περιεχομένου

Τόσο το Facebook και το Twitter, όσο και τα περισσότερα Μέσα Κοινωνικής Δικτύωσης διαθέτουν κάποια χαρακτηριστικά γνωρίσματα που τα διαφοροποιούν από τα υπόλοιπα ψηφιακά μέσα επικοινωνίας. Αυτά τα χαρακτηριστικά έχουν να κάνουν κυρίως με την μέτρηση της απήχησης του παραγόμενου υλικού καθώς και του χρήστη ο οποίος το δημοσιεύει. Έτσι:

- Κάθε χρήστης έχει τη δυνατότητα να επικοινωνήσει με ένα σύνολο άλλων χρηστών, με τους οποίους έχει μιας μορφής σύνδεση, κατά συνέπεια οι χρήστες αυτού του είδους ψηφιακών μέσων μπορούν να δημιουργήσουν ένα δικό τους δίκτυο επικοινωνίας. Κάποια μέσα ενδέχεται να υποστηρίζουν σύγχρονη μορφή σύνδεσης, όπως το Facebook, μιας και ο τρόπος δημιουργίας της σύνδεσης ανάμεσα σε δύο προφίλ, είναι αμφίδρομος μέσω της φιλίας μεταξύ δύο χρηστών. Ωστόσο, όταν πρόκειται για κάποια σελίδα στο Facebook ή για κάποιο προφίλ στο Twitter, η μορφή σύνδεσης είναι ασύγχρονη, μιας και δεν είναι απαραίτητο ένας χρήστης να ακολουθεί τους χρήστες που τον ακολουθούν.
- Κάθε ανάρτηση που γίνεται στα Μέσα Κοινωνικής Δικτύωσης από κάποιον χρήστη, έχει τη δυνατότητα να αποσπάσει μιας ή περισσότερων μορφών «αντιδράσεις» από τους υπόλοιπους χρήστες, οι οποίες δηλώνονται με το

πάτημα ενός «κουμπιού» που συνοδεύει την ανάρτηση. Στο Facebook, παραδείγματος χάρη, οι χρήστες μπορούν να δηλώσουν την ικανοποίησή, τη δυσφορία, τη λύπη, την έκπληξη, αλλά και το ότι βρήκαν κάτι αστείο μέσα από έξι διαφορετικές αντιδράσεις. Κάτι αντίστοιχο συμβαίνει και στις υπόλοιπες πλατφόρμες κοινωνικής δικτύωσης, όπως επίσης υπάρχει και η δυνατότητα σχολιασμού και κοινοποίησης μιας ανάρτησης.

Υπάρχουν δηλαδή δύο είδη δημοτικότητας, στις πλατφόρμες κοινωνικής δικτύωσης. Η μία έχει να κάνει με τους χρήστες και η άλλη με το περιεχόμενο που δημοσιεύεται. Όπως μπορεί να γίνει εύκολα αντιληπτό, τα δύο αυτά είδη δημοτικότητας επηρεάζονται μεταξύ τους και το ένα μπορεί να συμβάλει θετικά στην ανάπτυξη του άλλου.

3.2 Μέτρηση Δημοτικότητας

Η μέτρηση της απήχησης στα ψηφιακά μέσα επικοινωνίας αποκτά μια νέα πτυχή. Εκτός από την επισκεψιμότητα των ιστοσελίδων και των ιστολογίων, οι δημοσιογραφικοί ιστότοποι θα πρέπει πλέον να έχουν υπόψη τους και των αριθμό των ακολούθων στα Μέσα Κοινωνικής Δικτύωσης, στα οποία χρειάζεται να έχουν παρουσία, αλλά και τον αριθμό των αντιδράσεων που λαμβάνουν οι δημοσιεύσεις του περιεχομένου τους σε αυτά. Οι αντιδράσεις των κοινωνικών μέσων όμως ενδέχεται να ποικίλλουν όχι μόνο στον όγκο αλλά και ποιοτικά. Πιο συγκεκριμένα, στην έρευνα των Hu, Farnham και Talamadurula (2015), στην οποία καταγράφηκαν οι αναρτήσεις χρηστών στο Twitter κατά τη διάρκεια προβολής της ομιλίας του προέδρου των ΗΠΑ, παρατηρήθηκε ότι μεγάλο μέρος των αναφορών σε αυτή την ομιλία είχαν γενικό χαρακτήρα και ότι αυτές έγιναν μόνο εξαιτίας της δημοτικότητας του συγκεκριμένου θέματος. Συνεπώς, το διαδικτυακό περιεχόμενο που γίνεται δημοφιλές έχει την τάση να γίνεται ακόμα πιο δημοφιλές, αφού οι χρήστες τείνουν να το «εκμεταλλευτούν» λόγω της δημοτικότητάς του.

Μέχρι στιγμής έχουν αναφερθεί μερικές έννοιες, σε μια προσπάθεια προσέγγισης και μέτρησης της επιτυχίας, μιας και αυτή αποτελεί έναν γενικότερο όρο, ο οποίος μπορεί να αποδοθεί με ποικίλους τρόπους. Απήχηση, ποιότητα, θετικές αντιδράσεις, δημοτικότητα είναι κάποιες από αυτές, οι οποίες αντανακλούν το μέγεθος της επιτυχίας που έχει το περιεχόμενο των χρηστών στις πλατφόρμες των μέσων κοινωνικής δικτύωσης, αλλά και οι ίδιοι ή ο οργανισμός τους. Παρακάτω θα δοθούν οι περιγραφές

ορισμένων εννοιών που σχετίζονται με την επιτυχία μιας ανάρτησης στα μέσα κοινωνικής δικτύωσης:

Απήχηση: Η έννοια αυτή αναφέρθηκε προηγουμένως και σύμφωνα με το ελληνικό λεξικό που χρησιμοποιεί η ιστοσελίδα greek-language.gr σημαίνει «η εντύπωση που προκαλεί κάποιο γεγονός και η επίδραση που ασκεί ή οι αντιδράσεις που προκαλεί» (Dictionary of Standard Modern Greek, χ.χ.). Ωστόσο, στη γλώσσα των μέσων κοινωνικής δικτύωσης, ο όρος «απήχηση» έχει να κάνει με το πλήθος των ατόμων που είδαν οποιοδήποτε περιεχόμενο από μια σελίδα στο Facebook (“Ποια είναι η διαφορά ανάμεσα στις προβολές Σελίδας”, χ.χ.), είτε ύστερα από κάποια διαφημιστική καμπάνια, είτε όχι (“Σε τι διαφέρουν μεταξύ τους η οργανική απήχηση”, χ.χ.). Η έννοια της «απήχησης», όπως είναι συνηθισμένο να χρησιμοποιείται στο Facebook και στα αντίστοιχα μέσα, δεν εμπλέκει την δήλωση κάποιας μορφής αντίδρασης από την πλευρά του κοινού. Για το λόγο αυτό, αλλά και για την αποφυγή της σύγχυσης, στα πλαίσια της παρούσας εργασίας, δε θα χρησιμοποιηθεί ο όρος «απήχηση», για την μέτρηση της επιτυχίας. Θα πρέπει όμως να σημειωθεί ότι η απήχηση με τη σημασία της αναγνωρισιμότητας της σελίδας και του περιεχομένου της, διαδραματίζει σημαντικό ρόλο στην επιτυχία μιας δημοσίευσης στα κοινωνικά μέσα δικτύωσης, αφού είναι πολύ σημαντικό μια ανάρτηση να έχει γίνει ορατή σε ένα μεγάλο μέρος του κοινού ώστε να θεωρηθεί επιτυχημένη.

Δήλωση Ικανοποίησης: Η έννοια αυτή συμπεριλαμβάνει την έννοια της απήχησης μιας δημοσίευσης, αφού για να δηλώσει ένας χρήστης ότι του αρέσει μια ανάρτηση, σημαίνει ότι έφτασε με κάποιον τρόπο στην «αρχική του σελίδα» και έγινε ορατή σε αυτόν. Η δήλωση της ικανοποίησης του χρήστη είναι μια μορφή αλληλεπίδρασης, η οποία συμβαίνει με τη δήλωση μιας θετικής αντίδρασης για μια ανάρτηση σε ένα μέσο κοινωνικής δικτύωσης. Αξίζει να σημειωθεί ότι δεν είναι η μόνη μορφή αλληλεπίδρασης, αφού ο χρήστης μπορεί να δηλώσει μια αρνητική αντίδραση, να κάνει κάποιο σχόλιο ή να κοινοποιήσει την ανάρτηση. Ένα μεγάλο μέρος των χρηστών στις πλατφόρμες κοινωνικής δικτύωσης παρουσιάζει συνήθως μιας μορφής παθητικότητα απέναντι στο περιεχόμενο που παράγουν οι άλλοι χρήστες, για το λόγο αυτό είναι σημαντικό για τους παραγωγούς περιεχομένου να βρουν έναν τρόπο ώστε υπερνικήσουν την παθητικότητα των χρηστών. (Romero, Galuba, Asur και Huberman 2011) Σίγουρα, το γεγονός ότι κάποια ανάρτηση σε ένα μέσο κοινωνικής δικτύωσης προκαλεί την εμπλοκή μεγάλου μέρους των χρηστών, μέσω των αλληλεπιδράσεων, αποτελεί μια ένδειξη αυξημένου ενδιαφέροντος, ενώ αν η εκδήλωση του ενδιαφέροντος είναι θετικής μορφής (θετική

αντίδραση, κοινοποίηση, θετικό σχόλιο), τότε σίγουρα η ανάρτηση μπορεί να θεωρηθεί δημοφιλής. (Hu κ.ά, 2015)

Ποιότητα: Η ποιότητα του περιεχομένου που δημοσιεύεται στα μέσα κοινωνικής δικτύωσης ενδέχεται να επηρεάσει την αλληλεπίδραση των χρηστών με μία ανάρτηση. Αντίστοιχα και η ποιότητα συγγραφής των δημοσιογραφικών άρθρων, μπορεί να αποτελέσει έναν παράγοντα που θα οδηγήσει στη συγκέντρωση μεγάλου πλήθους θετικών αντιδράσεων, μέσω των οποίων δηλώνεται η ικανοποίηση των αναγνωστών. (Cheng κ.ά, 2016)

Η δημοτικότητα είναι μια έννοια που ορίζει το πόσο δημοφιλής είναι κάποιος ή κάτι. Έτσι, στα πλαίσια της συγκεκριμένης εργασίας η δημοτικότητα στον τομέα της διαδικτυακής δημοσιογραφίας, αφορά το πόσο δημοφιλές είναι το περιεχόμενο που δημοσιεύουν οι μεμονωμένοι χρήστες ή οι δημοσιογραφικοί οργανισμοί στα ψηφιακά μέσα ενημέρωσης. Ωστόσο, αξίζει να παρουσιαστεί ποια είναι η σχέση της δημοτικότητας σε σχέση με τις προηγούμενες έννοιες.

Η ποιότητα, όπως έχει ήδη αναφερθεί είναι ένας παράγοντας που συμβάλλει στη δημοτικότητα των δημοσιογραφικών άρθρων στο διαδίκτυο, όμως δεν είναι ο μόνος. Στο επόμενο κεφάλαιο θα γίνει λεπτομερής αναφορά στους παράγοντες στους οποίους μπορεί να οφείλεται η υψηλή δημοτικότητα των άρθρων. Η απήχηση, με την έννοια του πόσοι βλέπουν το εκάστοτε άρθρο είναι και αυτό ένας παράγοντας που ενδεχομένως συμβάλλει στη δημοτικότητα των δημοσιογραφικών ιστοριών. Αντιθέτως, η έννοια της δήλωσης της ικανοποίησης, μέσω των αντιδράσεων που προσφέρουν οι πλατφόρμες κοινωνικής δικτύωσης, μπορεί να αποτελέσει έναν τρόπο μέτρησης της δημοτικότητας. Για τη συγκεκριμένη έρευνα, η δημοτικότητα των δημοσιογραφικών ιστοριών θα μετρηθεί μέσα από τη δήλωση ικανοποίησης του κοινού, στις αντίστοιχες δημοσιεύσεις των ιστοριών στην πλατφόρμα του Medium.

3.2.1 Μέτρηση δημοτικότητας στην πλατφόρμα του Medium

Στο medium.com, οι χρήστες μπορούν να δημιουργήσουν το δικό τους προφίλ, να αποκτήσουν ακόλουθους αλλά και να ακολουθήσουν τους άλλους χρήστες. Ωστόσο, η μορφή διασύνδεσης είναι ασύγχρονη, όπως για παράδειγμα στο Twitter, μιας και δεν είναι απαραίτητο ένας χρήστης να ακολουθεί τους χρήστες που τον ακολουθούν. Η δημοτικότητα των χρηστών προκύπτει από το πλήθος των ακολούθων που έχουν

καταφέρει να αποκτήσουν. Ωστόσο, η δημοτικότητα των δημοσιεύσεών τους μπορεί να μετρηθεί με τρεις τρόπους:

- με το πλήθος των κλικ
- με το πλήθος των κοινοποιήσεων, ή
- με το πλήθος των αντιδράσεων που λαμβάνει η εκάστοτε δημοσίευση.

Το medium.com, διαθέτει μια αντίδραση ή οποία καλείται χειροκρότημα (claps). Η αντίδραση αυτή έχει θετική χροιά και χρησιμοποιείται για να επιδοκιμάσει τον συγγραφέα. Στην ουσία μπορεί να αποτελέσει δήλωση ικανοποίησης του αναγνώστη σε σχέση με αυτό που διάβασε. Μάλιστα ο χρήστης μπορεί να «χειροκροτήσει» ένα άρθρο έως και πενήντα φορές, πράγμα το οποίο καθιστά το «χειροκρότημα» μιας σύνθετης μορφή αντίδραση. Η αντίδραση του χειροκροτήματος όμως, χρησιμοποιείται και από την ίδια την πλατφόρμα για να καθορίσει ποια άρθρα θα βλέπει ο κάθε χρήστης, αφού λαμβάνει υπόψη το πλήθος των αντιδράσεων αυτών για να προτείνει το περιεχόμενο στους ακόλουθους των χρηστών. Είναι εύλογο, λοιπόν, ότι οι χρήστες δε θέλουν να τους προτείνεται περιεχόμενο, το οποίο δεν είναι καλό, συνεπώς θα μπορούσε να θεωρηθεί ότι τα χειροκροτήματα της πλατφόρμας σχετίζονται και με την ποιότητα των ιστοριών που δημοσιεύονται. Μάλιστα, δεδομένου ότι το πλήθος των θετικών αντιδράσεων, είναι ένα στοιχείο, το οποίο εμφανίζεται δημόσια, οι χρήστες ενδέχεται να επηρεαστούν από αυτό και να θεωρήσουν ως περισσότερο ποιοτικά τα άρθρα με το μεγαλύτερο πλήθος χειροκροτημάτων. (Sotirakou κ.ά., 2019)

3.3 Πρόβλεψη δημοτικότητας

Η πρόβλεψη της δημοτικότητας των δημοσιογραφικών άρθρων στο διαδίκτυο, είναι ένα σχετικά καινοφανές πρόβλημα, με τις μελέτες πάνω στο συγκεκριμένο ζήτημα να μην είναι ιδιαίτερα πολλές σε πλήθος. Ωστόσο, υπάρχουν αρκετές έρευνες τα τελευταία χρόνια σχετικές με την πρόβλεψη της δημοτικότητας του περιεχομένου που παράγεται και καταναλώνεται διαδικτυακά, οι οποίες αφορούν μέσα όπως το Twitter, αλλά και το YouTube. Σε ό,τι αφορά την πρόβλεψη των ειδήσεων που δημοσιεύονται στο διαδίκτυο, οι έρευνες προσανατολίζονται στην πρόβλεψη τόσο πριν, όσο και μετά από ένα σύντομο χρονικό διάστημα από τη δημοσίευση του περιεχομένου, με βάση την αξιοποίηση κάποιων πρώτων στοιχείων που μπορούν να αντληθούν από το εκάστοτε μέσο. (Cheng κ.ά., 2016) Ωστόσο, για την παρούσα μελέτη η προσοχή θα στραφεί στην πρόβλεψη πριν τη δημοσίευση των άρθρων. Με δεδομένο μάλιστα ότι τα

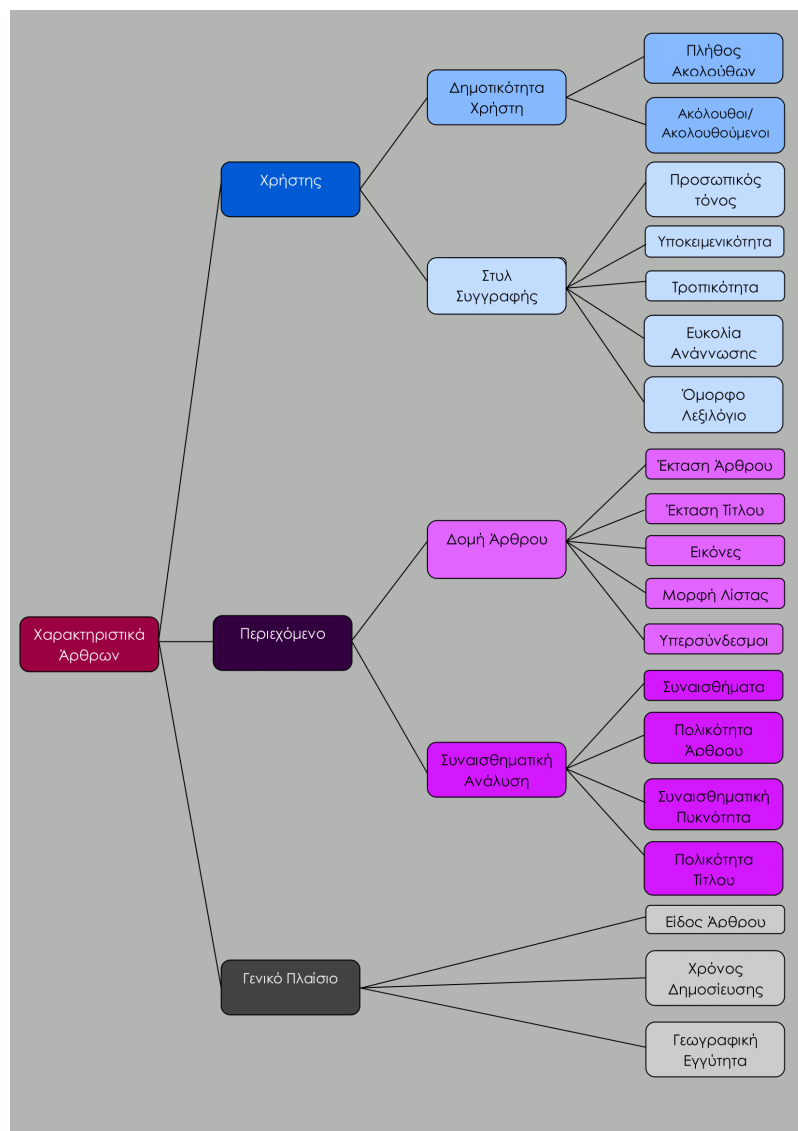
ειδησεογραφικά άρθρα είναι άμεσα εξαρτώμενα από το χρόνο δημοσίευσης, είναι σημαντικό η πρόβλεψη να γίνεται σύμφωνα με τα στοιχεία που είναι διαθέσιμα πριν από δημοσιοποίηση της εκάστοτε ιστορίας. (Bandari και Asur και Huberman,2012)

Για την πραγματοποίηση της πρόβλεψης, θα χρειαστεί η ανάπτυξη ενός μοντέλου, το οποίο θα αποτελείται από συγκεκριμένα χαρακτηριστικά του άρθρου, ώστε να «αποκαλυφθεί» ποια από αυτά συμβάλλουν στην υψηλή δημοτικότητα, καθώς και με ποιον τρόπο συνδυαστικά με τα υπόλοιπα. Το μοντέλο των χαρακτηριστικών συνδυαστικά με την τεχνολογία της Μηχανικής Μάθησης, θα χρησιμοποιηθεί για την πρόβλεψη της δημοτικότητας και στην συνέχεια για την παραγωγή κανόνων συγγραφής άρθρων τα οποία έχουν αυξημένη πιθανότητα να αποκτήσουν μεγάλο πλήθος αντιδράσεων θετικής χροιάς.

4. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΡΘΡΩΝ

Τα χαρακτηριστικά των άρθρων μπορούν να γίνουν αντιληπτά υπό τρεις οπτικές. Στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας και λαμβάνοντας υπόψη τις αντίστοιχες τρεις διαστάσεις των χαρακτηριστικών των ειδήσεων στο Twitter που πρότειναν και μελέτησαν στην έρευνά τους οι Orellana-Rodriguez και Keane (2018), οι οπτικές υπό τις οποίες γίνονται αντιληπτά τα χαρακτηριστικά των δημοσιογραφικών άρθρων είναι:

- η οπτική του χρήστη,
- η οπτική του περιεχομένου, και
- η οπτική του περιβάλλοντος.



Εικόνα 1: Μοντέλο χαρακτηριστικών άρθρων

Λαμβάνοντας υπόψη τη μοντελοποίηση των Sotirakou κ.ά. (2019), αλλά και των Arapakis, Peleja, Cambazoglu και Magalhães (2016), το από κάτω προς τα επάνω μοντέλο στο οποίο θα βασιστεί η παρούσα έρευνα, επεκτείνεται με την οπτική του χρήστη να αποτελείται από δύο υποδιαστάσεις που προσδιορίζουν τον συγγραφέα, τη δημοτικότητα και το προσωπικό στυλ συγγραφής. Η οπτική του περιεχομένου επίσης επεκτείνεται σε δύο υποδιαστάσεις, που προσδιορίζουν τη δομή του κειμένου, αλλά και το συναίσθημα που προκύπτει από την ανάγνωσή του. Στην Εικόνα 1, παρουσιάζονται ομαδοποιημένα για κάθε οπτική τα χαρακτηριστικά των άρθρων που θα μελετηθούν για το αν και πόσο συμβάλλουν στη συγκέντρωση υψηλής δημοτικότητας στην πλατφόρμα που δημοσιεύονται.

4.1 Χρήστης

Τα χαρακτηριστικά που γίνονται αντιληπτά υπό αυτή τη διάσταση, αφορούν τον συντάκτη του κάθε δημοσιογραφικού άρθρου. Οι χρήστες της πλατφόρμας του Medium διατηρούν τα χαρακτηριστικά των χρηστών των κοινωνικών μέσων δικτύωσης, αλλά ταυτόχρονα είναι και συντάκτες δημοσιογραφικών ιστοριών. Για το λόγο αυτό, τα χαρακτηριστικά που επιλέχθηκαν λαμβάνουν υπόψη αυτή τη διττή φύση των χρηστών.

4.1.1 Δημοτικότητα του Χρήστη

Οι συγγραφείς, που διατηρούν ένα προφίλ στις πλατφόρμες κοινωνικής δικτύωσης, μπορούν να ακολουθήσουν, αλλά και να ακολουθηθούν από άλλους χρήστες.

4.1.1.1 Το πλήθος των ατόμων που ακολουθούν τον χρήστη (Followers)

Το χαρακτηριστικό αυτό, φαίνεται πως είναι ένα από τα σημαντικότερα, είτε βλέποντας το άρθρο από την οπτική του χρήστη, δηλαδή του συγγραφέα, είτε βλέποντάς το από μια γενικότερη οπτική, και αυτό συμβαίνει γιατί το συγκεκριμένο χαρακτηριστικό αντικατοπτρίζει τη δημοτικότητα του συγγραφέα. Σύμφωνα με τη μελέτη των Sotirakou κ.ά. (2019), η οποία είναι προσανατολισμένη στο Medium, η δημοτικότητα του συγγραφέα είναι εκείνο το χαρακτηριστικό που συμβάλει περισσότερο, ανάμεσα και σε πολλά άλλα, στη δημοτικότητα ενός άρθρου. Όπως έχει ήδη αναφερθεί, το Medium είναι μια πλατφόρμα η οποία έχει πολλά κοινά χαρακτηριστικά με τα μέσα κοινωνικής δικτύωσης. Ένα από αυτά είναι η δυνατότητα που έχει κάθε χρήστης να αποκτήσει

ακόλουθους. Στην περίπτωση του medium.com, κάθε χρήστης δημιουργεί και διατηρεί το δικό του λογαριασμό και προφίλ στον ιστοχώρο αυτό. Έτσι, και σε αντιστοιχία με το Twitter (Orellana-Rodriguez και Keane, 2018), όπως και γενικότερα στην πλειονότητα των μέσων κοινωνικής δικτύωσης υπάρχει η δυνατότητα μέτρησης της δημοτικότητας του κάθε χρήστη. Η δημοτικότητα των χρηστών καθορίζεται από το πλήθος των ακολούθων που έχουν καταφέρει να αποκτήσουν. Στην έρευνά της, σχετικά με το YouTube, η Yoganarasimhan (2012) παρατήρησε ότι στην ουσία ο αριθμός των ακολούθων ενός χρήστη δείχνει το μέγεθος του κοινού στο οποίο απευθύνεται το βίντεο. Έτσι, αν κάποιος χρήστης διαθέτει ένα μεγάλο πλήθος ατόμων που τον ακολουθεί, τότε το περιεχόμενο που παράγει, θα διαδοθεί ταχύτερα και ευρύτερα. Είναι φανερό, λοιπόν ότι όταν ένας συντάκτης με μεγάλο πλήθος ακολούθων στο Medium, δημοσιεύει τη δημοσιογραφική του ιστορία, η ιστορία αυτή αναμένεται να γίνει γνωστή σε περισσότερα άτομα, αλλά και σε μικρότερο χρονικό διάστημα, σε σχέση με μια αντίστοιχη ιστορία που θα είχε παραχθεί από κάποιον χρήστη με σημαντικά πιο λίγους ακόλουθους. Αυτός είναι και ένας από τους κύριους λόγους, σύμφωνα με τους Paul, Hong και Chi (2012), που παρατηρείται μεγάλη διαφορά δημοτικότητας περιεχομένου στα κοινωνικά μέσα δικτύωσης. Κατά αντιστοιχία και στο Medium, οι νέοι συντάκτες θα πρέπει να καταβάλουν αρκετά μεγάλη προσπάθεια ώστε να παράξουν ποιοτικό περιεχόμενο που παράλληλα θα έλκει το ενδιαφέρον των αναγνωστών και θα συγκεντρώνει την αποδοχή τους μέσω των χειροκροτημάτων, σε αντίθεση με τους καταξιωμένους και δημοφιλείς συντάκτες οι οποίοι ενδέχεται να μην παράγουν εξίσου ποιοτικά άρθρα, αλλά τελικά καταφέρνουν να αποκομίσουν τα «χειροκροτήματα» των πολυπληθών ακολούθων τους. (Jin, Huang και Wang, 2017)

Στο σημείο αυτό αξίζει να διευκρινιστεί ότι οι ακόλουθοι που έχει κάποιος χρήστης αποτελούν μια ένδειξη δημοτικότητας του ίδιου του χρήστη και όχι απαραίτητα και του περιεχομένου που αυτός παράγει. Επιπροσθέτως, το γεγονός ότι το υλικό που δημοσιεύει ένας δημοφιλής χρήστης γίνεται ευρέως και ταχύτατα γνωστό, δε συνεπάγεται αυτόματα ότι θα ικανοποιήσει και το κοινό, ώστε αυτό να δηλώσει την ικανοποίησή του μέσω των χειροκροτημάτων του Medium, ή των άλλων αντιδράσεων στα υπόλοιπα κοινωνικά μέσα δικτύωσης. Σύμφωνα με την έρευνα των Cha, Haddadi, Benvenuto και Gummadi (2010) η ύπαρξη μεγάλου αριθμού ακολούθων σε ένα προφίλ στο Twitter δεν συνεπάγεται απαραίτητα ότι θα υπάρχει και η αντίστοιχη αλληλεπίδρασή τους με τις δημοσιεύσεις που γίνονται από το προφίλ αυτό. Επιπροσθέτως αυτό έρχεται να εξάρει και η μελέτη του Romero κ.ά. (2011), η οποία υποστηρίζει ότι η ύπαρξη πολύ μεγάλου αριθμού ακολούθων δρα ανασταλικά στην αποκόμιση αλληλεπιδράσεων,

καθώς οδηγεί σε μείωση της ευαισθητοποίησης, της εμπλοκής και της συμμετοχής του κοινού λόγω της εντύπωσης των ακολούθων ότι οι αλληλεπιδράσεις τους γίνονται λιγότερο προσωπικές.

Κατά συνέπεια, θα μπορούσε κανείς να ισχυριστεί ότι η υψηλή δημοτικότητα του χρήστη είναι αναγκαία ως ένα βαθμό συνθήκη, αλλά σίγουρα όχι ικανή από μόνη της για την υψηλή δημοτικότητα του άρθρου. Επιπλέον, ναι μεν η δημοτικότητα του άρθρου επηρεάζεται από την δημοτικότητα του συγγραφέα, ωστόσο όχι αναλογικά.

4.1.1.2 Το πλήθος των ατόμων που ακολουθεί ο χρήστης (Followees)

Εκτός από το πλήθος των ατόμων που ακολουθούν τον χρήστη, σημαντικό ρόλο διαδραματίζει και το πλήθος των ατόμων που ακολουθεί ο χρήστης. Σύμφωνα με τους Sotirakou κ.ά. (2019) , είναι ένας άλλος τρόπος για να προσδιοριστεί η δημοτικότητα του συγγραφέα. Ο συγγραφέας ακολουθώντας και αλληλεπιδρώντας με άλλους χρήστες δημιουργεί σχέσεις μαζί τους κι έτσι κερδίζει φήμη. Επιπλέον, μπορεί να δημιουργήσει με τους άλλους αρθρογράφους σχέσεις ανταποδοτικές, αφού διαβάζοντας τις ιστορίες τους και δηλώνοντας ότι του αρέσουν αυτές μέσω των «χειροκροτημάτων» του Medium, υπάρχει το ενδεχόμενο σε βάθος χρόνου να υιοθετήσουν και εκείνοι μια αντίστοιχη συμπεριφορά ως προς τις δημοσιεύσεις του. Ωστόσο, οι De Veirman, Cauberghe και Hudders (2017), δίνουν επιπλέον, μια άλλη διάσταση στον αριθμό των ατόμων που ακολουθεί κάποιος χρήστης, πέρα από αυτή της δικτύωσης και του ανταλλάγματος. Σύμφωνα με την έρευνα τους, το πλήθος των ακολουθούμενων, φαίνεται να είναι ένα χαρακτηριστικό που δεν έχει σχέση απαραίτητα άμεση με τη δημοτικότητα του περιεχομένου που δημοσιεύει ο χρήστης, αλλά έμμεση. Αυτό που κρίνεται σημαντικότερο είναι η αναλογία των ατόμων που ακολουθούν το χρήστη, προς εκείνους που ακολουθεί ο ίδιος. Πιο συγκεκριμένα, σημασία έχει το πλήθος των πρώτων να είναι μεγαλύτερο από το πλήθος των δεύτερων. Ένας βασικός κανόνας που ορίζει το κύρος ενός προφίλ στα μέσα κοινωνικής δικτύωσης είναι η διαφορά των δύο αυτών μεγεθών να είναι θετική . Στη μελέτη τους οι De Veirman κ.ά. (2017), δεν εμβαθύνουν περισσότερο στο πόσο μεγάλη θα πρέπει να είναι αυτή η διαφορά, ωστόσο φαίνεται ότι θα πρέπει να είναι αρκετά μεγάλη, ώστε να διασφαλιστεί ότι ο χρήστης διαθέτει κάποιο κύρος και ότι δεν ακολουθεί όποιον τον ακολουθεί – και το αντίστροφο – ανταποδοτικά.

4.1.2 Στυλ συγγραφής

Ο τρόπος, με τον οποίο ένας συγγραφέας συντάσσει τις δημοσιογραφικές του ιστορίες, είναι σε θέση να επηρεάσει τελικά την δημοτικότητα που μπορεί αυτές να αποκτήσουν, στην πλατφόρμα που δημοσιεύονται. Στα πλαίσια της παρούσας εργασίας, θα μελετηθούν πέντε χαρακτηριστικά που σχετίζονται με το στυλ συγγραφής.

4.1.2.1 Ο προσωπικός τόνος

Το χαρακτηριστικό αυτό αφορά τη χρήση προσωπικού τόνου από την πλευρά του συγγραφέα είτε για να αναφερθεί στον εαυτό του, είτε για να απευθυνθεί και να επικοινωνήσει με τους αναγνώστες του, μέσα στο άρθρο. Πρόκειται για ένα χαρακτηριστικό, το οποίο αξίζει να μελετηθεί, σύμφωνα με το μοντέλο που ανέπτυξαν οι Sotirakou κ.ά. (2019), μιας και αποτελεί ένα ισχυρό όπλο στη φαρέτρα των ειδικών στην επιστήμη του μάρκετινγκ και της επικοινωνίας. Το προσωποποιημένο ύφος μπορεί να αποτελέσει μια πολύ καλή επικοινωνιακή τακτική, αφού συμβάλει θετικά στην ανάπτυξη μιας καλής εικόνας για το προϊόν (με την ευρύτερη έννοια του όρου «προϊόν» στο τομέα του μάρκετινγκ) από το κοινό. Επιπλέον, εκφράσεις που απευθύνονται άμεσα στο πρόσωπο του αναγνώστη, μπορούν να χρησιμοποιηθούν και στο γραπτό λόγο, ως μια τεχνική προσέλκυσης του ενδιαφέροντος του αναγνωστικού κοινού. Σε ένα αγγλικό κείμενο, όπως είναι ένα άρθρο στην πλατφόρμα του Medium, το χαρακτηριστικό του προσωπικού τόνου είναι εύκολο να υπολογιστεί, λαμβάνοντας υπόψη τις προσωπικές αντωνυμίες στην αγγλική γλώσσα. Η Sotirakou κ.ά. (2019), στη μελέτη τους ερευνήσαν τη χρήση των προσωπικών αντωνυμιών:

- I, you, he, she, it, we, they
- Myself, yourself, himself, herself, itself, ourselves, yourselves, themselves
- Mine, yours, his, hers, its, ours, theirs

Ωστόσο, αξίζει να δοθεί στο συγκεκριμένο χαρακτηριστικό και μια άλλη πτυχή. Σύμφωνα με την έρευνα των Louis και Nenkova (2013), η χρήση του προσωπικού τόνου από το συγγραφέα μπορεί να συνδεθεί με την ποιότητα του άρθρου. Αυτή είναι μια ιδιαίτερα σημαντική παρατήρηση μιας και σύμφωνα με τους Keneshloo, Wang, Han και Ramakrishnan (2016) ένας από τους παράγοντες που διαμορφώνουν την υψηλή δημοτικότητα ενός άρθρου είναι η ποιότητα γραφής.

4.1.2.2 Υποκειμενικότητα

«Ο δημοσιογραφικός λόγος εμφανίζεται αντικειμενικός, χωρίς όμως να είναι, γιατί δε μπορεί εκ των πραγμάτων να είναι». (Χατζησαββίδης, 2000) Η αντικειμενικότητα στη δημοσιογραφία είναι κάτι το οποίο δύσκολα μπορεί να οριστεί αλλά και να επιτευχθεί. Σύμφωνα με την άποψη των Κόβατς και Ρόζενταλ (2004) η αντικειμενικότητα στην ουσία θα πρέπει να είναι ο τελικός στόχος, τον οποίο καλείται ο δημοσιογράφος να επιτύχει, έχοντας διασχίσει ένα πολύ δύσβατο μονοπάτι. Πολλές φορές η αντικειμενικότητα συγχέεται με την τιμιότητα και την ισορροπία ανάμεσα στις δύο πλευρές, ωστόσο και πάλι αυτό δε μπορεί να δώσει μια σωστή προσέγγιση της αλήθειας. Η διατήρηση της ισορροπίας αποτελεί ξεκάθαρα ένα στοιχείο υποκειμενικότητας. Ωστόσο, για να καταφέρουν σύμφωνα με τον Χατζησαββίδη (2000), οι δημοσιογράφοι να πείσουν το κοινό για την αντικειμενικότητα των όσων ισχυρίζονται, χρησιμοποιούν μια σειρά από τεχνικές και μέσα που επηρεάζουν τη μορφή, το λεξιλόγιο, τη σύνταξη αλλά και τη δομή του δημοσιογραφικού τους λόγου. Παρ' όλ' αυτά, ο ίδιος έχει οδηγηθεί σε μια παρατήρηση για την ύπαρξη της υποκειμενικότητας μέσα σε ένα δημοσιογραφικό κείμενο. Ισχυρίζεται, λοιπόν, ότι η ύπαρξη των επιθέτων «σπουδαίος», «αξιόλογος», «μεγάλος» καθιστά το δημοσιογραφικό λόγο υποκειμενικό. Προς αυτή την κατεύθυνση έχει κινηθεί και το μοντέλο των Sotirakou κ.ά. (2019), για τη μέτρηση της υποκειμενικότητας μέσα σε ένα κείμενο, χρησιμοποιώντας την βιβλιοθήκη TextBlob στη γλώσσα προγραμματισμού Python . Η βιβλιοθήκη αυτή δε λαμβάνει υπόψη μόνο τις προαναφερθείσες λέξεις, αλλά ένα ευρύ φάσμα λέξεων και φράσεων, στις οποίες μάλιστα αντιστοιχίζει μια αριθμητική τιμή από το μηδέν έως το ένα, με τις τιμές όσο το δυνατόν πιο κοντά στο μηδέν να αντιπροσωπεύουν την «απόλυτη» αντικειμενικότητα και τις τιμές πολύ κοντά στο ένα, την «πλήρη» υποκειμενικότητα. Για παράδειγμα, σύμφωνα με την TextBlob βιβλιοθήκη (TextBlob Sentiment: Calculating Polarity and Subjectivity, 2015), η αγγλική λέξη «great», που σημαίνει σπουδαίος έχει σκορ υποκειμενικότητας 0.75, ενώ η φράση «very great» 0.9750000000000001, μια τιμή πολύ κοντά στο 1 που φανερώνει μεγάλη υποκειμενικότητα στο λόγο του συγγραφέα.

4.1.2.3 Τροπικότητα (Επιστημική)

Η τροπικότητα είναι ένα χαρακτηριστικό του λόγου, μέσω του οποίου γίνεται αντιληπτός ο τρόπος με τον οποίο ο συγγραφέας εκφράζει τη στάση ή τη γνώμη του σε σχέση με το περιεχόμενο των όσων γράφει. (Piqué, Posteguillo και Andreu-Besó, 2001) Είναι ένα

σημαντικό γλωσσικό εργαλείο για την κατανόηση της διαπροσωπικής λειτουργίας και της έκφρασης των κοινωνικών ρόλων μεταξύ του συγγραφέα και του αναγνώστη. (Iwamoto, 2007) Η τροπικότητα μπορεί να είναι είτε επιστημική, είτε δεοντική.

Η επιστημική τροπικότητα υποδεικνύει το βαθμό της βεβαιότητας του συγγραφέα για τα όσα παρουσιάζει στο λόγο του και αφορά σε ένα ευρύ φάσμα σημασιών, του οποίου το ένα άκρο είναι η υπόθεση και το άλλο η πλήρης βεβαιότητα. Ορισμένες από τις σημασίες αυτού του φάσματος είναι η υπόθεση, η δυνατότητα, η πιθανότητα και η βεβαιότητα. (Χατζησαββίδης και Χατζησαββίδου, 2016)

Η δεοντική τροπικότητα υποδεικνύει την προσδοκία της πραγματοποίησης από την πλευρά του συγγραφέα για τα όσα παρουσιάζει στο λόγο του και αφορά σε ένα ευρύ φάσμα σημασιών, του οποίου το ένα άκρο είναι η απλή επιθυμία και το άλλο η υποχρέωση. Ορισμένες από τις σημασίες αυτού του φάσματος είναι η επιθυμία, η ευχή, η πρόθεση και η υποχρέωση. (Χατζησαββίδης και Χατζησαββίδου, 2016)

Και οι δύο αυτές μορφές τροπικότητας μπορούν να υπολογιστούν με παρόμοιο τρόπο. Στην Αγγλική γλώσσα, στην οποία είναι γραμμένα τα άρθρα του Medium, τα οποία λαμβάνονται υπόψη στην παρούσα έρευνα, η τροπικότητα υπολογίζεται μέσω των τροπικών ρημάτων της αγγλικής (modal verbs) που συναντώνται στο άρθρο. Επιπλέον, αξίζει να σημειωθεί ότι κάθε τροπικό ρήμα δεν ανήκει στην ίδια βαθμολογική βαθμίδα. Πιο συγκεκριμένα:

- τα ρήματα must, should, ought to, need to, has to, is to, έχουν υψηλή τιμή τροπικότητας,
- τα ρήματα will, would, shall, έχουν μεσαία τιμή τροπικότητας, ενώ
- τα ρήματα may, might, can, could, έχουν μικρή τιμή τροπικότητας.

Εκτός από τα τροπικά ρήματα, για τον υπολογισμό της συνολικής τροπικότητας ενός κειμένου, διαδραματίζουν σημαντικό ρόλο και κάποια επιρρήματα, επίθετα, ουσιαστικά και ρήματα που μπορούν να υποδηλώσουν το βαθμό της βεβαιότητας ή την προσδοκία πραγματοποίησης. (Iwamoto, 2007)

Στα πλαίσια της έρευνας των Sotirakou κ.ά. (2019), δίνεται μεγαλύτερη έμφαση στην επιστημική τροπικότητα, η οποία μάλιστα μπορεί να υπολογιστεί εύκολα μέσα από τη γλώσσα προγραμματισμού Python και του εργαλείου Pattern στα Αγγλικά (Smedt και Daelemans, 2012). Προς αυτή την κατεύθυνση θα κινηθεί και η παρούσα έρευνα, ώστε να εξακριβωθεί ο τρόπος που επηρεάζει ο βαθμός της βεβαιότητας του συγγραφέα την ικανοποίηση του αναγνωστικού κοινού. Το φάσμα που αντιπροσωπεύει τις διάφορες

διαβαθμίσεις βεβαιότητας αποτυπώνεται σαν ένα αριθμητικό σύνολο τιμών με τιμές από το -1 έως το 1, με το κάτω άκρο να δηλώνει την υπόθεση και το πάνω άκρο την απόλυτη σιγουριά. Πιο συγκεκριμένα, μια τιμή τροπικότητας μεγαλύτερη του +0.5 αντιπροσωπεύει ότι ο συγγραφέας εκφράζει σιγουριά στη δημοσιογραφική του ιστορία και ενδεχομένως γίνεται αντιληπτό πως πρόκειται για κάποιο γεγονός παρά για μια απλή υπόθεση ή πιθανότητα.

4.1.2.4 «Πλούσιο» λεξιλόγιο

Η χρήση πλούσιας φρασεολογίας και η επιλογή ενός μη κοινότυπου λεξιλογίου σε ένα δημοσιογραφικό άρθρο μπορεί να επιδράσει θετικά ως προς την άποψη που θα αποκομίσει ο αναγνώστης, αλλά και την ικανοποίηση που θα αισθανθεί διαβάζοντας το συγκεκριμένο άρθρο. Σύμφωνα με αρκετές έρευνες που έχουν γίνει στον τομέα της εκπαίδευσης, όπως αναφέρουν και στη μελέτη τους οι Louis και Nenkonva (2013), η επιλογή του λεξιλογίου αλλά και ο τρόπος διατύπωσης του γραπτού λόγου διαδραματίζουν σημαντικό ρόλο στον τρόπο που βαθμολογούνται οι φοιτητές από τους καθηγητές τους αλλά και στο πως αξιολογούνται από έμπειρους αναγνώστες γενικότερα. Λαμβάνοντας υπόψη τις συγκεκριμένες έρευνες θα μπορούσε κανείς να χρησιμοποιήσει το συγκεκριμένο χαρακτηριστικό του στυλ συγγραφής και για τη μελέτη της περίπτωσης των δημοσιογραφικών ιστοριών που δημοσιεύονται στην πλατφόρμα του Medium. Στα πλαίσια της έρευνάς τους οι Sotirakou κ.ά. (2019), χρησιμοποιούν σαν μέθοδο ανίχνευσης του «πλούσιου» λεξιλογίου, τον υπολογισμό του ποσοστού των σπάνιων λέξεων σε σχέση με τις συνολικές λέξεις του άρθρου, θεωρώντας ως σπάνιες λέξεις εκείνες που εμφανίζονται με συχνότητα μικρότερη του ενός δευτέρου ($1/2$) της μέσης συχνότητας εμφάνισης των λέξεων του αγγλικού λεξιλογίου.

4.1.2.5 Ευκολία ανάγνωσης

Μια αρκετά πρώιμη, αλλά με σημαντική επιρροή έρευνα (Hutto και Gilbert, 2014) στην αναγνωσιμότητα των άρθρων, ξεκίνησε με την υπόθεση ότι τα άρθρα που θα καταφέρουν να κερδίσουν το ενδιαφέρον του κοινού, είναι εκείνα που είναι εύκολα στην ανάγνωση. Το χαρακτηριστικό αυτό αποτελεί μια μετρική τόσο στον τομέα της δημοσιογραφίας, όσο και σε αυτούς του μάρκετινγκ, της επικοινωνίας και της εκπαίδευσης. Στον τομέα της δημοσιογραφίας, στον οποίο και επικεντρώνεται η παρούσα έρευνα, είναι σημαντικό για την επιτυχία τους, τα άρθρα να είναι προσβάσιμα προς όλες τις εκπαιδευτικές βαθμίδες καθώς επίσης και εύκολα κατανοητά. Για τη

μέτρηση της αναγνωσιμότητας, στο μοντέλο που ανέπτυξαν οι Sotirakou κ.ά. (2019) χρησιμοποιούν το «Flesch–Kincaid grade level» τεστ (Flesch, 1948). Το επίπεδο του βαθμού αναγνωσιμότητας υπολογίζεται από μια μαθηματική φόρμουλα, η οποία λαμβάνει υπόψη:

- Το μέσο πλήθος λέξεων ανά πρόταση και
- το μέσο πλήθος συλλαβών ανά λέξη.

Το αποτέλεσμα αυτής της φόρμουλας αφορά στο επίπεδο της εκπαιδευτικής κλίμακας της Αμερικής και οι τιμές που μπορεί να αποδώσει συνηθίζεται να είναι πάνω από 0 και μέχρι 18, με τις μικρότερες τιμές να αφορούν σε εύκολα στην ανάγνωση κείμενα. Πιο συγκεκριμένα, κείμενα με τιμές από μηδέν έως τρία (0-3) είναι εύκολα κατανοητά ακόμη και από ανθρώπους που είναι στα αρχικά στάδια εκμάθησης ανάγνωσης, ενώ τιμές από 15-18 συνήθως αφορούν σε ακαδημαϊκές δημοσιεύσεις. Αξίζει να σημειωθεί ότι μπορεί να υπάρχουν και τιμές εκτός του διαστήματος μηδέν έως δεκαοκτώ, σε ακραίες περιπτώσεις.

4.2 Περιεχόμενο

Το περιεχόμενο ενός άρθρου αφορά τόσο τα δομικά χαρακτηριστικά του όσο και τα χαρακτηριστικά που προκύπτουν ύστερα από τη συναισθηματική ανάλυση του κειμένου.

4.2.1 Δομή του άρθρου

Το ίδιο το κείμενο αλλά και οι εικόνες, οι τίτλοι και οι υπερσύνδεσμοι μπορούν να επηρεάσουν τη δημοτικότητα που μπορεί να συγκεντρώσει ένα άρθρο.

4.2.1.1 Η έκταση του άρθρου

Ένα από τα πέντε πρώτα πράγματα που μπορεί να δει ένας χρήστης, πριν διαβάσει μια δημοσιογραφική ιστορία στο Medium, είναι ο χρόνος ανάγνωσης. Πρόκειται για μια εκτίμηση, η οποία προκύπτει από το πλήθος των λέξεων και το πλήθος των εικόνων που περιέχονται στο κείμενο του άρθρου. Σύμφωνα με το Medium, ο χρόνος ανάγνωσης βασίζεται στη μέση ταχύτητα ανάγνωσης ενός ενήλικα (περίπου 275 WPM). Το συνολικό πλήθος των λέξεων μεταφράζεται σε λεπτά, ενώ για κάθε εικόνα

προστίθενται δώδεκα δευτερόλεπτα στο συνολικό χρόνο. Σύμφωνα με την έρευνα του Mission.org (McCready, 2016), η οποία μάλιστα δημοσιεύτηκε στην πλατφόρμα του medium.com, ο χρόνος ανάγνωσης διαδραματίζει σημαντικό ρόλο στη δημοτικότητα μιας δημοσιογραφικής ιστορίας. Ωστόσο, επειδή το χαρακτηριστικό αυτό δε μπορεί να υπολογιστεί εκ των προτέρων από τον συγγραφέα, στα πλαίσια της συγκεκριμένης έρευνας, θα μελετηθούν ξεχωριστά η έκταση του άρθρου και το πλήθος των εικόνων.

4.2.1.2 Η έκταση του τίτλου

Το Medium είναι μια διαδικτυακή πλατφόρμα στην οποία δημοσιεύεται μεγάλος όγκος πληροφορίας καθημερινά. Εύκολα, λοιπόν, γίνεται αντιληπτό ότι οι χρήστες δεν μπορούν να διαβάσουν όλη αυτή την πληροφορία, αλλά επιλέγουν με βάση κάποια άλλα χαρακτηριστικά τα άρθρα που θα αναγνώσουν. Ένα από αυτά, ίσως και το σημαντικότερο είναι ο τίτλος, ο οποίος αποτελεί τη σύνοψη του άρθρου. Οι τίτλοι των δημοσιογραφικών ιστοριών, όπως αναφέρει η Θεοδώρα Τσελίγκα-Γκαζιάνη στο συλλογικό έργο που επιμελήθηκε ο Περικλής Πολίτης (2008), είναι «αυτόνομα και νοηματικά συμπυκνωμένα μικροκείμενα, τα οποία διακρίνονται από το υπόλοιπο δημοσιογραφικό κείμενο ως προς τα γλωσσικά χαρακτηριστικά, τις λειτουργίες και τη μορφή τους». Σε έρευνα που πραγματοποιήθηκε μάλιστα για τους τίτλους της ελληνικής εφημερίδας «Έθνος», στην έντυπη και στη διαδικτυακή έκδοση, παρατηρήθηκε σημαντική διαφορά μεταξύ της έκτασής αυτών ανάμεσα στα δύο μέσα (εφημερίδα και διαδικτυακός ιστότοπος). Είναι εμφανές, λοιπόν, πως η έκταση του τίτλου ενός άρθρου εξαρτάται σε ένα βαθμό από το μέσο στο οποίο δημοσιεύεται. Στην περίπτωση ενός ηλεκτρονικού μέσου, με μεγάλο όγκο πληροφορίας, που οι αναγνώστες είναι κυρίως «σαρωτές τίτλων», φαίνεται πως λειτουργούν καλύτερα οι πιο σύντομοι τίτλοι. (Πολίτης, 2018).

4.2.1.3 Οι εικόνες

Όπως αναφέρθηκε και προηγουμένως, το πλήθος των εικόνων, είναι ένα χαρακτηριστικό, το οποίο επηρεάζει σε ένα βαθμό την απήχηση που πρόκειται να έχει ένα άρθρο στο Medium, αφού αποτελεί βασικό κομμάτι του υπολογισμού του χρόνου ανάγνωσης (McReady, 2016). Επιπλέον, η σύγχρονη μορφή της δημοσιογραφίας έχει ως βασικό χαρακτηριστικό γνώρισμα την χρήση πολυμέσων στις διαδικτυακές πλατφόρμες ενημέρωσης.(Πολίτης, 2008). Πολλές φορές μάλιστα, όταν γίνεται λόγος για τον τομέα του μάρκετινγκ και τις διαφημιστικές εκστρατείες, είναι πολύ σημαντικό να

διατηρείται μια «συγκεκριμένη» αναλογία σε σχέση με το πλήθος των λέξεων και το πλήθος των εικόνων. (Sotirakou κ.ά., 2019) Έτσι, ως απόρροια των παραπάνω, στα πλαίσια της παρούσας έρευνας θα χρειαστεί να μελετηθεί το χαρακτηριστικό των εικόνων με δύο τρόπους:

- Λαμβάνοντας υπόψη το συνολικό πλήθος τους σε ένα άρθρο
- Λαμβάνοντας υπόψη την αναλογία του συνολικού πλήθους τους σε σχέση με το συνολικό πλήθος λέξεων του άρθρου.

Αξίζει να σημειωθεί, ότι με τον ίδιο διπτό τρόπο θα μελετηθεί και η συμβολή του χαρακτηριστικού του πλήθους των λέξεων, στην έρευνα για τη δημοτικότητα των δημοσιογραφικών άρθρων του Medium.

4.2.1.4 Η μορφή λίστας σημείων (Bullet Points)

Αυτό το χαρακτηριστικό λαμβάνει υπόψη την ύπαρξη λίστας σημείων μέσα σε μία δημοσιογραφική ιστορία. Πιο συγκεκριμένα εξετάζεται αν το άρθρο (εξ' ολοκλήρου ή κάποιο μέρος των πληροφοριών που περιλαμβάνει) είναι δομημένο σαν λίστα. Η συγκεκριμένη μορφή άρθρου, είναι μια μορφή, η οποία είναι ιδιαίτερα δημοφιλής, αν μετρήσει κανείς την έννοια της δημοτικότητας από την πλευρά των κοινοποιήσεων των άρθρων (Okrent, 2014). Τέτοιου είδους άρθρα φαίνεται πως έχουν υψηλή δημοτικότητα λόγω της διαδοχικότητας, της λογικής σειράς και του ξεκάθਾਰου τρόπου παρουσίασης των πληροφοριών. Ωστόσο, επειδή στην περίπτωση της συγκεκριμένης διπλωματικής εργασίας, η δημοτικότητα δεν ορίζεται με γνώμονα τις κοινοποιήσεις του άρθρου, αλλά με τα χειροκροτήματα που συγκεντρώνει στην πλατφόρμα του Medium, η μορφή λίστας σημείων είναι ένα χαρακτηριστικό το οποίο έχει ιδιαίτερο ενδιαφέρον να μελετηθεί σε σχέση με το αν και πως επιδρά με τα υπόλοιπα χαρακτηριστικά για την ανάδειξη των δημοφιλών δημοσιογραφικών ιστοριών.

4.2.1.5 Οι υπερσύνδεσμοι

Ένα άρθρο που δημοσιεύεται σε μια ηλεκτρονική πλατφόρμα ενδέχεται, εκτός από το κείμενο και τα πολυμέσα, να περιέχει και ορισμένους συνδέσμους προς εξωτερικούς ιστοχώρους ή προς άλλα άρθρα της πλατφόρμας στην οποία δημοσιεύεται. Η ύπαρξη υπερσυνδέσμων προς εξωτερικούς ιστοχώρους μπορεί να σχετίζεται με την παροχή συμπληρωματικών πληροφοριών καθώς και με την αναφορά επαληθευμένων πηγών. Επιπροσθέτως, η παροχή επιπλέον υλικού μπορεί να αποτελέσει ένδειξη εκτεταμένης

δημοσιογραφικής προσπάθειας και κατ' επέκταση να επιδράσει θετικά ως προς την εικόνα που θα σχηματίσει ο αναγνώστης για τον συγγραφέα. (Sotirakou κ.ά., 2019) Από την άλλη πλευρά, το μεγάλο πλήθος υπερσυνδέσμων σε μια δημοσιογραφική ιστορία, μπορεί να αποπροσανατολίσει τον αναγνώστη και να τον οδηγήσει σε απώλεια συνοχής της αφήγησης του αρθρογράφου. (Πολίτης, 2018). Στα πλαίσια, της συγκεκριμένης έρευνας και θεωρώντας για το συγκεκριμένο χαρακτηριστικό σημαντικότερη την παρουσία παρά το πλήθος του, η προσοχή θα στραφεί στην ύπαρξη ή στη μη ύπαρξη υπερσυνδέσμων μέσα σε ένα άρθρο στην πλατφόρμα του Medium.

4.2.2 Συναισθηματική Ανάλυση

Πέρα από τα δομικά χαρακτηριστικά, μπορεί να γίνει και η συναισθηματική ανάλυση του άρθρου, ώστε να αποκαλυφθούν τα συναισθήματα που προκαλούνται κατά την ανάγνωσή του.

4.2.2.1 Συναισθήματα

Η ανάγνωση ενός δημοσιογραφικού άρθρου – όπως και ενός κειμένου γενικότερα - μπορεί να προκαλέσει στον αναγνώστη ποικίλα συναισθήματα. Υπάρχουν μελέτες που αποδεικνύουν ότι το διαδικτυακό περιεχόμενο (οποιασδήποτε μορφής όπως ειδήσεις, άρθρα, πολυμέσα κλπ), το οποίο προκαλεί στους αναγνώστες κυρίως τα συναισθήματα του θυμού, του άγχους και του δέους, κατέχει μεγαλύτερες πιθανότητες να γίνει δημοφιλές. (Sotirakou κ.ά., 2019) Για τη μελέτη της περίπτωσης του Medium, θα εξεταστούν οχτώ βασικά συναισθήματα για την ύπαρξή τους, αλλά και τη συμβολή τους ή όχι στην υψηλή δημοτικότητα, η οποία προσμετρείται ανάλογα με το πλήθος των χειροκροτημάτων που συγκεντρώνουν τα άρθρα στη συγκεκριμένη πλατφόρμα. Τα συναισθήματα αυτά είναι:

- Χαρά
- Θλίψη
- Φόβος
- Θυμός
- Εμπιστοσύνη
- Έκπληξη
- Αηδία
- Προσμονή

Για τον υπολογισμό αυτών των οχτώ βασικών συναισθημάτων λαμβάνεται υπόψη το πλήθος των συναισθηματικά φορτισμένων λέξεων του εκάστοτε συναισθήματος, σε σχέση με το συνολικό πλήθος λέξεων που περιέχει το άρθρο. Οι συναισθηματικά φορτισμένες λέξεις προκύπτουν από τη σουίτα λεξικών NRC (Mohammad, Kiritchenko και Zhu, 2013) και τα οχτώ συναισθήματα βασίζονται στο μοντέλο του Plutchik. Με βάση τα παραπάνω, στα πλαίσια της παρούσας έρευνας θα εξεταστεί για κάθε συναίσθημα αν και σε πόσο μεγάλο ποσοστό (επί τοις εκατό) συμβάλλει στην επιτυχία των άρθρων στο medium.com.

4.2.2.2 Πολικότητα του άρθρου

Μέσω του χαρακτηριστικού αυτού, εξετάζεται αν ένα άρθρο χαρακτηρίζεται από θετικότητα, αρνητικότητα ή ουδετερότητα. Υπάρχουν διάφορες μελέτες, τις οποίες παρουσιάζουν συγκεντρωτικά στην έρευνα τους οι Sotirakou κ.ά. (2019) που σχετίζουν την πολικότητα, και πιο συγκεκριμένα την αρνητικότητα, με την επιτυχία ενός άρθρου. Οι αναγνώστες συνήθως έλκονται από τις αρνητικές ειδήσεις, γεγονός που σχετίζεται με την ανάγκη τους για επιβίωση, την ασφάλειά τους, αλλά και την αποφυγή δυσάρεστων συμβάντων στο εγγύτερο περιβάλλον τους. Επιπλέον, τόσο η θετικότητα όσο και η αρνητικότητα, διαδραματίζουν σημαντικό ρόλο στο διαμοιρασμό των άρθρων στα μέσα κοινωνικής δικτύωσης, ωστόσο μόνο η αρνητικότητα είναι αυτή που συμβάλλει περισσότερο (σε σχέση με τη θετικότητα και την ουδετερότητα) στην απόκτηση μεγάλου πλήθους αντιδράσεων. Στην μελέτη του Medium, ως ένδειξη της δημοτικότητας αναλογίζεται το πλήθος των αντιδράσεων που κερδίζει ένα άρθρο, συνεπώς οι τρεις μορφές πολικότητας θα αποτελέσουν χαρακτηριστικά, τα οποία αξίζει να μελετηθούν σχετικά με το πόσο συμμετέχουν στην επιτυχία του περιεχομένου της συγκεκριμένης πλατφόρμας. Σε αντιστοιχία με τα οκτώ διαφορετικά συναισθήματα, οι τρεις διαφορετικές μορφές πολικότητας θα εξεταστούν ως τρία διαφορετικά χαρακτηριστικά του άρθρου και για κάθε ένα από αυτά θα μελετηθεί σε ποιο ποσοστό συναντάται μέσα στο κείμενο, στην περίπτωση που συναντάται. Για τον υπολογισμό των τριών ειδών πολικότητας θα χρησιμοποιηθεί το λεξικό Vader (Hutto και Gilbert, 2014), το οποίο αποδίδει κατευθείαν αυτές τις τιμές στη γλώσσα προγραμματισμού Python, ενώ οι τιμές που αποδίδονται μπορεί να ξεκινάνε από μηδέν και να φτάνουν έως τη μονάδα. Όπως είναι εμφανές, αφορούν το ποσοστό της κάθε μορφής πολικότητας που συναντάται μέσα στο άρθρο. Επιπλέον, εκτός από την ξεχωριστή μέτρηση για κάθε μορφή πολικότητας, αξίζει να παρουσιαστεί και η συνολική πολικότητα του άρθρου. Πρόκειται για ένα χαρακτηριστικό, το οποίο εύκολα υπολογίζεται μέσω της γλώσσας

προγραμματισμού Python και της βιβλιοθήκης TextBlob, και το οποίο δίνει σε ένα διάστημα τιμών από το μείον ένα (-1) έως το ένα (1) την πολικότητα του άρθρου, με τις τιμές κοντά στο κάτω άκρο να δηλώνουν την αρνητικότητα, κοντά στο άνω άκρο τη θετικότητα και στη μέση τιμή την ουδετερότητα.

4.2.2.3 Πυκνότητα συναισθημάτων

Η πυκνότητα συναισθημάτων προσδιορίζει το συνολικό συναισθηματικό βάρος που περιέχει ένα κείμενο. (Varshney και Wagh, 2017) Πρόκειται για έναν διαφορετικό τρόπο μέτρησης του συναισθήματος ενός κειμένου. Υπάρχουν αρκετοί τρόποι με τους οποίους μπορεί να υπολογιστεί το συνολικό συναισθηματικό βάρος που φέρει μια δημοσιογραφική ιστορία. Στην περίπτωση της παρούσας έρευνας, για τον υπολογισμό του συγκεκριμένου χαρακτηριστικού λαμβάνεται υπόψη το άθροισμα των θετικά και των αρνητικά φορτισμένων λέξεων, αναλογικά με το συνολικό πλήθος λέξεων του άρθρου. Οι θετικά και αρνητικά φορτισμένες λέξεις προκύπτουν από τη σουίτα λεξικών NRC (Mohammad κ.ά. ,2013) και το μοντέλο του Plutchik. Το χαρακτηριστικό της συναισθηματικής πυκνότητας εκφράζεται με ένα ποσοστό επί τοις εκατό, με παρόμοιο τρόπο που εκφράζονται και τα οχτώ βασικά συναισθήματα.

4.2.2.4 Πολικότητα του τίτλου του άρθρου

Όπως έχει ήδη αναφερθεί (Πολίτης, 2008), οι αναγνώστες άρθρων από ψηφιακές πλατφόρμες λειτουργούν περισσότερο ως σαρωτές τίτλων. Εκτός από την έκταση του τίτλου, η πολικότητά του ενδέχεται να διαδραματίζει σημαντικό ρόλο στην επιλογή του αναγνώστη για το αν θα το διαβάσει ή όχι. Ένα άρθρο, του οποίου, ο τίτλος δεν είναι ελκυστικός δύσκολα θα γίνει δημοφιλές ακόμα κι αν πληροί ορισμένα κριτήρια που εγγυώνται την ποιότητα και εξασφαλίζουν την επιτυχία. Στην περίπτωση των τίτλων, και αυτό αποτελεί κυρίως μια επιλογή εξαιτίας της μικρής έκτασής τους, η προσοχή θα στραφεί στη συνολική πολικότητα, κι όχι ξεχωριστά στις τρεις μορφές της. Συνεπώς, για την πλατφόρμα του Medium, στους τίτλους θα αντιστοιχιστεί μια τιμή από το μείον ένα (-1) έως το ένα (1), ώστε να εξεταστεί αν τελικά παίζει ρόλο και με ποιο τρόπο η πολικότητα του τίτλου ενός άρθρου.

4.3 Γενικό Πλαίσιο

Υπάρχουν χαρακτηριστικά που δεν αφορούν ούτε τον συγγραφέα, αλλά ούτε και το περιεχόμενο του άρθρου. Τέτοια μπορούν να είναι το είδος που ανήκει το άρθρο καθώς και κάποια χωροχρονικά στοιχεία.

4.3.1 Είδος άρθρου

Η δημοτικότητα ενός άρθρου επηρεάζεται και εξαρτάται σε ένα βαθμό, σύμφωνα με τον Arapakis κ.ά (2016), από το είδος στο οποίο ανήκει. Το ενδιαφέρον του αναγνωστικού κοινού, παραδείγματος χάρη είναι μεγαλύτερο όταν το άρθρο αφορά σε διασημότητες και έχει ένα πιο «ψυχαγωγικό» περιεχόμενο, παρά όταν αφορά σε κάποιον επιστημονικό τομέα, όπως η ανθρωπολογία. Επιπλέον, υπάρχουν κάποια είδη άρθρων τα οποία απαιτούν περισσότερη προσπάθεια για να γίνουν κατανοητά. Τέτοια για παράδειγμα μπορεί να είναι κάποια επιστημονικά άρθρα, ειδήσεις που αφορούν στα οικονομικά και το χρηματιστήριο ή άρθρα που απαιτούν την προϋπαρξη τεχνικών γνώσεων από την πλευρά του αναγνώστη. Σε ό,τι αφορά το είδος των άρθρων στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας θα ληφθούν υπόψη έξι βασικές κατηγορίες άρθρων που μπορεί να συναντήσει κανείς κατά την περιήγησή του στον ιστοχώρο του medium.com. Αυτές έχουν να κάνουν με τις ειδήσεις, την τεχνολογία, την υγεία, την εργασία, τα αθλητικά και τον τρόπο ζωής και αναφέρονται με τα ονόματα News, Technology, Health, Business, Sports, και Lifestyle αντίστοιχα. Αξίζει να σημειωθεί πως αυτό είναι το βασικό χαρακτηριστικό που θα καθοριστεί για τη μελέτη των υπόλοιπων χαρακτηριστικών που συμβάλλουν στην επιτυχία ενός άρθρου. Πιο συγκεκριμένα, για κάθε είδος ξεχωριστά θα παραχθούν οι κανόνες που διασφαλίζουν επιτυχημένα άρθρα. Επιπροσθέτως, κάθε είδος από τα προαναφερθέντα είναι εμφανές πως έχει ένα διαφορετικό κοινό – στόχος και για το λόγο αυτό το πλήθος των χειροκροτημάτων που θα ορίζει το ύψος της δημοτικότητας για κάθε ένα από αυτά θα είναι διαφορετικό.

4.3.2 Χρόνος δημοσίευσης

Τα άρθρα που χρησιμοποιούνται στη συγκεκριμένη έρευνα αφορούν ένα χρονικό διάστημα διάρκειας ενός έτους. Τα χαρακτηριστικά που θα εξεταστούν ως προς το κατά πόσο συμβάλλουν στην υψηλή δημοτικότητα ενός άρθρου είναι η ημέρα με τιμές από ένα (1) έως τριανταένα (31), ο μήνας με τιμές από ένα (1) έως δώδεκα (12) και το έτος το οποίο μπορεί να είναι το 2017 ή το 2018, μιας και τα δεδομένα αναφέρονται σε αυτές

τις περιόδους. Αυτό που έχει, ωστόσο, μεγαλύτερη σημασία, και ενδεχομένως να επηρεάζει τη δημοτικότητα των δημοσιογραφικών ιστοριών του Medium είναι η ώρα δημοσίευσης.

4.3.3 Γεωγραφική εγγύτητα

Η γεωγραφική εγγύτητα είναι ακόμη ένα χαρακτηριστικό το οποίο μπορεί να επηρεάσει τη δημοτικότητα των δημοσιογραφικών άρθρων. Στα πλαίσια της μελέτης τους, οι Sotirakou κ.ά. (2019) χρησιμοποιούν το σύστημα Mordecai για να εξάγουν τα τοπωνύμια και να πάρουν τις συντεταγμένες τους. Η έρευνά τους έχει προσανατολιστεί στις Ηνωμένες Πολιτείες Αμερικής, μιας και οι αναγνώστες του Medium προέρχονται από εκεί κατά πλειονότητα. Για την παρούσα μελέτη, θα εξεταστούν επίσης τα τοπωνύμια που ανήκουν στις ΗΠΑ και πιο συγκεκριμένα θα μελετηθεί αν η ύπαρξη ενός τέτοιου τοπωνυμίου επηρεάζει ή όχι τη δημοτικότητα του άρθρου παράλληλα με τα υπόλοιπα χαρακτηριστικά.

5. ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

5.1 Τι είναι η Μηχανική Μάθηση

Η Μηχανική Μάθηση είναι ένας κλάδος της επιστήμης της Τεχνητής Νοημοσύνης, ο οποίος ασχολείται με την ανάπτυξη υπολογιστικών συστημάτων που μπορούν να αποκτούν μιας μορφής γνώση καθώς και να βελτιώνουν την απόδοσή τους με βάση την πρότερη γνώση και την εμπειρία που έχουν αποκτήσει. Σύμφωνα με έναν από τους πιο κλασικούς ορισμούς για το επιστημονικό πεδίο της Μηχανικής Μάθησης, «ένα πρόγραμμα υπολογιστή μαθαίνει από την εμπειρία E , ως προς μια κλάση έργων T (Tasks) και αποδόσεις P (Performance Measures), εάν οι αποδόσεις του στις κλάσεις T , όπως μετριοούνται από τα P , βελτιώνονται με την εμπειρία E ». (Mitchell, 1997) Για να γίνει περισσότερο κατανοητή, όμως, η έννοια της Μηχανικής Μάθησης, αξίζει να μελετηθεί η έννοια της μάθησης γενικότερα. Στον όρο «μάθηση» μπορούν να αποδοθούν ποικίλες σημασίες όπως:

- «Χτίσιμο» γνώσης μέσω μελέτης, πείρας ή διδασκαλίας
- Απόκτηση ενημερότητας μέσω πληροφοριών ή παρατήρησης
- Ικανότητα απομνημόνευσης
- Λήψη και ακολούθηση οδηγιών

Ωστόσο, αυτές οι σημασίες της μάθησης, όταν μιλάμε για υπολογιστές είναι δύσκολο να αξιολογηθούν. Πιο συγκεκριμένα, δεν μπορεί κανείς να αποφανθεί με σιγουριά, για το αν ένας υπολογιστής έχει αποκτήσει όντως γνώση. Για το λόγο αυτό, αντί να ελέγχεται η ικανότητα εκμάθησης των αλγορίθμων, ελέγχεται η ικανότητά τους να απαντούν σε ερωτήματα. Η μάθηση, μπορεί να γίνει αντιληπτή μέσα από την αλλαγή συμπεριφοράς ενός αλγορίθμου στο μέλλον. Η παρατήρηση μιας καλύτερης απόδοσης ενός αλγορίθμου, μπορεί να δείξει ότι ο αλγόριθμος έχει μάθει, όμως η έννοια της μάθησης προϋποθέτει την ύπαρξη σκέψης και σκοπού, δύο έννοιες τις οποίες δεν μπορούν να λάβουν υπόψη οι υπολογιστικοί αλγόριθμοι. (Witten, Frank και Hall, 2005) Έτσι, όταν γίνεται λόγος για μάθηση, στην ουσία πρόκειται για εκπαίδευση των αλγορίθμων με βάση τα δεδομένα. Κατά συνέπεια, ο τομέας της Μηχανικής Μάθησης ασχολείται με την κατασκευή αλγορίθμων ικανών να μαθαίνουν μέσα από τα δεδομένα. Η διαδικασία της εκμάθησης γίνεται σε δύο στάδια. Στο πρώτο στάδιο γίνεται η εκπαίδευση του αλγορίθμου, ενώ στο δεύτερο ελέγχεται η αποδοτικότητα του αλγορίθμου και πιστοποιείται η γνώση που έχει κατακτήσει. Η νέα γνώση που έχει αποκτηθεί μπορεί να χρησιμοποιηθεί για την επίλυση πραγματικών προβλημάτων.

5.2 Είδη Μηχανικής Μάθησης

Μπορεί ο κλάδος της Μηχανικής Μάθησης να απέχει πολύ από την ανάπτυξη υπολογιστικών συστημάτων που να μαθαίνουν τόσο καλά όσο ο άνθρωπος, παρ' όλ' αυτά οι αλγόριθμοι μπορούν να μιμηθούν τον ανθρώπινο τρόπο εκμάθησης και έτσι να κερδίσουν πολύτιμη γνώση. Σε αναλογία με τους τρόπους που μαθαίνουν οι άνθρωποι, διακρίνονται τρία είδη Μηχανικής Μάθησης:

1. Επιτηρούμενη ή Επιβλεπόμενη Μάθηση: Σε αυτό το είδος μάθησης, το υπολογιστικό πρόγραμμα δέχεται παραδειγματικές εισόδους και επιθυμητές εξόδους, με σκοπό να «ανακαλύψει» έναν κανόνα, ο οποίος μπορεί να γενικευτεί και σε άγνωστες εισόδους και να παράξει ύστερα από την εφαρμογή του ορισμένες εξόδους. Χρησιμοποιείται κυρίως για προβλήματα ταξινόμησης (classification), παρεμβολής (regression), πρόβλεψης (prediction) αλλά και διερμηνείας (interpretation)
2. Μη Επιτηρούμενη Μάθηση: Σε αυτό το είδος μάθησης, το υπολογιστικό πρόγραμμα προσπαθεί μόνο του να ανακαλύψει συγκεκριμένα μοτίβα που υπάρχουν στα δεδομένα, χωρίς να λάβει κάποια καθοδήγηση για τις επιθυμητές εξόδους. Χρησιμοποιείται κυρίως για προβλήματα ανάλυσης συσχετισμών (association analysis) και ομαδοποίησης (clustering).
3. Ενισχυτική Μάθηση: Σε αυτό το είδος μάθησης, το υπολογιστικό πρόγραμμα μαθαίνει μέσω άμεσης αλληλεπίδρασης με το περιβάλλον, μέσα στο οποίο καλείται να φέρει εις πέρας έναν στόχο, χωρίς να έχει κάποιου είδους ενημέρωση για το αν είναι κοντά στον στόχο του ή όχι. Χρησιμοποιείται κυρίως για προβλήματα σχεδιασμού (planning).

Αξίζει να αναφερθεί ότι μεταξύ της επιτηρούμενης και της μη επιτηρούμενης, υπάρχει και η ημι-επιτηρούμενη μάθηση, κατά την οποία ένα μέρος από τα δεδομένα εκμάθησης και πιο συγκεκριμένα ένα μέρος εκ των επιθυμητών αποτελεσμάτων αποκρύπτεται από το υπολογιστικό πρόγραμμα. (Γεωργούλη, 2015)

6. ΕΡΕΥΝΑ

6.1 Σκοπός της Έρευνας

Σκοπός της συγκεκριμένης έρευνας είναι η παραγωγή κανόνων συγγραφής διαδικτυακών δημοσιογραφικών άρθρων, με αυξημένη πιθανότητα να καταστούν δημοφιλή στην πλατφόρμα στην οποία ανεβαίνουν.

6.2 Μεθοδολογία

1^ο Στάδιο: Η Συλλογή των Δεδομένων

Για την πραγματοποίηση της έρευνας έχουν χρησιμοποιηθεί πάνω από διακόσιες χιλιάδες άρθρα, τα οποία ελήφθησαν από την πλατφόρμα του Medium. Η επιλογή αυτής της πλατφόρμας έγινε γιατί διαθέτει χαρακτηριστικά τόσο από τον τομέα της δημοσιογραφίας, όσο και από τους τομείς των μέσων κοινωνικής δικτύωσης και της επικοινωνίας. Επιπλέον, οι δημοσιογραφικές ιστορίες του Medium, κατηγοριοποιούνται με βάση το είδος τους και έτσι υπάρχει η δυνατότητα παραγωγής κανόνων, για κάθε είδος ξεχωριστά.

Τα άρθρα που έχουν συλλεχθεί αφορούν ένα χρονικό διάστημα διάρκειας ενός έτους, από τον Σεπτέμβριο του 2017 έως και τον Σεπτέμβριο του 2018 με τη μέθοδο της απόξεσης ή συγκομιδής δεδομένων (data scraping). Η μέθοδος αυτή, αποτελεί μια εξέχουσα τεχνική για την αυτοματοποιημένη συλλογή δεδομένων που βρίσκονται στο διαδίκτυο. Πρόκειται για μία τεχνική, η οποία προσφέρει πολλές δυνατότητες στην έρευνα για τα ψηφιακά και τα κοινωνικά μέσα επικοινωνίας, αφού μέσω αυτής μπορεί να υπάρξει πρόσβαση σε ένα μεγάλο πλήθος δεδομένων, τα οποία περιέχουν σημαντικές πληροφορίες προς επεξεργασία και ανάλυση. (Marres και Weltevrede, 2013) Η μέθοδος που χρησιμοποιήθηκε για την εφαρμογή της τεχνικής της συγκομιδής δεδομένων στην πλατφόρμα του Medium, είναι ένα πρόγραμμα στη γλώσσα προγραμματισμού Python, που αναπτύχθηκε για αυτό το σκοπό.

2^ο Στάδιο: Η Προετοιμασία των Δεδομένων

Τα άρθρα που συλλέχθηκαν αρχικά, βρίσκονταν σε μία πρώιμη ακατέργαστη μορφή, η οποία καθιστούσε την διαχείρισή τους δύσκολη. Για το λόγο αυτό, τα δεδομένα χρειάστηκε να περάσουν από επεξεργασία, ώστε να «καθαριστούν» και να διατηρηθεί μόνο η χρήσιμη πληροφορία. Πιο συγκεκριμένα, αφαιρέθηκαν άρθρα που δεν είναι γραμμένα στην αγγλική γλώσσα, άρθρα για τα οποία υπήρχαν ελλιπείς πληροφορίες,

HTML κώδικας, μη αποδεκτοί χαρακτήρες και μη έγκυρες λέξεις. Και στη συγκεκριμένη φάση η επεξεργασία των δεδομένων, έγινε με κώδικα που γράφτηκε στη γλώσσα προγραμματισμού Python.

3^ο Στάδιο: Η Εξαγωγή των Χαρακτηριστικών

Τα χαρακτηριστικά που προέκυψαν από τη μελέτη της βιβλιογραφίας, με τη βοήθεια βιβλιοθηκών καθώς και κώδικα στη γλώσσα προγραμματισμού της Python, εξήχθησαν από τα επεξεργασμένα δεδομένα του προηγούμενου σταδίου, με σκοπό να αξιοποιηθούν από τον αλγόριθμο Μηχανικής Μάθησης του επόμενου σταδίου.

4^ο Στάδιο: Η Αξιοποίηση της Μηχανικής Μάθησης

Η εύρεση των κανόνων για την παραγωγή δημοφιλών άρθρων, στηρίζεται στην πρόβλεψη της δημοτικότητας των άρθρων στην πλατφόρμα του Medium. Η πρόβλεψη της δημοτικότητας είναι ένα πρόβλημα, το οποίο μπορεί να επιλυθεί με τη βοήθεια της Μηχανικής Μάθησης. Πιο συγκεκριμένα και στα πλαίσια της παρούσας έρευνας, χρησιμοποιήθηκε το είδος της επιβλεπόμενης μηχανικής μάθησης με επαγωγή. Σκοπός των αλγορίθμων αυτού του είδους, είναι η κατάληξη σε συμπεράσματα μέσα από παραδείγματα εισόδων και εξόδων, κάνοντας χρήση του προτασιακού λογισμού. Οι αλγόριθμοι επιτηρούμενης επαγωγικής μάθησης χρησιμοποιούνται κυρίως για προβλήματα ταξινόμησης ή παρεμβολής.

Το πρόβλημα της πρόβλεψης της δημοτικότητας αντιμετωπίστηκε ως πρόβλημα ταξινόμησης. Τα προβλήματα που λύνονται με τη μέθοδο της ταξινόμησης, ουσιαστικά διαχωρίζουν τα δεδομένα σε διακριτές κλάσεις. Έτσι, για την συγκεκριμένη έρευνα, χρειάστηκε πρώτα από όλα να χωριστούν τα άρθρα που έχουν ληφθεί σε δύο κατηγορίες. Πιο συγκεκριμένα, διαιρέθηκαν σε άρθρα υψηλής δημοτικότητας και άρθρα χαμηλής δημοτικότητας, ανάλογα με το πλήθος των χειροκροτημάτων που έχουν συγκεντρώσει. Στον Πίνακα 1, φαίνεται για κάθε είδος, το πλήθος των άρθρων, το μέσο πλήθος χειροκροτημάτων καθώς και η τιμή κατωφλίου, η οποία αποτελεί την κατώτατη τιμή για να θεωρηθεί ότι ένα άρθρο ανήκει στην κατηγορία με τα άρθρα υψηλής δημοτικότητας.

Πίνακας 1: Στατιστικά στοιχεία ανά είδος άρθρου

Είδος Άρθρου	Πλήθος Άρθρων	Μέση Τιμή Χειροκροτημάτων	Τιμή Κατωφλίου
Ειδησεογραφικά	17762	150	195
Τεχνολογικά	30883	236	350
Επιχειρηματικά	35768	65	102
Υγείας	49072	39	53
Αθλητικά	12664	19	49
Τρόπου Ζωής	54561	347	493

Η τιμή κατωφλίου προκύπτει ύστερα από την ταξινόμηση των άρθρων με βάση τα χειροκροτήματα που συγκέντρωσαν και τον ορισμό του 10% εξ' αυτών με τα περισσότερα ως υψηλής δημοτικότητας και του υπόλοιπου 90% ως χαμηλής δημοτικότητας.

Αφού τα δεδομένα χωρίστηκαν σε δύο κλάσεις και αποδόθηκε ένας αριθμός για την κάθε κλάση (μηδέν για τα μη δημοφιλή και ένα για τα δημοφιλή άρθρα), δόθηκαν ως είσοδο σε έναν αλγόριθμο επιβλεπόμενης επαγωγικής μάθησης. Αξίζει να σημειωθεί ότι το 80% των συνολικών άρθρων χρησιμοποιήθηκε για την εκπαίδευση του αλγορίθμου, ενώ το 20% για την αξιολόγηση. Ο αλγόριθμος που επιλέχτηκε ήταν αυτός του δέντρου απόφασης (Decision Tree).

```

1 estimator = DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=7,
2     max_features=None, max_leaf_nodes=None,
3     min_samples_leaf=4, min_samples_split=8, presort=False, random_state=0,
4     splitter='best')

```

```

1 estimator.fit(X_train, y_train)

```

```

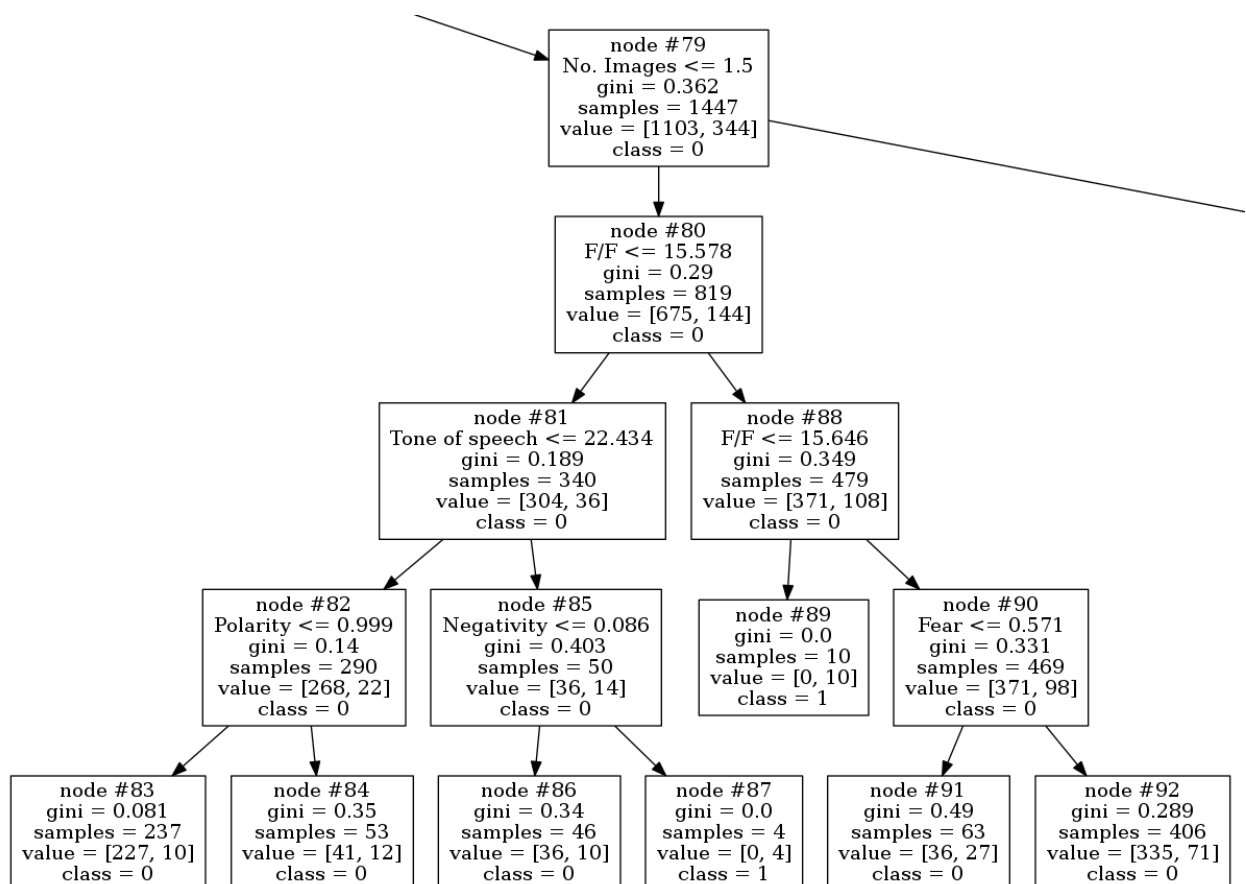
DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=7,
    max_features=None, max_leaf_nodes=None,
    min_impurity_decrease=0.0, min_impurity_split=None,
    min_samples_leaf=4, min_samples_split=8,
    min_weight_fraction_leaf=0.0, presort=False, random_state=0,
    splitter='best')

```

Εικόνα 2: Αλγόριθμος Δένδρου Απόφασης

Τα δέντρα απόφασης αποτελούν ίσως έναν από τους πιο καθιερωμένους αλγορίθμους επιβλεπόμενης επαγωγικής μάθησης για προβλήματα στα οποία απαιτείται ταξινόμηση. Στην περίπτωση των άρθρων του Medium, τα άρθρα μπορούν να ανήκουν σε δύο κατηγορίες, για το λόγο αυτό το δέντρο απόφασης είναι δυαδικό. Η εφαρμογή του αλγορίθμου έχει ως αποτέλεσμα τη δημιουργία ενός δενδρικού σχηματισμού, του οποίου τα φύλλα αντιστοιχούν στις κατηγορίες (κλάσεις) ταξινόμησης. Ο δενδρικός σχηματισμός μπορεί να θεωρηθεί και σαν ένα σύνολο κανόνων ταξινόμησης. (Γεωργούλη, 2015)

Με βάση την παραπάνω θεώρηση, το δέντρο που παράγεται ύστερα από την εφαρμογή του αλγορίθμου μπορεί να οδηγήσει σε ορισμένους κανόνες, σχετικούς με το πως τα άρθρα καταλήγουν στην κάθε κλάση. Οι κανόνες αυτοί προκύπτουν αν ακολουθήσει κανείς το μονοπάτι από τη ρίζα του δένδρου μέχρι και το κάθε φύλλο. Κάθε κόμβος του δένδρου έχει τη μορφή μιας αριθμητικής έκφρασης στην οποία εμπλέκεται ένα χαρακτηριστικό του άρθρου (πχ. Αριθμός εικόνων ≥ 2.5), άρα ο τελικός κανόνας για κάθε μονοπάτι που οδηγεί σε κόμβο – φύλλο, θα περιέχει ένα σύνολο αριθμητικών εκφράσεων που συνδέονται με το λογικό «ΚΑΙ».

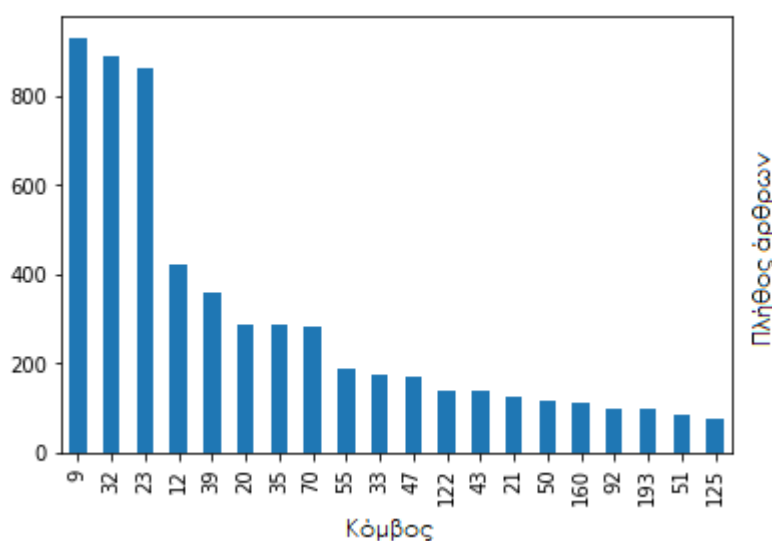


Εικόνα 3: Κόμβοι Δένδρου Απόφασης

Στην Εικόνα 3, παρουσιάζεται ένα μέρος του δενδρικού σχηματισμού που δημιουργήθηκε ύστερα από την εφαρμογή του αλγορίθμου Μηχανικής Μάθησης, για την παραγωγή των κανόνων που αφορούν στα άρθρα επιχειρηματικού περιεχομένου. Επειδή το δένδρο απόφασης είναι δυαδικό, υπάρχουν δύο ενδεχόμενα για κάθε συνθήκη που ελέγχεται σε κάθε κόμβο – μη φύλλο. Έτσι, αν ακολουθηθεί το μονοπάτι προς τα αριστερά, η συνθήκη του παραπάνω επιπέδου είναι αληθής, ενώ αντίστοιχα αν ακολουθηθεί το μονοπάτι προς τα δεξιά σημαίνει πως η συνθήκη είναι ψευδής. Πιο συγκεκριμένα, η παρούσα έρευνα ασχολείται με χαρακτηριστικά τα οποία είτε είναι αριθμητικά, είτε έχουν αποδοθεί ως τέτοια. Κατά συνέπεια, και με δεδομένο ότι η συνθήκη που ελέγχεται είναι της μορφής «Χαρακτηριστικό άρθρου $\leq$ Συγκεκριμένη Τιμή», το αριστερό μονοπάτι αφορά τον συγκριτικό τελεστή «μικρότερο ή ίσο» ($\leq$), ενώ το δεξιό τον «μεγαλύτερο» ($>$).

5^ο Στάδιο: Η Επιλογή των Κατάλληλων Κόμβων – Φύλλων

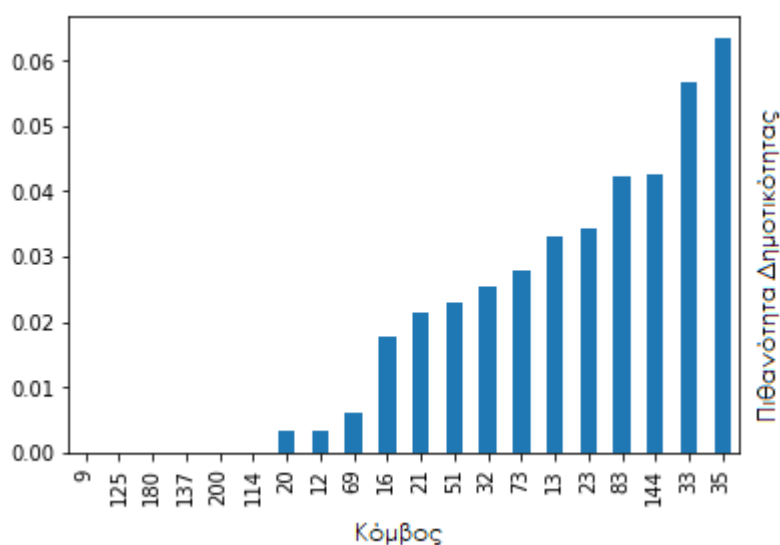
Από τα παραπάνω συνάγεται ότι κάθε κόμβος – φύλλο μπορεί να οδηγήσει στην εξαγωγή ενός συμπεράσματος, το ερώτημα όμως που τίθεται είναι το κατά πόσο το συμπέρασμα αυτό είναι έγκυρο και ισχυρό. Οι κόμβοι – φύλλα έχουν κάποια χαρακτηριστικά, των οποίων οι τιμές ποικίλλουν. Ένα φύλλο του δυαδικού δένδρου έχει συγκεκριμένο πλήθος άρθρων καθώς και συγκεκριμένη πιθανότητα να «δώσει» άρθρο υψηλής δημοτικότητας.



Εικόνα 4: Πλήθος άρθρων ανά κόμβο (Οι 20 μεγαλύτεροι)

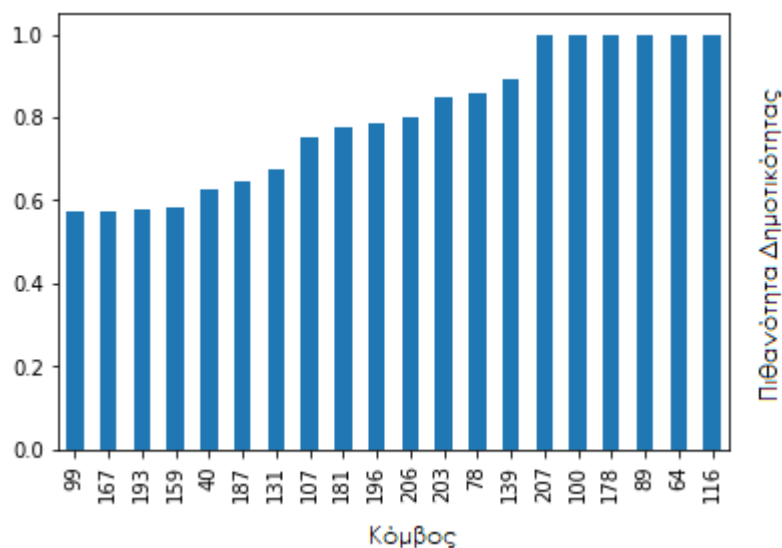
Στην Εικόνα 4, παρουσιάζεται το πλήθος των άρθρων που έχει κάθε κόμβος – φύλλο του δενδρικού σχηματισμού ύστερα από την εφαρμογή του αλγορίθμου του Δέντρου Απόφασης. Το πλήθος αυτό αφορά στο σύνολο δεδομένων εκπαίδευσης για τα άρθρα επιχειρηματικού περιεχομένου και ενδεικτικά αναφέρεται στους είκοσι μεγαλύτερους κόμβους. Είναι εμφανές, λοιπόν, πως τα άρθρα δεν είναι ίσα κατανομημένα σε όλα τα φύλλα, με κάποια από αυτά να είναι πολυπληθέστερα σε σχέση με κάποια άλλα τα οποία δεν έχουν σχεδόν καθόλου άρθρα.

Δεδομένου ότι όλα τα φύλλα δεν έχουν το ίδιο πλήθος άρθρων, θα ήταν λανθασμένη τακτική να ληφθούν υπόψη εκείνα που έχουν μικρό αριθμό άρθρων, αφού το συμπέρασμα που εξάγεται δε μπορεί να γενικευτεί με ασφάλεια. Έτσι, για το παρόν πρόβλημα θα απορριφθούν τα φύλλα που έχουν πλήθος μικρότερο από το 0.8% των δεδομένων που χρησιμοποιούνται για την εκμάθηση.



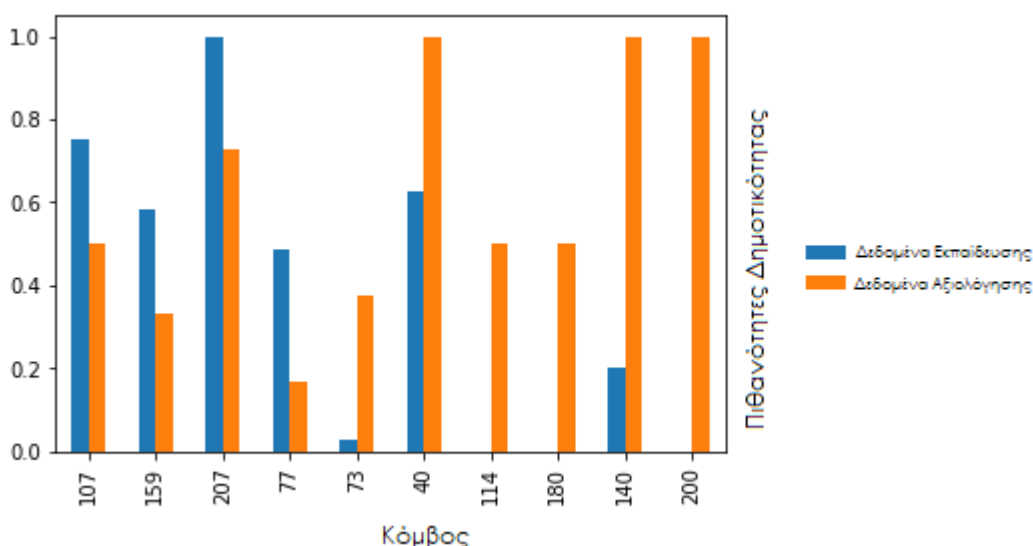
Εικόνα 5: Πιθανότητες Δημοτικότητας (Οι 20 χαμηλότερες)

Όπως φαίνεται στην Εικόνα 5 καθώς και στην Εικόνα 6, η πιθανότητα του κάθε κόμβου – φύλλο (για τα επιχειρηματικά άρθρα, σύμφωνα με το σύνολο δεδομένων εκπαίδευσης) ποικίλλει από 0% έως και 100%. Επειδή, ο στόχος της παρούσας έρευνας αφορά την εύρεση κανόνων για τα δημοφιλή άρθρα, η προσοχή στρέφεται σε φύλλα που βελτιώνουν την αρχική πιθανότητα για υψηλή δημοτικότητα, η οποία είναι 10%. Έτσι, για την εξαγωγή των κανόνων η προσοχή στράφηκε σε φύλλα με πιθανότητα υψηλής δημοτικότητας τουλάχιστον τριπλάσια από την αρχική, δηλαδή μεγαλύτερη ή ίση του 30%.



Εικόνα 6: Πιθανότητες Δημοτικότητας (Οι 20 υψηλότερες)

Η πιθανότητα για την οποία γίνεται λόγος, είναι η πιθανότητα να έχει ένας κόμβος – φύλλο δημοφιλές άρθρο, σύμφωνα με το αποτέλεσμα της εφαρμογής του αλγορίθμου στα δεδομένα εκμάθησης. Ωστόσο, υπάρχει περίπτωση η πιθανότητα του ίδιου κόμβου – φύλλου να διαφέρει όταν ο αλγόριθμος του δέντρου απόφασης εφαρμόζεται για τα δεδομένα αξιολόγησης.



Εικόνα 7: Διαφορές Πιθανοτήτων Δημοτικότητας (Οι 10 μεγαλύτερες)

Στην περίπτωση που οι δύο αυτές πιθανότητες διαφέρουν σε μεγάλο βαθμό (όπως παραδείγματος χάρη φαίνεται στην Εικόνα 7 για τα επιχειρηματικά άρθρα), είναι

προφανές πως γίνεται λόγος για έναν ασταθή κόμβο, ο οποίος δεν θα οδηγήσει σε κάποιο αξιόπιστο συμπέρασμα. Για το λόγο αυτό επιλέχθηκαν, για την εξαγωγή των κανόνων, οι κόμβοι – φύλλα, των οποίων οι πιθανότητες διαφέρουν το πολύ κατά 0.09.

Κατά συνέπεια για την επιλογή των κόμβων – φύλλων που παράγουν τους κανόνες για την αύξηση της δημοτικότητας των δημοσιογραφικών ιστοριών ελέγχθηκε η ταυτόχρονη ικανοποίηση των παρακάτω συνθηκών:

- Πλήθος άρθρων $> 0,8\% \cdot$ Πλήθος δεδομένων εκμάθησης
- Πιθανότητα δημοφιλών άρθρων $> 30\%$
- Πιθανότητα δημοφιλών άρθρων με βάση τα δεδομένα εκμάθησης - πιθανότητα δημοφιλών άρθρων με βάση τα δεδομένα εκμάθησης < 0.09

Αξίζει να σημειωθεί, ότι για κάθε είδος άρθρου, εφαρμόστηκε ξεχωριστά ο αλγόριθμος του δέντρου απόφασης και έτσι παράχθηκε ένα διαφορετικό σύνολο κανόνων για κάθε ένα από αυτά, αφού όπως φαίνεται και στον Πίνακα 1, τα στατιστικά στοιχεία της κατανομής των τιμών των χειροκροτημάτων, των μέσων τιμών αλλά και του πλήθους των άρθρων διαφέρουν από κατηγορία σε κατηγορία. Οι κανόνες που παράχθηκαν δεν είναι ισάξιοι καθώς δε αυξάνουν κατά το ίδιο ποσοστό την πιθανότητα για δημοφιλή άρθρα. Για το λόγο αυτό, για κάθε κανόνα δίνεται και η νέα αυξημένη πιθανότητα να παράξει δημοφιλές άρθρο, αν ακολουθηθεί.

Επιπλέον, για κάθε είδος άρθρου υπολογίστηκε και η τιμή σφάλματος. Η τιμή αυτή είναι ξεχωριστή και αφορά το σύνολο όλων των κανόνων που ανήκουν στο εκάστοτε είδος. Η τιμή σφάλματος υπολογίστηκε ως ο λόγος του αθροίσματος του πλήθους των δημοφιλών άρθρων του κάθε κόμβου – φύλλου επί την διαφορά των πιθανοτήτων του συνόλου δεδομένων εκμάθησης και αξιολόγησης, προς το μέγεθος του συνόλου δεδομένων. Πιο συγκεκριμένα, ο υπολογισμός έγινε σύμφωνα με τον μαθηματικό τύπο:

$$\frac{\sum_{i=1}^k n_i |p_i - p'_i|}{N}, \text{ όπου } n_i: \text{ το πλήθος των δημοφιλών άρθρων στον κόμβο } i$$

k : το πλήθος των κόμβων φύλλων

N : το πλήθος των άρθρων στο σύνολο δεδομένων εκμάθησης

p_i : η πιθανότητα στον κόμβο i για τα δεδομένα εκμάθησης

p'_i : η πιθανότητα στον κόμβο i για τα δεδομένα αξιολόγησης

Πρόκειται στην ουσία για μια μετρική, η οποία χρησιμοποιείται σαν μια επιπλέον δικλείδα ασφαλείας, πέρα από την διαφορά των δύο πιθανοτήτων και αφορά όχι κόμβους αλλά ένα ολόκληρο σύνολο δεδομένων.

7. ΑΠΟΤΕΛΕΣΜΑΤΑ

Στο κεφάλαιο αυτό, παρουσιάζονται τα αποτελέσματα που προέκυψαν από την έρευνα που διεξήχθη. Πιο συγκεκριμένα, για κάθε είδος άρθρου παρουσιάζεται ένα σύνολο κανόνων παραγωγής άρθρων με αυξημένη πιθανότητα υψηλής δημοτικότητας σε σχέση με την αρχική. Επιπλέον, μαζί με τους κανόνες δίνεται και η πιθανότητα του καθενός να «δώσει» δημοφιλές άρθρο, ενώ για κάθε είδους άρθρου παρουσιάζεται και το σφάλμα.

7.1 Κανόνες για ειδησεογραφικά άρθρα

Κανόνας 1 - Πιθανότητα Δημοφιλούς Άρθρου: 88,59 %

Πλήθος Ακολούθων : 3206 – 102679 & Πλήθος Εικόνων > 3.5 &

Προσωπικές αντωνυμίες > 8.587 % & Υποκειμενικότητα > 0.452 &

Αναλογία Ακόλουθων - Ακολουθούμενων : 6.406 - 155.913

Κανόνας 2 - Πιθανότητα Δημοφιλούς Άρθρου: 100,00%

Πλήθος Ακολούθων > 102679 & Πλήθος Λέξεων > 421.5

Σφάλμα: 0.0199

7.2 Κανόνες για τεχνολογικά άρθρα

Κανόνας 1 - Πιθανότητα Δημοφιλούς Άρθρου: 32,99 %

Πλήθος Ακολούθων : 41.5 - 509.5 & Πλήθος Λέξεων > 812.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων : 1.159 - 5.905 &

Προσωπικές αντωνυμίες > 15.053 & Χρήση Τοπωνυμίων = Ψευδής

Κανόνας 2 - Πιθανότητα Δημοφιλούς Άρθρου: 31,90 %

Πλήθος Ακολούθων : 509.5 - 5004.5 & Πλήθος Λέξεων : 767.5 - 1193.5 &

Χρήση Λίστας Σημείων = Αληθής & Αναλογία Ακόλουθων - Ακολουθούμενων <= 11.504 & Προσμονή > 4.202%

Κανόνας 3 - Πιθανότητα Δημοφιλούς Άρθρου: 54,77 %

Πλήθος Ακολούθων > 509.5 & Πλήθος Λέξεων : 767.5 - 2914.5

Χρήση Λίστας Σημείων = Αληθής &

Αναλογία Ακόλουθων - Ακολουθούμενων > 11.504 &

Συναισθηματική Πυκνότητα > 14.414% & Προσωπικές αντωνυμίες > 13.563%

Σφάλμα: 0.0223

7.3 Κανόνες για επιχειρηματικά άρθρα

Κανόνας 1 - Πιθανότητα Δημοφιλούς Άρθρου: 34,34 %

Πλήθος Ακολούθων : 112.5 - 802.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων > 5.164 & Πλήθος Εικόνων > 1.5 &

Χρήση Υπερσυνδέσμων = Ψευδής & Αρνητικότητα > 0.033

Κανόνας 2 - Πιθανότητα Δημοφιλούς Άρθρου: 32,69 %

Πλήθος Ακολούθων : 802.5 - 22747.5 & Πλήθος Λέξεων <= 722.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων : 20.341 - 347.402 &

Αναγνωσιμότητα <= 16.25

Κανόνας 3 - Πιθανότητα Δημοφιλούς Άρθρου: 39,61 %

Πλήθος Ακολούθων : 802.5 - 3790.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων > 2.711 & Πλήθος Λέξεων > 722.5 &

Πλήθος Εικόνων <= 2.5 & Χρήση Τοπωνυμίων = Ψευδής

Κανόνας 4 - Πιθανότητα Δημοφιλούς Άρθρου: 57,65 %

Πλήθος Ακολούθων : 802.5 - 3610.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων : 2.711 - 337.208 &

Πλήθος Λέξεων > 722.5 & Πλήθος Εικόνων > 2.5

Σφάλμα: 0.0278

7.4 Κανόνες για άρθρα υγείας

Κανόνας 1 - Πιθανότητα Δημοφιλούς Άρθρου: 37,73 %

Πλήθος Ακολούθων : 162.5 - 804.0 & Πλήθος Λέξεων > 812.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων > 1.123 & Αρνητικότητα > 0.065

Κανόνας 2 - Πιθανότητα Δημοφιλούς Άρθρου: 41,91 %

Πλήθος Ακολούθων : 1872.5 - 10073.5 & Πλήθος Λέξεων : 351.5 - 1140.0 &

Σπάνιο Λεξιλόγιο <= 3.104% & Προσωπικές αντωνυμίες > 13.052%

Κανόνας 3 - Πιθανότητα Δημοφιλούς Άρθρου: 52,10 %

Πλήθος Ακολούθων : 804.0 - 10073.5 & Πλήθος Λέξεων > 1140.0 &

Αναλογία Ακόλουθων - Ακολουθούμενων : 1.489 - 18.16 &

Αρνητικότητα <= 0.106

Σφάλμα: 0.0226

7.5 Κανόνες για αθλητικά άρθρα

Κανόνας 1 - Πιθανότητα Δημοφιλούς Άρθρου: 34,61 %

Πλήθος Ακολούθων > 1955.5 & Πλήθος Λέξεων : 261.0 - 1305.0 &

Τροπικότητα <= 0.661 & Πλήθος Λέξεων Τίτλου > 5.5

Χρήση Λίστας Σημείων = Ψευδής & Σπάνιο Λεξιλόγιο <= 4.083%

Σφάλμα: 0.0379

7.6 Κανόνες για άρθρα τρόπου ζωής

Κανόνας 1 - Πιθανότητα Δημοφιλούς Άρθρου: 33,85 %

Πλήθος Ακολούθων : 826.5 - 4098.5 & Πλήθος Λέξεων Τίτλου > 5.5 &

Πλήθος Λέξεων > 1476.0 & Θετικότητα > 0.123

Κανόνας 2 - Πιθανότητα Δημοφιλούς Άρθρου: 40,49 %

Πλήθος Ακολούθων : 4098.5 - 7976.5 & Πλήθος Λέξεων > 754.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων > 1.031 - 120.281

Κανόνας 3 - Πιθανότητα Δημοφιλούς Άρθρου: 95,06 %

Πλήθος Ακολούθων : 9047.0 - 38530.5 & Πλήθος Λέξεων > 550.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων <= 230.884 & Έκπληξη <= 4.936% &

Θετικότητα > 0.097

Κανόνας 4 - Πιθανότητα Δημοφιλούς Άρθρου: 99,79 %

Πλήθος Ακολούθων : 9047.0 - 38530.5 & Πλήθος Λέξεων > 550.5 &

Αναλογία Ακόλουθων - Ακολουθούμενων > 230.893 & Λύπη <= 9.0% &

Αναλογία Εικόνων - Λέξεων > 0.001& Σπάνιο Λεξιλόγιο > 0.757%

Σφάλμα: 0.0188

8. ΣΥΖΗΤΗΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Τα αποτελέσματα που προέκυψαν από την έρευνα έδωσαν ένα ικανοποιητικό πλήθος κανόνων (από έναν έως τέσσερις), για κάθε είδος άρθρου. Επίσης, οι κανόνες που παρήχθησαν αύξησαν κατά ένα μεγάλο εύρος τιμών την πιθανότητα απόκτησης υψηλής δημοτικότητας, αφού ανάλογα με την περίπτωση, η νέα πιθανότητα μπορεί να φτάσει από 30% έως και 100% σε σχέση με την αρχική (10%). Ωστόσο, για να δομηθούν οι κανόνες, δεν «αξιοποιήθηκε» όλη η παλέτα των χαρακτηριστικών και αυτό γιατί κάποια από αυτά φαίνεται πως είναι περισσότερο σημαντικά. Πιο συγκεκριμένα, μόνο ένα ήταν το χαρακτηριστικό που ενεπλάκη σε όλους τους κανόνες, το πλήθος των ατόμων που ακολουθούν τον χρήστη. Κάθε κανόνας απευθύνεται σε χρήστες διαφορετικής δημοτικότητας, πράγμα το οποίο είναι ιδιαίτερα σημαντικό, μιας και δεν θα ήταν σωστό να υπάρχει μια καθολική αντιμετώπιση, η οποία δεν λαμβάνει υπόψη την δημοτικότητα του χρήστη. Σε ό,τι αφορά τα υπόλοιπα χαρακτηριστικά, αυτά που εμπλέκονται κυρίως είναι εκείνα που γίνονται αντιληπτά από τη μεριά του χρήστη, αλλά και του περιεχομένου. Στις παρακάτω ενότητες, αναλύονται με μεγαλύτερη λεπτομέρεια τα χαρακτηριστικά που συμμετέχουν στους κανόνες καθώς και η τιμή που πρέπει να λάβουν.

8.1 Ειδησεογραφικά Άρθρα

Στα ειδησεογραφικά άρθρα, η έρευνα κατέληξε σε δύο κανόνες, με μάλιστα ιδιαίτερα μεγάλη πιθανότητα απόκτησης υψηλής δημοτικότητας, για τα άρθρα που συμμορφώνονται με αυτούς.

Ο πρώτος κανόνας αφορά λογαριασμούς, οι οποίοι έχουν από 3206 έως 102679 άτομα που τους ακολουθούν. Αυτό σημαίνει ότι ο κανόνας απευθύνεται σε ένα ευρύ κοινό – στόχος, που ωστόσο αποκλείει τους πολύ αδύναμους από άποψη δημοτικότητας χρήστες. Πρόκειται για έναν κανόνα που αυξάνει από το 10% στο 88,59%, την πιθανότητα να παραχθεί δημοφιλές άρθρο. Για να συμβεί αυτό, θα πρέπει ο συγγραφέας να ακολουθεί αρκετά λιγότερους λογαριασμούς, σε σχέση με αυτούς που τον ακολουθούν και πιο συγκεκριμένα, το πλήθος των ακολούθων του θα πρέπει να είναι από 6 έως 156 φορές μεγαλύτερο, από το πλήθος των ακολουθούμενων από αυτόν. Με άλλα λόγια, θα πρέπει ο λογαριασμός να έχει αποκτήσει στο παρελθόν κύρος και οι ακόλουθοί του να μην προέρχονται λόγω αμοιβαίας ακολούθησης. Σε ό,τι αφορά το στυλ συγγραφής, το άρθρο θα πρέπει να περιέχει προσωπικές αντωνυμίες σε

ποσοστό μεγαλύτερο του 8,5%. Επιπλέον, ο χρήστης θα πρέπει να είναι περισσότερο αντικειμενικός, ωστόσο όχι απόλυτα, ενώ επιθυμητή είναι η έκφραση προσωπικής γνώμης. Πρόκειται για ένα αποτέλεσμα το οποίο είναι αναμενόμενο, μιας και πρόκειται για ειδησεογραφικά άρθρα, όπου η αντικειμενικότητα είναι απαραίτητη για την μετάδοση της είδησης, ωστόσο για να καταφέρει να κερδίζει την ικανοποίηση του αναγνωστικού κοινού, χρειάζεται και η άποψη του συγγραφέα στο συγκεκριμένο θέμα. Τέλος σε ό,τι αφορά το περιεχόμενο του άρθρου, θα πρέπει να περιέχει αρκετό συνοδευτικό υλικό πολυμέσων και ιδιαίτερα εικόνες, θα πρέπει να είναι από τέσσερις και πάνω στο πλήθος. Ο κανόνας αυτός, μπορεί εύκολα να ακολουθηθεί από κάποιον συντάκτη, εφόσον έχει καταφέρει χτίσει την κοινωνική του εικόνα και την αξιοπιστία του κοινού σε παρελθόντα χρόνο, λαμβάνοντας απλώς υπόψη πολύ λίγα πράγματα σχετικά με τον τρόπο συγγραφής και τα δομικά στοιχεία του άρθρου.

Ο δεύτερος κανόνας για τα ειδησεογραφικά άρθρα που εγγυάται 100% επιτυχία, αφορά λογαριασμούς με πάνω από 102679 ακόλουθους και για άρθρα τουλάχιστον 421 λέξεων. Είναι εμφανές ότι πρόκειται για λογαριασμούς εξαιρετικά υψηλής δημοτικότητας, που παράγουν υλικό που πάντοτε συγκεντρώνει την ικανοποίηση του κοινού.

8.2 Τεχνολογικά Άρθρα

Στα τεχνολογικά άρθρα η έρευνα κατέληξε σε τρεις κανόνες που αυξάνουν την πιθανότητα των άρθρων να γίνουν δημοφιλή από τρεις έως 5,5 φορές.

Ο πρώτος κανόνας απευθύνεται σε χρήστες που δεν έχουν καταφέρει να κερδίσουν φήμη στην πλατφόρμα που δημοσιεύουν, μιας και αφορά λογαριασμούς που έχουν αποκτήσει από 42 έως 510 ακόλουθους. Παρά το γεγονός ότι μόλις το 1/3 των άρθρων που παράγει καταφέρνουν να γίνονται δημοφιλή, αυτός ο κανόνας είναι ιδιαίτερα χρήσιμος για νέους χρήστες ή οργανισμούς. Για να παράξουν ένα επιτυχημένο άρθρο, οφείλουν να έχουν ένα σχετικό κύρος, με σίγουρα μεγαλύτερο πλήθος ακολούθων έναντι ακολουθούμενων. Προφανώς, η διαφορά δε θα πρέπει να είναι μεγάλη (1,5 έως 6 φορές περισσότεροι ακόλουθοι). Σε ότι αφορά τα δομικά στοιχεία, το άρθρο θα πρέπει να περιέχει πάνω από 812 λέξεις, ένα ικανοποιητικό δηλαδή πλήθος λέξεων που μπορούν να κερδίσουν τον αναγνώστη και να τον πείσουν για το περιεχόμενων των λόγων του συγγραφέα. Από άποψη συγγραφικού στυλ, ο συγγραφέας θα πρέπει να

έχει ένα σχετικά προσωποποιημένο κείμενο, χρησιμοποιώντας προσωπικές αντωνυμίες σε ποσοστό άνω του 15%, ενώ καλό είναι να εκλείπουν οι γεωγραφικοί προσδιορισμοί.

Ο δεύτερος κανόνας απευθύνεται στην αμέσως επόμενη κλίμακα δημοτικότητας λογαριασμών που γράφουν για το συγκεκριμένο είδος άρθρων. Οι λογαριασμοί που έχουν αποκτήσει από 510 έως 5005 ακόλουθους μπορούν επίσης να αυξήσουν την πιθανότητα παραγωγής δημοφιλών άρθρων στο 32%. Θα πρέπει οι ακόλουθοί τους να μην είναι πάνω από 11,5 φορές περισσότεροι από τους ακολουθούμενους. Επιπλέον, θα πρέπει τα κείμενα που γράφουν να έχουν έκταση από 768 έως 1195 λέξεις. Σύμφωνα με τον δεύτερο κανόνα, τα τεχνολογικά άρθρα θα πρέπει να είναι γραμμένα ή να περιέχουν κάποια μορφή λίστας σημείων. Τέλος, θα πρέπει στο κείμενο να δηλώνεται η προσμονή, με πάνω από 4,2% των λέξεων να εκφράζουν το συγκεκριμένο συναίσθημα. Το τελευταίο χαρακτηριστικό υποδηλώνει ότι τα άρθρα που ασχολούνται με ένα θέμα που προσμένει το κοινό, έχουν μεγαλύτερες πιθανότητες υψηλής δημοτικότητας.

Ο τρίτος κανόνας για τα τεχνολογικά άρθρα αφορά, όσους λογαριασμούς έχουν πάνω από 510 ακόλουθους και αναλογία ακολούθων – ακολουθούμενων πάνω από 11,5. Οι μεμονωμένοι χρήστες ή οι οργανισμοί με τα συγκεκριμένα χαρακτηριστικά, για παράξουν άρθρα με πάνω από 54% πιθανότητα να γίνουν δημοφιλή, θα πρέπει να γράφουν άρθρα 768 έως 2914 λέξεων με ένα σχετικά προσωποποιημένο κείμενο (χρήση προσωπικών αντωνυμιών σε ποσοστό 13% και πάνω) καθώς και με συναισθηματική πυκνότητα πάνω από 14%.

8.3 Επιχειρηματικά Άρθρα

Στα άρθρα επιχειρηματικού χαρακτήρα, κατά την έρευνα εντοπίστηκαν τέσσερις κανόνες που ανεβάζουν την αρχική πιθανότητα επιτυχίας του 10%, σε τιμές από 32,69% έως 57,65%.

Ο πρώτος κανόνας αφορά χρήστες με μικρό πλήθος ακολούθων, από 113 έως 802 που ωστόσο έχουν αρκετό κύρος και ακολουθούν πάνω από πέντε φορές λιγότερους χρήστες σε σχέση με τους ακολούθους τους. Για να παράξει η συγκεκριμένη κατηγορία ένα δημοφιλές άρθρο με πιθανότητα 34,34% θα πρέπει να συμπεριλάβει στο κείμενο τουλάχιστον δύο εικόνες, ενώ δε θα πρέπει να συμπεριλάβει κανένα υπερσύνδεσμο προς κάποια εξωτερική πηγή. Τέλος, θα πρέπει να έχει στο κείμενο ένα μικρό ποσοστό αρνητικά φορτισμένων λέξεων (τουλάχιστον 3,3%). Σίγουρα, μια αρνητική είδηση, ειδικά

στον τομέα της εργασίας, μπορεί να καταστήσει ένα άρθρο δημοφιλές, ωστόσο θα πρέπει να έχει δημοσιευτεί από κάποιον αξιόπιστο χρήστη ή οργανισμό.

Ο δεύτερος κανόνας απασχολεί χρήστες που έχουν λίγο περισσότερους ακόλουθους (803 - 22747), ωστόσο επαυξημένο κύρος με αναλογία ακολούθων - ακολουθούμενων από 20 έως 347. Σε αυτή την περίπτωση, δεν χρειάζεται το άρθρο να είναι εκτενές για να γίνει δημοφιλές, καθώς 722 λέξεις είναι «υπεραρκετές». Επίσης, παρά το γεγονός ότι το θέμα ενδεχομένως απαιτεί ειδικές γνώσεις, δεν θα πρέπει το άρθρο να είναι ιδιαίτερα δυσανάγνωστο, ή τουλάχιστον να είναι ευκολότερα αναγνώσιμο από μια επιστημονική έρευνα. Σύμφωνα με την κλίμακα Flesch – Kincaid η ανώτερη αποδεκτή τιμή για την επιτυχία του άρθρου μπορεί να είναι 16.25.

Ο τρίτος κανόνας εξειδικεύει το κοινό του προηγούμενου κανόνα σε αυτούς που έχουν από 803 έως 3790 ακόλουθους. Για να ανέλθει η πιθανότητα παραγωγής δημοφιλούς άρθρου στο 39,61%, θα πρέπει τουλάχιστον ο χρήστης να έχει τουλάχιστον 2,7 φορές περισσότερους ακόλουθους σε σχέση με τα άτομα που ακολουθεί. Επιπλέον θα πρέπει να έχει έκταση μεγαλύτερη από 722 λέξεις, το πολύ δύο εικόνες και να μην περιέχει γεωγραφικό προσανατολισμό.

Ο τέταρτος κανόνας είναι σχεδόν παρόμοιος με τον προηγούμενο και αφορά χρήστες με 803 έως 3610 ακόλουθους και αναλογία ακολούθων – ακολουθούμενων από 2,7 έως 337. Για να παραχθεί δημοφιλές άρθρο, θα πρέπει σε αυτή την περίπτωση το πλήθος των λέξεων να είναι πάνω από 722 και οι εικόνες που το συνοδεύουν τουλάχιστον τρεις.

8.4 Άρθρα Υγείας

Στα άρθρα υγείας εντοπίστηκαν κατά την έρευνα τρεις κανόνες, οι οποίοι αυξάνουν την πιθανότητα υψηλής δημοτικότητας σε τιμές από 37,73% έως 52,10%.

Ο πρώτος κανόνας αφορά άρθρα χρηστών, των οποίων οι ακόλουθοι κυμαίνονται από 163 έως 804, ενώ είναι τουλάχιστον πάνω από 1,1 φορά μεγαλύτεροι από το πλήθος των λογαριασμών που ακολουθεί ο χρήστης που το δημοσιεύει. Οι χρήστες αυτοί μπορούν να αυξήσουν την πιθανότητα απόκτησης υψηλής δημοτικότητας των άρθρων τους απλά χρησιμοποιώντας αρνητικά φορτισμένες λέξεις σε ποσοστό μεγαλύτερο του 6,5% .

Ο δεύτερος κανόνας απευθύνεται σε χρήστες με 1873 έως 10073 ακόλουθους. Για να παράξουν οι χρήστες άρθρα, που έχουν πιθανότητα της τάξης του 41.9% να

αποκτήσουν υψηλή δημοτικότητα , θα πρέπει να γράψουν ένα άρθρο έκτασης 352 έως 1140 λέξεων, στο οποίο να μην χρησιμοποιούν σπάνιο λεξιλόγιο σε μεγάλο βαθμό (το πολύ 3,1% των συνολικών λέξεων να είναι σπάνιες). Επιπλέον, θα πρέπει να χρησιμοποιούν αρκετές προσωπικές αντωνυμίες σε ποσοστό πάνω από 13%.

Ο τρίτος κανόνας έχει να κάνει με ένα μεγαλύτερο εύρος δημοτικότητας των χρηστών, αφού απευθύνεται σε εκείνους που έχουν από 804 έως 10073 ακόλουθους. Οι συγκεκριμένοι χρήστες για να αυξήσουν την πιθανότητα του άρθρου τους να γίνει δημοφιλές, στο 52,1% θα πρέπει να καταβάλλουν χρόνο γράφοντας ένα εκτενές κείμενο άνω των 1140 λέξεων, στο οποίο η αρνητικά φορτισμένες λέξεις να μην ξεπερνούν το 10,6% των συνολικών λέξεων. Επιπλέον, θα πρέπει να έχουν 1,5 έως 18 φορές περισσότερους χρήστες που τους ακολουθούν, σε σχέση με τους χρήστες που ακολουθούν οι ίδιοι.

8.5 Αθλητικά Άρθρα

Ο μοναδικός κανόνας για τα άρθρα αθλητικού περιεχομένου, ο οποίος παράγει δημοφιλή άρθρα με πιθανότητα 34,61% απευθύνεται σε χρήστες με πάνω από 1955 ακόλουθους. Σε αυτή την περίπτωση ο συγγραφέας θα πρέπει να συντάξει ένα κείμενο των 261 έως 1305 λέξεων, στο οποίο θα πρέπει να εκφράζει σιγουριά στα όσα γράφει, αλλά όχι απολυτότητα. Επιπλέον, θα πρέπει η χρήση του σπάνιου λεξιλογίου να είναι περιορισμένη και πιο συγκεκριμένα οι όχι συχνά χρησιμοποιούμενες λέξεις μην αποτελούν παραπάνω από το 4% των συνολικών λέξεων. Τέλος, για την επιτυχία του άρθρου είναι επιθυμητή η απουσία μορφής λίστας σημείων από το κείμενο, ενώ ο τίτλος θα πρέπει να μην είναι σύντομος καθώς απαιτείται να αποτελείται από τουλάχιστον έξι λέξεις.

8.6 Άρθρα Τρόπου Ζωής

Για τα lifestyle άρθρα, εντοπίστηκαν τέσσερις κανόνες που αυξάνουν την πιθανότητα απόκτησης υψηλής δημοτικότητας σε ένα εύρος τιμών από 33,85% έως και 99,79%.

Ο πρώτος κανόνας απευθύνεται σε λογαριασμούς με μέτριο πλήθος ακολούθων και πιο συγκεκριμένα σε χρήστες που ακολουθούνται από 827 έως 4098 άτομα. Οι χρήστες αυτοί, αν λάβουν υπόψη τους τρία συγκεκριμένα χαρακτηριστικά, που αφορούν το περιεχόμενο των άρθρων, τότε η μία στις τρεις δημοσιογραφικές ιστορίες που θα

δημοσιεύουν θα γίνεται δημοφιλής, συγκεντρώνοντας υψηλό πλήθος θετικών αντιδράσεων. Έτσι, θα πρέπει να γράψουν μια εκτενή ιστορία άνω των 1476 λέξεων, εκ των οποίων πάνω από το 12,3% θα πρέπει να είναι θετικά φορτισμένες. Τέλος, θα πρέπει να δώσουν στην ιστορία τους έναν τίτλο τουλάχιστον έξι λέξεων, για να καταφέρουν να κερδίσουν την ικανοποίηση των αναγνωστών τους.

Ο δεύτερος κανόνας, απευθύνεται σε λογαριασμούς που έχουν λίγο υψηλότερη δημοτικότητα και συγκεκριμένα από 4099 – 7976 ακόλουθους. Σε αυτή την περίπτωση είναι αρκετά εύκολο να παραχθεί μια δημοφιλής ιστορία με πιθανότητα 40,49%, μιας και η αναλογία ακολούθων και ακολουθούμενων αρκεί να είναι μεγαλύτερη του 1 και μικρότερη του 120. Με άλλα λόγια, μπορεί ο χρήστης απλά να ακολουθείται από λίγα ή πολύ περισσότερα άτομα παραπάνω από όσα ακολουθεί. Επίσης αρκεί να γράψει μια ιστορία άνω των 755 λέξεων για να τετραπλασιάσει την αρχική πιθανότητα απόκτησης υψηλής δημοτικότητας.

Ο τρίτος κανόνας αφορά τους λογαριασμούς που έχουν αποκτήσει πολύ μεγάλο αριθμό ακολούθων και συγκεκριμένα πάνω από 9048. Σε αυτή την περίπτωση αν ένας συγγραφέας συμμορφωθεί με αυτόν θα έχει 95,06% πιθανότητα να δημοσιεύσει ένα άρθρο που θα γίνει δημοφιλές. Δεν υπάρχει κάποιος ουσιαστικός περιορισμός στην αναλογία ακόλουθων και ακολουθούμενων, παρά μόνο ότι αυτή θα πρέπει να είναι μικρότερη του 230. Σε ό,τι αφορά το περιεχόμενο του άρθρου, θα πρέπει η έκτασή του να ξεπερνά τις 550 λέξεις, εκ των οποίων πάνω από το 9,7% θα πρέπει να εκφράζουν θετικά συναισθήματα, ενώ είναι επιθυμητό το πλήθος αυτών που εκφράζουν το συναίσθημα της έκπληξης να μην ξεπερνάνε το 4,93% του συνολικού πλήθους λέξεων.

Ο τέταρτος και τελευταίος κανόνας για τα άρθρα τρόπου ζωής αφορά επίσης τους χρήστες ίδιας δυναμικής από άποψη ακολούθων με τον προηγούμενο κανόνα. Σε αυτή την περίπτωση όμως, γίνεται αναφορά σε λογαριασμούς πολύ υψηλού κύρους, μιας και οι ακόλουθοι πρέπει να είναι πάνω από 230 φορές περισσότεροι από τους ακολουθούμενους, ώστε η ιστορία να έχει 99,79% για να γίνει δημοφιλής. Επιπλέον, θα πρέπει να έχει έκταση πάνω από 550 λέξεις και τουλάχιστον 0,001 εικόνα ανά λέξη του κειμένου. Αξίζει επίσης να σημειωθεί, ότι θα πρέπει να χρησιμοποιείται μη κοινότυπο λεξιλόγιο σε ποσοστό άνω του 0,75% και το πλήθος των λέξεων που υποδηλώνουν το συναίσθημα της λύπης να μην υπερβαίνουν το 9% των συνολικών λέξεων του άρθρου.

9. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

Φτάνοντας στο τέλος αυτής της εργασίας, έχουν παραχθεί συνολικά 17 κανόνες συγγραφής άρθρων που αφορούν έξι διαφορετικές κατηγορίες και βελτιώνουν την πιθανότητα απόκτησης υψηλής δημοτικότητας από το αρχικό 10% σε ένα ποσοστό μεταξύ του 30% και του 100%. Πρόκειται για ένα αρκετά μεγάλο πλήθος κανόνων, οι οποίοι μπορούν να διευκολύνουν τους χρήστες να γράψουν άρθρα, ικανά να αποκτήσουν μεγάλη δημοτικότητα. Από τα 32 χαρακτηριστικά που σχημάτισαν το από κάτω προς τα πάνω μοντέλο, 21 ήταν εκείνα τα οποία συμμετείχαν στους κανόνες που μπορούν να επηρεάσουν τη δημοτικότητα των άρθρων. Πιο συγκεκριμένα, χαρακτηριστικά που έχουν να κάνουν με τη δημοτικότητα του χρήστη, το στυλ συγγραφής αλλά και τη δομή του άρθρου ήταν εκείνα που συνέβαλαν στην απόκτηση υψηλής δημοτικότητας. Ωστόσο, το χαρακτηριστικό εκείνο το οποίο ενεπλάκη σε όλους τους κανόνες είναι το πλήθος των ακολούθων του χρήστη, το οποίο λόγω της υψηλής σημαντικότητάς του, αποτελεί και διακριτικό γνώρισμα των κανόνων. Μάλιστα, ανάλογα με το πόσα άτομα έχει καταφέρει να συγκεντρώσει ένας χρήστης ως ακόλουθους, θα πρέπει να συμβουλευτεί διαφορετικούς κανόνες για τη συγγραφή των άρθρων του.

Παρατηρώντας τους κανόνες μπορεί να εξάγει κανείς το συμπέρασμα, ότι στις περισσότερες περιπτώσεις είναι αναγκαίο ο λογαριασμός του χρήστη ή του οργανισμού που δημοσιεύει τις δημοσιογραφικές ιστορίες, να έχει ξεκάθαρα περισσότερους ακόλουθους σε σχέση με τους λογαριασμούς που ο ίδιος ακολουθεί. Σε ό,τι αφορά τα υπόλοιπα χαρακτηριστικά, όμως, δε μπορεί να εξαχθεί ένα καθολικό συμπέρασμα μιας και οι κανόνες διαφέρουν μεταξύ τους σημαντικά, ειδικά όταν πρόκειται για διαφορετικά είδη άρθρων.

Η χρήση των δυαδικών δένδρων απόφασης, για την παραγωγή των κανόνων, είναι μια ιδιαίτερα σημαντική επιλογή, ωστόσο είχε ορισμένους περιορισμούς, οι οποίοι οφείλονται στη φύση αυτού του αλγορίθμου. Για κάθε χαρακτηριστικό που συμμετέχει στον κανόνα, θα ήταν ιδανικό να δίνεται ένα διάστημα τιμών με διασαφηνισμένες, τόσο την κατώτερη όσο και την ανώτερη τιμή που δύναται να λάβει. Ωστόσο, το δένδρο απόφασης επειδή έχει ως στόχο να εντοπίσει τη βέλτιστη λύση σε ένα πρόβλημα ταξινόμησης και όχι την παραγωγή κανόνων, δεν λαμβάνει υπόψη αυτή την ανάγκη για διάστημα τιμών. Μπορεί σε ορισμένες περιπτώσεις να αποδίδονται το ανώτερο και το κατώτερο όριο τιμών, εντούτοις αξίζει να σημειωθεί ότι αυτό δε γίνεται εσκεμμένα ή καθολικά, αλλά μόνο για λόγους εύρεσης της καλύτερης λύσης στο πρόβλημα.

Όπως διευκρινίστηκε και παραπάνω, αναπτύχθηκε ένα μοντέλο που επιλέγει μεγάλους, σταθερούς και με πιθανότητα για καλό άρθρο άνω του 30%, κόμβους – φύλλα. Ωστόσο, στον δένδρικό σχηματισμό υπάρχουν και κόμβοι οι οποίοι δεν πληρούν την προϋπόθεση να περιλαμβάνουν μεγάλο πλήθος άρθρων, παρ' όλ' αυτά είναι ικανοί να παράγουν δημοφιλή άρθρα με πιθανότητα κοντά στο 100%. Σε μία μελλοντική έρευνα, η οποία θα αποτελεί εξέλιξη της παρούσας, αξίζει να μελετηθούν και αυτοί οι κόμβοι – φύλλα, ώστε να γίνει εμφανές αν μπορούν ή όχι να δώσουν κάποια χρήσιμη πληροφορία για τα χαρακτηριστικά που μπορεί να καταστήσουν τα διαδικτυακά άρθρα δημοφιλή στην πλατφόρμα που δημοσιεύονται.

Τέλος, θα ήταν χρήσιμο να γίνει στο μέλλον και η αντίστοιχη διαδικασία για τα δημοσιογραφικά άρθρα χαμηλής δημοτικότητας, ώστε να αποκαλυφθούν ποια είναι εκείνα τα χαρακτηριστικά, τα οποία όταν λαμβάνουν συνδυαστικά συγκεκριμένες τιμές μπορούν δώσουν με σχεδόν απόλυτη σιγουριά, άρθρο που δεν μπορεί να γίνει δημοφιλές. Σε αυτή την περίπτωση, τα αποτελέσματα θα λειτουργήσουν σαν «παράδειγμα προς αποφυγή», ωστόσο αυτό είναι κάτι το οποίο θα μπορούσε να χρησιμοποιηθεί μόνο συμπληρωματικά μαζί με τους κανόνες που παρήχθησαν από τη συγκεκριμένη έρευνα.

ΠΑΡΑΡΤΗΜΑ Ι

Στο Παράρτημα Ι, παρουσιάζονται οι κανόνες, όπως παράχθηκαν από τον αλγόριθμο του δέντρου απόφασης, στη γλώσσα προγραμματισμού Python. Σε περίπτωση που υπάρχει παραπάνω από μια αναφορά για κάποιο χαρακτηριστικό σε έναν κανόνα, διατηρείται η τελευταία, θεωρώντας ότι το δέντρο απόφασης, αποκτά μεγαλύτερη γνώση καθώς προχωράει σε μεγαλύτερο βάθος.

Επιχειρηματικά Άρθρα:

```
[['XFollowed By <= 802.5',
  'XF/F > 5.164',
  'XFollowed By > 104.5',
  'XNo. Images > 1.5',
  'XFollowed By > 112.5',
  'XLinks present <= 0.5',
  'XNegativity > 0.033'],
 ['XFollowed By > 802.5',
  'XF/F > 2.711',
  'XNo. Words <= 722.5',
  'XFollowed By <= 22747.5',
  'XF/F > 20.341',
  'XF/F <= 347.402',
  'XFlesh <= 16.25'],
 ['XFollowed By > 802.5',
  'XF/F > 2.711',
  'XNo. Words > 722.5',
  'XFollowed By <= 17458.5',
  'XNo. Images <= 2.5',
  'XContains USA <= 0.5',
  'XFollowed By <= 3790.5'],
 ['XFollowed By > 802.5',
  'XF/F > 2.711',
  'XNo. Words > 722.5',
  'XFollowed By <= 17458.5',
  'XNo. Images > 2.5',
  'XFollowed By <= 3610.5',
  'XF/F <= 337.208']]
```

Άρθρα Υγείας:

```
[['XFollowed By <= 804.0',
  'XFollowed By > 7.5',
  'XNo. Words > 812.5',
  'XFollowed By > 162.5',
  'XF/F > 0.44',
  'XF/F > 1.123',
  'XNegativity > 0.065'],
 ['XFollowed By > 804.0',
  'XFollowed By <= 10073.5',
  'XNo. Words <= 1140.0',
```

'XNo. Words > 351.5',
'XFollowed By > 1872.5',
'XBeautiful phrasing <= 3.104',
'XTone of speech > 13.052'],
['XFollowed By > 804.0',
'XFollowed By <= 10073.5',
'XNo. Words > 1140.0',
'XF/F <= 24.955',
'XF/F > 1.489',
'XF/F <= 18.16',
'XNegativity <= 0.106']]

Άρθρα Τρόπου Ζωής:

['XFollowed By <= 9047.0',
'XFollowed By > 826.5',
'XNo. Words > 754.5',
'XFollowed By <= 4098.5',
'XNo. TitleWords > 5.5',
'XNo. Words > 1476.0',
'XPositivity > 0.123'],
['XFollowed By <= 9047.0',
'XFollowed By > 826.5',
'XNo. Words > 754.5',
'XFollowed By > 4098.5',
'XF/F > 1.031',
'XF/F <= 120.281',
'XFollowed By <= 7976.5'],
['XFollowed By > 9047.0',
'XFollowed By > 38530.5',
'XNo. Words > 550.5',
'XF/F <= 230.893',
'XF/F <= 230.884',
'XSurprise <= 4.936',
'XPositivity > 0.097'],
['XFollowed By > 9047.0',
'XFollowed By > 38530.5',
'XNo. Words > 550.5',
'XF/F > 230.893',
'XSadness <= 9.0',
'Ximg/word > 0.001',
'XBeautiful phrasing > 0.757']]

Ειδησεογραφικά Άρθρα:

['XFollowed By > 3206.0',
'XFollowed By <= 102679.0',
'XNo. Images > 3.5',
'XTone of speech > 8.587',
'XSubjectivity > 0.452',

'XF/F > 6.406',
'XF/F <= 155.913'],
['XFollowed By > 3206.0', 'XFollowed By > 102679.0', 'XNo. Words > 421.5']]

Αθλητικά Άρθρα:

['XFollowed By > 1955.5',
'XNo. Words <= 1305.0',
'XModality <= 0.661',
'XNo. TitleWords > 5.5',
'XBullets present <= 0.5',
'XNo. Words > 261.0',
'XBeautiful phrasing <= 4.083']]

Τεχνολογικά Άρθρα:

['XFollowed By <= 509.5',
'XFollowed By > 41.5',
'XNo. Words > 812.5',
'XF/F > 1.159',
'XTone of speech > 15.053',
'XContains USA <= 0.5',
'XF/F <= 5.905'],
['XFollowed By > 509.5',
'XNo. Words > 767.5',
'XBullets present > 0.5',
'XF/F <= 11.504',
'XAnticipation > 4.202',
'XNo. Words > 1193.5',
'XFollowed By <= 5004.5'],
['XFollowed By > 509.5',
'XNo. Words > 767.5',
'XBullets present > 0.5',
'XF/F > 11.504',
'XNo. Words <= 2914.5',
'XDensity > 14.414',
'XTone of speech > 13.563']]

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

- Arapakis, I., Peleja, F., Cambazoglu, B. B., & Magalhães, J. (2016). Linguistic Benchmarks of Online News Article Quality. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers* (Vol. 4, pp. 1893-1902)
- Bandari, R., Asur, S., & Huberman, B. A. (2012, May). The pulse of news in social media: Forecasting popularity. In *Sixth International AAAI Conference on Weblogs and Social Media*.
- Boler, M. (Ed.). (2010). *Digital media and democracy: Tactics in hard times*. Mit Press.
- Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P. (2010, May). Measuring user influence in twitter: The million follower fallacy. In *fourth international AAAI conference on weblogs and social media*.
- Cheng, Y., Wang, F., Zhang, P., & Hu, J. (2016, June). Risk prediction with electronic health records: A deep learning approach. In *Proceedings of the 2016 SIAM International Conference on Data Mining* (pp. 432-440). Society for Industrial and Applied Mathematics.
- De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *International Journal of Advertising*, 36(5), 798-828.
- Dictionary of Standard Modern Greek. (χ.χ.). Ανακτήθηκε 13 Ιανουαρίου, 2020, από http://www.greek-language.gr/greekLang/modern_greek/tools/lexica/triantafyllides/search.html?lq=%CE%B1%CF%80%CE%AE%CF%87%CE%B7%CF%83%CE%B7&dq=
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221.
- Hu, Y., Farnham, S., & Talamadupula, K. (2015, April). Predicting user engagement on twitter with real-world events. In *Ninth International AAAI Conference on Web and Social Media*.
- Hutto, C. J., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.

Iwamoto, N. (2007). Modality and point of view in media discourse. *Retrieved December, 17, 2013.*

Jin, Y., Huang, J., & Wang, X. (2017). What influences content popularity? An empirical investigation of voting in social Q&A communities. In *Pacific Asia Conference on Information Systems (PACIS)*. Association For Information Systems.

Keneshloo, Y., Wang, S., Han, E. H., & Ramakrishnan, N. (2016, June). Predicting the popularity of news articles. In *Proceedings of the 2016 SIAM International Conference on Data Mining* (pp. 441-449). Society for Industrial and Applied Mathematics.

Louis, A., & Nenkova, A. (2013). What makes writing great? First experiments on article quality prediction in the science journalism domain. *Transactions of the Association for Computational Linguistics*, 1, 341-352.

Marres, N., & Weltevrede, E. (2013). Scraping the social? Issues in live social research. *Journal of cultural economy*, 6(3), 313-335.

McCready, R., (2016, May 16). What makes an article popular on Medium? We analyzed 10,000+ data points to find out. Medium. Ανακτήθηκε από <https://medium.com/the-mission/after-10-000-data-points-we-figured-out-how-to-write-a-perfect-medium-post-58c41c314f6a> 19/12/2019

Medium (website). (χ.χ.). Ανακτήθηκε 13 Ιανουαρίου, 2020, από το Wikipedia: [https://en.wikipedia.org/wiki/Medium_\(website\)](https://en.wikipedia.org/wiki/Medium_(website))

Mitchell, T. M. (1997). Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45(37), 870-877.

Mohammad, S. M., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. *arXiv preprint arXiv:1308.6242*.

Okrent, A. (2014). The listicle as literary form. *University of Chicago Magazine*, 106(3), 52-53.

Orellana-Rodriguez, C., & Keane, M. T. (2018). Attention to news and its dissemination on Twitter: A survey. *Computer Science Review*, 29, 74-94.

Paul, S. A., Hong, L., & Chi, E. H. (2012). Who is authoritative? understanding reputation mechanisms in quora. *arXiv preprint arXiv:1204.3724*.

Piqué, J., Posteguillo, S., & Andreu-Besó, J. V. (2001). A pragmatic analysis framework for the description of modality usage in academic English contexts. *ELIA*, 2, 213-224.

Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2011, September). Influence and passivity in social media. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 18-33). Springer, Berlin, Heidelberg.

Smedt, T. D., & Daelemans, W. (2012). Pattern for python. *Journal of Machine Learning Research*, 13(Jun), 2063-2067.

Sotirakou, C., Trilling, D., Germanakos, P., & Mourlas, C. (2019, October). Opening the black box of perceived quality: Predicting endorsement on a blog site. In *2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI)* (pp. 388-392). IEEE.

TextBlob Sentiment: Calculating Polarity and Subjectivity (2015, Ιούνιος 7). Ανακτήθηκε 8 Δεκεμβρίου, 2019, από https://planspace.org/20150607-textblob_sentiment/

Varshney, V., & Wagh, R. S. (2017). Weighted Sentiment Score Formulation Using Sentence Level Sentiment Density for Opinion Analysis. *International Journal of Computational Intelligence Research*, 13(2), 285-298.

Witten, I. H., Frank, E., & Hall, M. A. (2005). Practical machine learning tools and techniques. Morgan Kaufmann, 578.

Yoganarasimhan, H. (2012). Impact of social network structure on content propagation: A study using YouTube data. *Quantitative Marketing and Economics*, 10(1), 111-150.

Γεωργούλη, Α. (2015). Μηχανική Μάθηση.

Καϊμάκη, Β., Καραμπάσης, Ζ., Παναγιωτάκης, Μ., Τσενέ, Λ. και Τσιμιτάκης, Μ. (2016). Μέσα Κοινωνικής Δικτύωσης και Επικοινωνία. χ.τ Εκδόσεις: Μεταμεσονύκτιες Εκδόσεις. Ανακτήθηκε 27 Ιανουαρίου, 2019, από <https://www.openbook.gr/mesa-koinwnikis-diktywsis-kai-epikoinwnia/>

Κόβατς, Μ. & Ρόζεντιλ, Τ. (2004). Εισαγωγή στη δημοσιογραφία Τι πρέπει να γνωρίζουν οι άνθρωποι των μέσων και τι πρέπει να αναμένει η κοινή γνώμη. (Ε. Μπαρτζινόπουλος, μετ.). Αθήνα: Εκδόσεις Καστανιώτη.

Πολίτης, Π. (Επιμ.). (2008). Ο λόγος της μαζικής επικοινωνίας Το ελληνικό παράδειγμα. Θεσσαλονίκη: Ίδρυμα Τριανταφυλλίδη.

Ποια είναι η διαφορά ανάμεσα στις προβολές Σελίδας, την απήχηση και τις εμφανίσεις στο Facebook. (χ.χ.). Ανακτήθηκε 13 Ιανουαρίου, 2020, από https://www.facebook.com/help/274400362581037?helpref=uf_permalink

Σε τι διαφέρουν μεταξύ τους η οργανική απήχηση, η πληρωμένη απήχηση και η απήχηση δημοσιεύσεων. (χ.χ.). Ανακτήθηκε 13 Ιανουαρίου, 2020, από

https://www.facebook.com/help/285625061456389?helpref=uf_permalink

Χατζησαββίδης, Σ. (2000). Ελληνική γλώσσα και δημοσιογραφικός λόγος. Αθήνα: Εκδόσεις Gutenberg.

Χατζησαββίδης, Σ., & Χατζησαββίδου, Α. (2016). Γραμματική Νέας Ελληνικής Γλώσσας. Α΄, Β΄, Γ΄ Γυμνασίου. Α΄, Β΄, Γ΄ Γυμνασίου, Διοφάντους.