



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

— ΙΔΡΥΘΕΝ ΤΟ 1837 —

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΚΑΙ ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΕΠΙΚΟΙΝΩΝΙΑΣ ΚΑΙ ΜΕΣΩΝ ΜΑΖΙΚΗΣ ΕΝΗΜΕΡΩΣΗΣ

ΠΜΣ «ΕΠΙΚΟΙΝΩΝΙΑ ΚΑΙ ΜΜΕ: ΠΟΛΙΤΙΚΗ ΕΠΙΚΟΙΝΩΝΙΑ ΚΑΙ ΝΕΑ ΜΕΣΑ»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αξιοποίηση δεδομένων μεγάλου όγκου, ανάλυση συναισθήματος και
εξόρυξη απόψεων στην πολιτική επικοινωνία:
Αυτοματοποιημένη ανάλυση τεχνικών προπαγάνδας



Μαλτέζος Βασίλειος - Α.Μ. 9983201737020

Επιβλέπων Καθηγητής: Κωνσταντίνος Μουρλάς

Συντονιστής: Νίκος Παπαναστασίου

ΑΘΗΝΑ, ΙΑΝΟΥΑΡΙΟΣ 2018

Πίνακας περιεχομένων

1.	Εισαγωγή.....	2
2.	Τα στάδια μέχρι την υπολογιστικοποίηση της ανάλυσης	4
3.	Η βασική δομή της επικοινωνίας.....	5
4.	Το θεωρητικό πλαίσιο της ανάλυσης	9
4.1.	Ορισμός	9
4.2.	Το μοντέλο προπαγάνδας Herman & Chomsky.....	11
4.3.	Τα χαρακτηριστικά της προπαγάνδας εν γένει.....	12
4.3.1.	Αποσαφήνιση και διαχωρισμός του όρου βάσει χαρακτηριστικών.....	12
4.3.2.	Παραδοχές – Παρατηρήσεις για την προπαγάνδα	16
4.4.	Τα προς ανάλυση τεχνάσματα	19
5.	Παρόμοια εγχειρήματα.....	29
6.	Ποιοτική ανάλυση.....	34
6.1.	Αρχικές διευκρινίσεις για την ανάλυση των τεχνικών	34
6.2.	Δεδομένα.....	35
6.3.	Ταξινόμηση	37
6.4.	Παρατηρήσεις και συμπεράσματα από την ποιοτική ανάλυση	39
7.	Έρευνα.....	44
7.1.	Ερευνητικά ερωτήματα	44
7.2.	Ερευνητικές υποθέσεις	44
7.3.	Κριτήρια αποτελεσματικότητας.....	45

8.	Χαρακτηριστικά της Μεθόδου.....	47
8.1.	Πλεονεκτήματα.....	47
8.2.	Περιορισμοί και προκλήσεις.....	48
9.	Διαδικασία αυτοματοποιημένης ανάλυσης.....	51
10.	Αποτελέσματα.....	60
10.1.	Αρχικές δοκιμές.....	62
10.2.	Βελτίωση και επαναξιολόγηση.....	67
10.2.1.	Δοκιμές και αποτελέσματα.....	70
11.	Συμπεράσματα.....	89
12.	Βελτιώσεις – Οδηγίες για μελλοντικά εγχειρήματα.....	91
13.	Χρησιμότητα του λογισμικού.....	92
14.	Αναφορές.....	94
15.	Παράρτημα.....	99

*Στη μνήμη του Alan Matheson Turing,
στους αγαπημένους μου καθηγητές Δρ.
Κωνσταντίνο Μουρλά και Δρ. Νίκο
Παπαναστασίου καθώς και στην υπ. Δρ.
Κατερίνα Σωτηράκου για την
πολύτιμη βοήθειά τους*

1. Εισαγωγή

Ήδη από το 1936 ο *Alan Matheson Turing* έθεσε τα θεμέλια για την ανάπτυξη τεχνητής νοημοσύνης με τη θεωρία του υπολογισμού. Ο Άγγλος μαθηματικός και, όπως αποκαλείται, «πατέρας της επιστήμης των υπολογιστών» απέδειξε πως μία υπολογιστική ακολουθία «γ» μπορεί να πραγματοποιηθεί όχι μόνο από τη λογική επεξεργασία από ένα ανθρώπινο μυαλό αλλά και από μία υπολογιστική μηχανή στην οποία εισάγονται τα δεδομένα σε μορφή που να μπορεί να τα επεξεργαστεί καθώς και ο αλγόριθμος επεξεργασίας των δεδομένων ή αλλιώς ο τρόπος που προγραμματίζεται η μηχανή να τα επεξεργαστεί ώστε να δώσει το αποτέλεσμα για το ζητούμενο ερώτημα και έτσι να εκτελέσει ορθά και με ακρίβεια την υπολογιστική ακολουθία (Turing, 1936). Έκτοτε και διαδοχικά, με τη δημιουργία των πρώτων υπολογιστικών μηχανών, του πρώτου υπολογιστή, των προσωπικών ηλεκτρονικών υπολογιστών, του Παγκόσμιου Ιστού (Διαδικτύου) και την διαρκώς βελτιστοποιούμενη τεχνολογία δικτύων καθώς και αποθήκευσης και επεξεργασίας δεδομένων οι δυνατότητες για αυτόματους υπολογισμούς ξεπέρασαν ίσως και τις προσδοκίες του ίδιου του Turing και έφεραν εις πέρας επιτυχώς πλέον τη δοκιμασία που ο ίδιος ο Turing έθεσε για την επίτευξη μίας τεχνητής νοημοσύνης με νοητική ικανότητα όμοια με του ανθρώπου (Turing, 1950).

Η χρήση της τεχνητής νοημοσύνης για την επεξεργασία γραπτού λόγου έχει ήδη χρησιμοποιηθεί για την εξαγωγή ποικίλων συμπερασμάτων σχετικά με το περιεχόμενο και τον σκοπό του λόγου. Οι αλγόριθμοι Υπολογιστικής Γλωσσολογίας (Computational Linguistics) και μέθοδοι Μηχανικής Μάθησης είναι τα εργαλεία με τα οποία καθίσταται δυνατή η αυτόματη ανάλυση κειμένων η οποία αποτελεί μία προσπάθεια εξαγωγής τέτοιων συμπερασμάτων με τη μέγιστη δυνατή ακρίβεια με τη χρήση ειδικού λογισμικού (Conroy, et al., 2015). Μία προσπάθεια που έχει ήδη γίνει από άλλους, η οποία πλησιάζει αρκετά αυτό που θα προσπαθήσουμε να πετύχουμε στην παρούσα φάση, είναι η προσπάθεια ανακάλυψης κειμενικών στοιχείων εντός ενός άρθρου που δηλώνουν ότι σε αυτό το άρθρο εντοπίζεται μία ψευδής είδηση (“Fake news”) (Tacchini, et al., 2017). Άλλες προσπάθειες επικεντρώθηκαν στο να μπορεί να γίνει

έλεγχος του κειμένου με σκοπό να εντοπιστούν τυχόν στερεοτυπικές απόψεις που μπορεί να εμπεριέχονται σε αυτό (Fairbanks, et al., 2018). Η προσπάθεια που ακολουθεί προσεγγίζει στη μέθοδο τις παραπάνω ως προς το κομμάτι του αυτόματου ελέγχου γραπτού λόγου, είναι όμως εντελώς διαφορετική ως προς το σκοπό της. Ο σκοπός του παρόντος επιστημονικού εγχειρήματος θα είναι η δημιουργία ενός προγράμματος, μίας «τεχνητής νοημοσύνης» που ευελπιστούμε ότι αναλύοντας δεδομένα μεγάλου όγκου και αυξανόμενα εξελισσόμενη με τη χρήση αλγορίθμων από συγκεκριμένα δεδομένα, θα μπορεί να εντοπίσει προπαγανδιστικά τεχνάσματα ή/και συγκεκριμένους τύπους προπαγάνδας, οι οποίοι είναι επιστημονικά ορισμένοι σαφώς και εντοπίζονται στο λόγο.

Αξίζει μετά από τα παραπάνω να αναρωτηθούμε:

1. Είναι δυνατόν ένα πρόγραμμα να εντοπίζει λέξεις, φράσεις, προτάσεις ή ομάδες των προηγούμενων που χρησιμοποιούνται για προπαγανδιστικούς σκοπούς και προκειμένου να «κατασκευάσουν την πραγματικότητα» και να δημιουργήσουν στερεοτυπικές αντιλήψεις στους δέκτες αυτής της πληροφορίας;
2. Σε ποιο βαθμό είναι δυνατόν και πως μπορεί να εξασφαλιστεί η μεγαλύτερη δυνατή επιτυχία σε ένα τέτοιο εγχείρημα;
3. Ποιες συγκεκριμένες τεχνικές μπορούν να εντοπιστούν με όσο το δυνατόν πιο αντικειμενικά κριτήρια;
4. Ποιο θεωρητικό πλαίσιο μπορεί να εξασφαλίσει τον, κατά το δυνατό, αντικειμενικότερο χαρακτήρα αυτών των κριτηρίων, με τα οποία θα εντοπίζει το πρόγραμμα τις προπαγανδιστικές τεχνικές;

Αρχικά, για να απαντηθούν οι συγκεκριμένες ερωτήσεις, το θεωρητικό πλαίσιο πρέπει να ορίζει και τις τεχνικές και το τρόπο που αυτές εκφράζονται σε επίπεδο λόγου, ως τεχνικές προπαγάνδας. Κατόπιν θα πρέπει να ακολουθήσει μία ποιοτική ανάλυση του λόγου με βάση τα καθορισμένα από πριν χαρακτηριστικά των τεχνικών προπαγάνδας που παρατηρούνται στο λόγο και καθορίζονται από το θεωρητικό πλαίσιο. Έπειτα, είναι απαραίτητο τα δεδομένα, που θα αποτελέσουν την «τροφή» με την οποία ο αλγόριθμος θα κάνει αυτόματα τις αναλύσεις, να εισαχθούν στο προγραμματιστικό περιβάλλον με τρόπο τέτοιο ώστε να φανερώνεται η εκάστοτε αιτία που οδήγησε στο συμπέρασμα ότι στο συγκεκριμένο κομμάτι γραπτού λόγου παρατηρείται μία συγκεκριμένη

τεχνική προπαγάνδας. Η επιλογή του κατάλληλου αλγορίθμου και των κατάλληλων παραμέτρων στον κώδικα είναι επίσης καθοριστική για την ακρίβεια των αποτελεσμάτων του. Τέλος τα κριτήρια αξιολόγησης εκτός από ποσοτικά θα πρέπει να ελεγχθούν και ποιοτικά.

Αυτό είναι ένα γενικό πλάνο απαιτήσεων ώστε να μπορεί να γίνει ορθή χρήση της τεχνητής νοημοσύνης στα πλαίσια του σκοπού που έχουμε θέσει. Το εγχείρημα είναι διεπιστημονικό και περίπλοκο γι' αυτό και κατά τη διαδικασία εφαρμογής των παραπάνω βημάτων θα υπάρχει έλεγχος και από επιβλέποντες ειδικούς στον τομέα της πληροφορικής και ειδικότερα της μηχανικής μάθησης και στον τομέα της ανάλυσης προπαγάνδας αντίστοιχα.

2. Τα στάδια μέχρι την υπολογιστικοποίηση¹ της ανάλυσης

Αρχικά θα παρατεθεί ένα στέρεο θεωρητικό πλαίσιο γύρω από την προπαγάνδα και τα τεχνάσματά της.

Πρώτα θα ορισθεί, έπειτα θα συγκριθεί με παραπλήσιους όρους προς αποφυγή συγχύσεων στον ορισμό και στην ανάλυση της προπαγάνδας και μετά θα εξεταστούν τα χαρακτηριστικά της και τα αποτελέσματά της.

Κατόπιν θα ακολουθήσει ακριβής ορισμός των επιμέρους τεχνασμάτων προπαγάνδας που πρόκειται να εντοπιστούν και μετά από ανάλυση κειμένων σε ποιοτικό επίπεδο θα προκύψουν τα δεδομένα που θα εισαχθούν ψηφιακά στο προγραμματιστικό περιβάλλον «Jupyter» (<https://jupyter.org/>), όπου θα ακολουθήσει επεξεργασία τους μέσω εντολών στη γλώσσα προγραμματισμού Python.

Της θεωρητικής ανάλυσης έπεται μία ανασκόπηση προηγουμένων ερευνών – εργασιών που επιχειρήσαν να αυτοματοποιήσουν την ανάλυση προπαγάνδας ή κάποιων τεχνικών της. Η εισαγωγή των δεδομένων, που θα προκύψουν έπειτα από ποιοτική ανάλυση βασισμένη στο θεωρητικό

¹ Ως υπολογιστικοποίηση ορίζεται η μετάβαση από την ανάλυση που πραγματοποιείται από άνθρωπο στην ανάλυση που πραγματοποιείται από υπολογιστή.

πλαίσιο και μετά από τον εντοπισμό της ύπαρξης ή μη τεχνασμάτων προπαγάνδας, τα οποία θα ορισθούν σαφώς, με τις ενδείξεις “Yes/No”, θα γίνει σε υπολογιστικό φύλλο του προγράμματος “Excel” (<https://products.office.com/el-gr/excel>). Αφού αποθηκευτούν τα δεδομένα σε μορφή – επέκταση “.xlsx”, θα περαστούν στο προαναφερθέν προγραμματιστικό περιβάλλον προκειμένου να εφαρμοστεί σε αυτά, έπειτα από επεξεργασία που απαιτείται και θα αναλυθεί στη συνέχεια, αλγόριθμος μηχανικής μάθησης.

Στο επόμενο κεφάλαιο θα περιλαμβάνονται η διαδικασία της ανάλυσης και του προγραμματισμού καθώς και τα ευρήματα και τα αποτελέσματα της διαδικασίας. Από το ποσοστό ακρίβειας πρόβλεψης αλλά και από τις λέξεις του κειμένου, που λειτούργησαν ως κριτήρια κατηγοριοποίησης των κειμένων σε αυτών που παρουσιάζεται κάποιο τέχνασμα προπαγάνδας και σε αυτών που δεν παρουσιάζεται, θα κριθεί πόσο επιτυχημένο ήταν το πείραμα αυτοματοποίησης μία βασικής ανάλυσης προπαγάνδας και θα εξαχθούν συμπεράσματα, διαπιστώσεις, παρατηρήσεις και εκτιμήσεις.

3. Η βασική δομή της επικοινωνίας

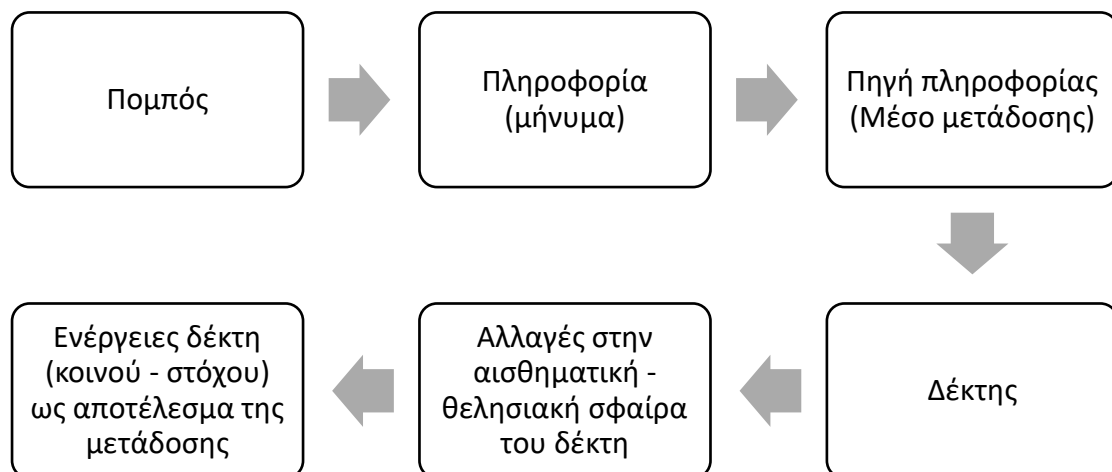
Η προπαγάνδα είναι μορφή επικοινωνίας και η καταγραφή των συστατικών της στοιχείων είναι απαραίτητη γιατί η ανάλυση του κάθε στοιχείου είναι διαφορετική διαδικασία και το κάθε ένα από αυτά έχει διαφορετικά χαρακτηριστικά και φύση.

Ο Dr. Harold D. Lasswell το 1948 προσπαθώντας να επεξηγήσει τη δομή και τη λειτουργία της επικοινωνίας σε κοινωνικό επίπεδο έφτιαξε ένα μοντέλο επικοινωνίας (Πίνακας 1) το οποίο αποτελεί, τρόπον τινά, μία «ανατομία» της μετάδοσης πληροφορίας μέσα στην κοινωνία καθώς διαχωρίζει τα βασικά διακριτά συστατικά της στοιχεία και καταδεικνύει τα βασικά επίπεδα ανάλυσής της (Lasswell, 1948, p. 216).

Ερώτηση	Στοιχείο	Είδος Ανάλυσης
Ποιος;	Πομπός (Απώτερη Προέλευση)	Ανάλυση ελέγχου της μετάδοσης
Τι διαδίδει;	Μήνυμα	Ανάλυση Περιεχομένου
Με ποιο μέσο;	Μέσο επικοινωνίας	Ανάλυση του Μέσου μετάδοσης
Σε ποιους;	Κοινό – στόχος (Target Audience)	Ανάλυση του κοινού
Με τι αποτέλεσμα;	Αποτελέσματα της διάδοσης	Ανάλυση των αποτελεσμάτων

Πίνακας 1 - Μοντέλο επικοινωνίας του Lasswell (Ελεύθερη Μετάφραση)

Ένα άλλο σχήμα επικοινωνίας εξίσου βασικής – απλοϊκής μορφής είναι αυτό που απεικονίζεται στον Πίνακα 2 (Κισελёв, 2011, p. 3).



Πίνακας 2 - Μοντέλο επικοινωνίας (Κισελёв, 2011, p. 3)

Τα μοντέλα αυτά επικοινωνίας λαμβάνουν υπόψη περίπλοκα μοντέλα ανάλυσης προπαγάνδας του Αμερικανικού στρατού. Συγκεκριμένα, τα σύγχρονα μοντέλα ανάλυσης προπαγάνδας MARCO² και SCAME³ που χρησιμοποιούνται από της μονάδες ψυχολογικών επιχειρήσεων του στρατού των ΗΠΑ (Department of the US Army, 2008) στηρίζονται πάνω στο μοντέλο του Lasswell έχοντας βέβαια εξελιχθεί ως την πολυπλοκότητα και τις λεπτομέρειές του.

Εν προκειμένω, θα αναλυθεί το περιεχόμενο του μηνύματος, η πληροφορία δηλαδή, μέσω αλγοριθμικής ανάλυσης κειμένου. Μας ενδιαφέρει επομένως να επικεντρωθούμε στον τρόπο ανάλυσης του μηνύματος από αυτά τα δύο μοντέλα. Εξετάζονται, λοιπόν, και καταγράφονται σε μορφή επίσημης αναφοράς ο σκοπός του μηνύματος, τα είδη των επιχειρημάτων που χρησιμοποιούνται σε αυτό, το «ηθικό» της πηγής με βάση το μήνυμα (υψηλό – χαμηλό) και εν τέλει έμμεσες πληροφορίες που μπορούν να εξαχθούν από αυτό (Department of the US Army, 2008, pp. 98-102). Σε αυτά τα στοιχεία με τη δική μας προσέγγιση θα μπορούσαν να απαντηθούν και να εξαχθούν περισσότερες και πιο ακριβείς πληροφορίες σε συνδυασμό και την ποιοτική ανάλυση με την οποία ως τώρα εξετάζονται. Επίσης το σημαντικότερο είναι ότι μπορεί να εξεταστεί πολύ μεγαλύτερος όγκος πληροφορίας για να βγουν συμπεράσματα για την προπαγάνδα που αναλύεται. Κάτι επίσης πολύ σημαντικό που μπορούμε να παρατηρήσουμε σε αυτό το σημείο είναι ότι η μέθοδος ανάλυσης του μηνύματος είναι και κάτι που διενεργείται ήδη ποιοτικά στην πράξη και αναζητά στοιχεία που θα προσπαθήσουμε να

² Message – Audience – Reaction/Response – Carrier – Origin (M.A.R.C.O) (Department of the US Army, 2008, p. 101)

³ Source – Content – Audience – Media – Effect (S.C.A.M.E) (Department of the US Army, 2008, p. 98)

αναζητήσουμε και εμείς όπως το είδος του επιχειρήματος που παρατηρείται στο μήνυμα.

4. Το θεωρητικό πλαίσιο της ανάλυσης

“η αρχή της σοφίας είναι ο ορισμός των εννοιών⁴”

4.1. Ορισμός

Ο όρος «προπαγάνδα» σήμαινε αρχικά «διάδοση»⁵ (Merriam-Webster, 2018) και είχε ουδέτερη σημασία. Ο όρος πρωτοεμφανίστηκε το 1622 στο «Ιερό Συμβούλιο για την Διάδοση της Πίστης» (“*Sacra Congregatio de Propaganda Fide*”) το οποίο δημιουργήθηκε από τον Πάπα Γρηγόριο τον 15^ο και έκτοτε έχει απωλέσει σταδιακά την ουδέτερη σημασία του, αποκτώντας αρνητική χροιά (O'Donnell & Jowett, 2014, p. 2), (Bernays, 1928, p. 21), (Gulseth, 2004, p. 26).

Ένας σύγχρονος ορισμός της προπαγάνδας είναι «η έκφραση απόψεων και ενεργειών που πραγματοποιούνται εσκεμμένα από άτομα ή ομάδες με σκοπό να επηρεάσουν τις απόψεις ή τις ενέργειες άλλων ατόμων ή ομάδων με προκαθορισμένους σκοπούς και μέσω ψυχολογικού τύπου χειραγώγησης.» - Lasswell όπως αναφέρεται σε (Ellul, 1965, p. xii).

Στη δέκατη έκδοση του λεξικού *Webster's Collegiate* η προπαγάνδα ορίζεται ως «ιδέες, γεγονότα ή ισχυρισμοί που διαδίδονται εσκεμμένα με σκοπό να εξυπηρετήσουν το σκοπό κάποιου ή να βλάψουν έναν αντίπαλο σκοπό.

Άλλος ορισμός από λεξικό: Σύνολο πληροφοριών, στερεοτυπικής ή παραπλανητικής φύσεως, που χρησιμοποιείται με σκοπό να διαδώσει ορισμένες αντιλήψεις (Oxford University Press, 2018).

⁴ «Αρχή Σοφίας ονομάτων επίσκεψις» - Αντισθένης ο Αθηναίος

⁵ Propaganda: Το θηλυκό γένος της γερονδιακής μορφής “propagandus” του απαρεμφάτου της λατινικής γλώσσας “propagare”, το οποίο μεταφράζεται ως «διαδίδειν». (Merriam-Webster, 2018)

Ο Leonard W. Doob την ορίζει ως μία προσπάθεια «να αλλάξουν προσωπικότητες και να ελεγχθούν άτομα σε σχέση με σκοπούς που θεωρούνται μη επιστημονικοί ή αμφιβόλου αξίας σε μία ορισμένη κοινωνία και συγκεκριμένη περίοδο.» – όπως αναφέρεται σε (Ellul, 1965, p. xiii).

Για τον όρο «προπαγάνδα» υπάρχουν πολλοί ορισμοί γι' αυτόν τον λόγο θα λάβουμε υπόψη και θεωρητικά και κατά την έρευνα τον πληρέστερο σύγχρονο ορισμό που εντοπίσαμε: «Η προπαγάνδα είναι ουδέτερα ορισμένη σαν μία συστηματοποιημένη μορφή σκόπιμης προσπάθειας πειθούς που επιχειρεί να επηρεάσει τα συναισθήματα, τις συμπεριφορές, τις απόψεις ή τις ενέργειες συγκεκριμένων κοινών – στόχων για ιδεολογικούς, πολιτικούς ή εμπορικούς σκοπούς με μία ελεγχόμενη μετάδοση μεροληπτικών (μονόπλευρων) μηνυμάτων (τα οποία μπορεί να είναι αληθή ή ψευδή) μέσω μαζικών «καναλιών» πληροφορίας.» (Nelson, 1996, pp. 222-223)

Ο λόγος που παραθέσαμε τόσους ορισμούς για την προπαγάνδα είναι ο σαφής καθορισμός κοινά αποδεκτών χαρακτηριστικών της προπαγάνδας από τους αναγνωρισμένους επιστήμονες στην επιστημονική κοινότητα που ειδικεύονται σε αυτόν τον τομέα και έχουν ασχοληθεί σε ερευνητικό επίπεδο με την προπαγάνδα. Ο συγκεκριμένος προκαθορισμένος σκοπός, το μεροληπτικό μήνυμα και κυρίως η προσπάθεια επιρροής της αντίληψης αποτελούν κοινή συνισταμένη στους επιστημονικά αποδεκτούς ορισμούς. Στα χαρακτηριστικά αυτά και στον ορισμό του Nelson ο οποίος αναφέρει και το χαρακτηριστικό των καναλιών πληροφορίας που απαιτούνται για να μεταδοθεί το μήνυμα, εν προκειμένω ο γραπτός λόγος, θα βασιστεί και το μετέπειτα θεωρητικό

πλαίσιο των τεχνασμάτων. Τα τεχνάσματα αυτά, όπως θα καταστεί σαφές στη συνέχεια, στοχεύουν στην επιρροή της αντίληψης με επιχειρήματα που μπορεί να βασίζονται ή να μη βασίζονται σε λογική επιχειρηματολογία και τεκμηρίωση, ενώ ενίοτε είναι ορθολογικοφανή, δηλαδή φαίνεται να είναι λογικά ενώ αποτελούν λογικές πλάνες.

4.2. Το μοντέλο προπαγάνδας Herman & Chomsky

Στην παρούσα προσπάθεια ανάλυσης κειμένων για τον εντοπισμό προπαγανδιστικών τεχνασμάτων, πολλά από τα κείμενα που χρησιμοποιούνται προέρχονται από διαφορετικά Μέσα για αυτό το λόγο έκρινα αναγκαίο να παραθέσω το μοντέλο των Herman & Chomsky το οποίο εξηγεί τα «φίλτρα» από τα οποία περνάει μία είδηση και διαμορφώνεται μέσα από αυτά η τελική μορφή που παρουσιάζονται οι ειδήσεις, η οποία εμπεριέχει τεχνάσματα προπαγάνδας. Το παρακάτω μοντέλο αποτελείται από πέντε παράγοντες που επηρεάζουν καθοριστικά τη διαμόρφωση των ειδήσεων σε ένα μέσο στις δημοκρατικές κοινωνίες.

- ❖ Μέγεθος, ιδιοκτησία και προσανατολισμός του κέρδους ενός ΜΜΕ
- ❖ Η προσέγγιση του διαφημιζόμενου κοινού
- ❖ Πηγές πληροφοριών του Μέσου (με ιδιαίτερη σημασία)
- ❖ Αρνητικές αντιδράσεις σε μία δήλωση ή πρόγραμμα που παρουσιάστηκε στο Μέσο
- ❖ Άντι-ιδεολογία (πρωτοεμφανίστηκε ως «Άντι-κομμουνισμός σαν μηχανισμός ελέγχου»)

(Herman & Chomsky, 1994) – Ελεύθερη μετάφραση

Το μοντέλο περιγράφεται επιγραμματικά διότι εν προκειμένω δεν αφορά τον τρόπο ανάλυσης των κειμένων αλλά παρατίθεται για να αποδειχθεί ότι προπαγάνδα υπάρχει και μάλιστα παρατηρείται συστηματικά από τα ΜΜΕ στις δημοκρατικές κοινωνίες. Επιπρόσθετα, τα Μέσα έχουν κίνητρα για τη διενέργεια αυτής τα οποία εντάσσονται σε σταθερές τυπολογίες.

4.3. Τα χαρακτηριστικά της προπαγάνδας εν γένει

4.3.1. Αποσαφήνιση και διαχωρισμός του όρου βάσει χαρακτηριστικών

Ένας λεπτομερέστερος ορισμός της προπαγάνδας θα ήταν χρήσιμος ώστε να προχωρήσουμε στο βασικό εγχείρημα μας και αν αυτοματοποιήσουμε την εύρεση κάποιων τεχνικών της.

Κατ' αρχήν πρέπει να αποσαφηνιστεί η διαφορά της με παρεμφερείς έννοιες καθώς και η τυχόν σύνδεση της με αυτές.

Σύγκριση όρων «προπαγάνδα – πειθώ»

Η πειθώ για παράδειγμα, που είναι παρεμφερής έννοια, απαιτεί διάδραση που στόχο έχει το να πειστεί λιγότερο ή περισσότερο η μία ή η άλλη πλευρά για τις απόψεις της μία ή της άλλης πλευράς (O'Donnell & Jowett, 2014, p. 1). Μπορεί η προπαγάνδα να έχει σκοπό να πείσει και να περιλαμβάνει κατά μία στενή έννοια την πειθώ όμως κάθε προσπάθεια πειθούς δεν εντάσσεται αυτομάτως στην προπαγάνδα καθώς η πρώτη λαμβάνει χώρα και στη συζήτηση με αναφορές επί του θέματος, παρατηρήσεις και ορθολογικά επιχειρήματα (Pratkanis & Aronson, 2001). Στην προπαγάνδα από την άλλη απαιτείται μία διάδοση μηνύματος που εκμεταλλεύεται στερεότυπα και συναισθήματα κάποιες φορές και ήδη υπάρχοντα στο στόχο (άτομο ή κοινό) της προπαγάνδας. Ιδανικά, έχει ως

απώτερο στόχο να πείσει μεν το κάθε άτομο ξεχωριστά σε ένα κοινό – στόχο για ορισμένες απόψεις κατά τέτοιον τρόπο ώστε το ίδιο το άτομο να έχει την εντύπωση ότι τις απέκτησε χωρίς επιρροή, «από μόνο του» και βασισμένο στην παρατηρητικότητα του και τη δική του αντίληψη με τη χρήση της οποίας έφτασε να καταλάβει αυτήν την «αλήθεια» (Pratkanis & Aronson, 2001). Αυτό δε συμβαίνει σε μία προσπάθεια πειθούς ενός ατόμου από άλλο καθώς λαμβάνει χώρα μία διαδικασία διάδρασης όχι όμως και διάδοσης. Ενδεχομένως λοιπόν οι δύο έννοιες ,πειθώ και προπαγάνδα, να σχετίζονται ως προς τον στόχο να πειστεί ένα υποκείμενο για κάτι υπάρχον όμως παρατηρήσιμες διαφορές στα χαρακτηριστικά τους.

Σύγκριση όρων «προπαγάνδα – επιχείρημα»

Η σχέση της έννοιας της προπαγάνδας με την επιχειρηματολογία είναι επίσης άξιο να σημειωθεί. Κατ' αρχήν δεν μπορεί να υπάρξει σύγχυση των εννοιών καθώς η προπαγάνδα απαιτεί διάδοση που έχει βραχυπρόθεσμους και μακροπρόθεσμους σκοπούς οι οποίοι μπορεί να διαφέρουν και μεταξύ τους και εν τέλει, αν την εκλάβουμε ως ένα μεροληπτικό μήνυμα που σκοπό έχει να διαστρεβλώσει την αντίληψη της πραγματικότητας υπέρ ή σε βάρος κάποιου ατόμου ή ομάδας ατόμων, τότε πρέπει να λάβουμε υπόψη και τον τρόπο που χρησιμοποιείται ένα οποιοδήποτε μήνυμα πριν το χαρακτηρίσουμε ως προπαγάνδα (Ross, 2002, p. 20).

Ένα επιχείρημα μπορεί να είναι απολύτως ορθολογικό και δεν είναι απαραίτητα προβληματικό ως προς τη λογική του ούτε επικαλείται αναγκαστικά το συναίσθημα. Αντίθετα, στην προπαγάνδα την θέση της λογικής λαμβάνει η προβληματική ή εσφαλμένη λογική ή η επίκληση στο συναίσθημα ή ένας συνδυασμός των δύο (Shabo, 2008, p. 5). Όταν το

μήνυμα που διαδίδεται είναι μεροληπτικό, όπως ορίστηκε από τον Nelson ότι ισχύει στην προπαγάνδα, τότε η λογική του προκύπτει ότι δεν είναι αμερόληπτη και ως προς αυτό είναι επομένως προβληματική. Το επιχείρημα εντούτοις μπορεί να αποτελεί και μέθοδο προπαγάνδας όταν έχει ορισμένα χαρακτηριστικά και χρησιμοποιείται για την αναγνώριση προπαγάνδας (Central Intelligence Agency, 1969, p. 3). Ένα παράδειγμα μπορεί να είναι ένα επιστημολογικά ελλιπές επιχείρημα του τύπου: «καμία ασπιρίνη δεν έχει αποδειχθεί καλύτερη από αυτήν (τη Χ ασπιρίνη που διαφημίζεται εν προκειμένω) και πρέπει να την αγοράσετε» (Ross, 2002, p. 23). Πρόκειται για ένα επιχείρημα που αποτελεί εμπορική προπαγάνδα (Γεωργαλάς, 1967, p. 85) και παραλείπει σημαντικές για τη σημασία του μηνύματος λεπτομέρειες όπως το ότι κάθε ασπιρίνη είναι η ίδια ουσία με τα ίδια χαρακτηριστικά και ιδιότητες ανεξαρτήτως μάρκας δεδομένου ότι δεν έχουν προστεθεί άλλα συστατικά και παραμένει ασπιρίνη (Ross, 2002, p. 23). Έτσι αποδεικνύεται πως ένα επιχείρημα είναι δυνατόν να αποτελεί και μορφή προπαγάνδας. Η διαπίστωση αυτή είναι σημαντική για την έρευνα μας αφού κάποια τεχνάσματα προπαγάνδας που ορίζονται από το IPA (Institute for Propaganda Analysis, 1938, p. 19) αποτελούν κατά άλλους και τύπους επιχειρημάτων που επικαλούνται το συναίσθημα.

Σύγκριση όρων «προπαγάνδα – ψυχολογικός πόλεμος»

Ένας επιστημονικός ορισμός για τον όρο «ψυχολογικός πόλεμος» συναντάται στο εγχειρίδιο ψυχολογικού πολέμου της αμερικανικής στρατιωτικής μονάδας ειδικών επιχειρήσεων. Εκεί, ως ψυχολογικός πόλεμος ορίζεται «η απόπειρα επιρροής της συμπεριφοράς ενός κοινού-στόχου (target audience) με σκοπό εξυπηρετηθούν οι απώτεροι σκοποί του φορέα που τον διεξάγει.» (US Army, 2005, p. 9). Το χαρακτηριστικό της

επιρροής μέσω ελεγχόμενου μηνύματος επιδεικνύει την συγγένεια με τον όρο προπαγάνδα ή και την υπαγωγή σε αυτόν.

Στη Βρετανική Εγκυκλοπαίδεια, ως ψυχολογικός πόλεμος ορίζεται η χρήση προπαγάνδας ενάντια σε έναν εχθρό, υποστηριζόμενης από στρατιωτικά, οικονομικά και πολιτικά μέσα όπου και όπως απαιτηθεί (The Editors of Encyclopaedia Britannica, 2015). Σε αυτόν τον ορισμό παρατηρούμε πως μία προπαγανδιστική επιχείρηση μπορεί να θεωρηθεί και επιχείρηση ψυχολογικού πολέμου όταν συντρέχουν κάποιες προϋποθέσεις όπως η γεωγραφική θέση της διενέργειας της προπαγάνδας σε εχθρική περιοχή ή γενικότερα η διενέργειά της σε εχθρικό και επιθετικό περιβάλλον. Επίσης όμως παρατηρούμε ότι αυτές οι συνθήκες δεν σχετίζονται με το μήνυμα που προπαγανδίζεται και το οποίο εν προκειμένω μας ενδιαφέρει. Επομένως, μπορούμε να συμπεράνουμε ότι οποιοδήποτε είδους ανάλυση της προπαγάνδας σε επίπεδο μηνύματος μπορεί να εφαρμοστεί και όταν η προπαγάνδα αυτή αποτελεί κομμάτι ψυχολογικής επιχείρησης. Το ίδιο μπορούμε να συμπεράνουμε και από τον ορισμό που δίνει ο Paul M. A. Linebarger για τον ψυχολογικό πόλεμο που περιλαμβάνει την «χρήση προπαγάνδας εναντίον του εχθρού, μαζί με επιχειρησιακά μέτρα» (O'Donnell & Jowett, 2014, p. 232).

Εν κατακλείδι, μία άποψη που ορίζεται ως επιχείρημα μπορεί να χρησιμοποιηθεί για προπαγάνδα. Μία προσπάθεια πειθούς, όταν συντρέχουν και άλλες προϋποθέσεις μπορεί να θεωρηθεί προπαγανδιστική επιχείρηση και όταν αυτό συμβαίνει σε επίπεδο κράτους και αφορά την ασφάλεια του τότε μπορεί να θεωρηθεί και ψυχολογική επιχείρηση. Η εννοιολογική αυτή σύνδεση όμως δεν επηρεάζει την προσπάθεια που θα ακολουθήσει στο πρακτικό της μέρος.

4.3.2. Παραδοχές – Παρατηρήσεις για την προπαγάνδα

Σύμφωνα με το εγχειρίδιο για την μελέτη της προπαγάνδας από το πανεπιστήμιο της Οξφόρδης, μπορούμε από τη μέχρι τώρα μελέτη για τον όρο να του αποδώσουμε ορισμένα χαρακτηριστικά.

«Η προπαγάνδα δεν είναι εγγενώς κακή ή ανήθικη» (Auerbach & Castronovo, 2013, p. 5). Μπορούμε να ελέγξουμε αυτήν παραδοχή χρησιμοποιώντας ιστορικά παραδείγματα που αποδεικνύουν τη μη εγγενή ανηθικότητα και φαυλότητα της χρήσης της προπαγάνδας. Η προπαγάνδα για την απόκτηση εργασιακών δικαιωμάτων, για την κατάργηση της δουλείας και για την προστασία του περιβάλλοντος είναι παραδείγματα χρήσης προπαγανδιστικών τεχνικών μέσα από την ιστορία αλλά και τη σημερινή δημόσια σφαίρα που έχουν «καλό και ηθικό» σκοπό. Ένα τέτοιο συμπέρασμα πρέπει βέβαια να συνοδεύεται με τον απαραίτητο σκεπτικισμό λόγω της υποκειμενικής σημασίας των όρων «καλό – κακό» και «ηθικό – ανήθικο». Αλλά παρατηρείται σαφώς.

Η προπαγάνδα εμπεριέχει την έννοια της διάδοσης αλλά δεν είναι απαραίτητο κάτι που διαδίδεται να αποτελεί αναγκαστικά και προπαγάνδα (Auerbach & Castronovo, 2013, p. 5). Ένα δελτίο καιρού για παράδειγμα διαδίδει πληροφορία αλλά δεν μπορεί να είναι προπαγάνδα αν δεν αφορά κάποιου είδους γενικότερη άποψη ή ιδεολογία και να πείσει και να επηρεάσει υπέρ ή κατά αυτής, όπως για παράδειγμα την κλιματική αλλαγή.

Για τον ίδιο λόγο μπορούμε να συμπεράνουμε ότι η σχέση μεταξύ προπαγάνδας και πληροφορίας είναι ρευστή και ποικίλλει ανάλογα με το γενικό πλαίσιο στο οποίο εντάσσεται η πληροφορία και τη λειτουργία της πληροφορίας (Auerbach & Castronovo, 2013, p. 5). Αυτό και το προηγούμενο χαρακτηριστικό μας οδηγούν σε έναν περιορισμό που

τίθεται στην έρευνα, δηλαδή η ανάλυση προπαγάνδα να γίνεται σε κείμενα και φράσεις που εντάσσονται σε ένα πολιτικό πλαίσιο, διότι σκοπός είναι αντίστοιχα η αναζήτηση πολιτικής προπαγάνδας.

Ένα άλλο χαρακτηριστικό που είναι διαφωτιστικό για την έρευνα μας είναι ότι παρά το γεγονός ότι «η προπαγάνδα δεν είναι μία αυστηρή κατηγορία επικοινωνίας με ακριβή και απόλυτα τυπικά χαρακτηριστικά», συγκεκριμένες τεχνικές της μπορούν να παρατηρηθούν και να μελετηθούν με αντικειμενικό τρόπο έχοντας όμως μεταβλητά αποτελέσματα (Auerbach & Castronovo, 2013, p. 6). Τέτοιες τεχνικές θα μελετηθούν στην παρούσα έρευνα και συγκεκριμένα του IPA που αναφέρονται και στο εγχειρίδιο του πανεπιστημίου της Οξφόρδης.

Η προπαγάνδα, επίσης, μπορεί να αλλάζει από μέσο σε μέσο όμως δεν μπορεί να οριστεί μέσω των χαρακτηριστικών ενός συγκεκριμένου μέσου (Auerbach & Castronovo, 2013, p. 7).

Η προπαγάνδα στην ανάλυσή της απαιτεί ιδιαίτερη εξέταση στη ροή της πληροφορίας (πως διαδίδεται) και το περιεχόμενο της (τι διαδίδεται) (Auerbach & Castronovo, 2013, p. 8). Είναι ιδιαίτερα σημαντικό πως σε επίπεδο κειμένου υπάρχουν τα εργαλεία προγραμματισμού ώστε να εντοπίζεται και το πως αλλά και το τι διαδίδεται. Πρέπει ξανά να τονιστεί ότι αυτό μέχρι σήμερα μπορεί να επιτευχθεί σε επίπεδο κειμένου, λέξεων αλλά ακόμη και εκεί απαιτείται μία περαιτέρω βαθύτερη και ποιοτική ανάλυση.

Η προπαγάνδα συμπεραίνεται επίσης πως μπορεί να παράγει ακούσια αποτελέσματα και από τη μεριά του πομπού της πληροφορίας και από τη μεριά του δέκτη (Auerbach & Castronovo, 2013, p. 10). Δεν είναι σαφώς ορισμένο και απόλυτα προβλέψιμο το πως ακριβώς θα επηρεάσει της

απόψεις του κοινού στόχου της μία συγκεκριμένη τεχνική προπαγάνδας και ειδικά σε ατομικό επίπεδο.

Σε αντίθεση με την ευρεία πεποίθηση ότι η προπαγάνδα συναντάται κατά κύριο λόγο σε ολοκληρωτικά και ανελεύθερα καθεστώτα, έχει συμπεραθεί και παρατηρηθεί ότι αποτελεί αναπόσπαστο κομμάτι των δημοκρατικών κοινωνιών (Auerbach & Castronovo, 2013, p. 11; Bernays, 1928; Herman & Chomsky, 1994). Γνωρίζουμε με βεβαιότητα επομένως ότι είτε σε επίπεδο Μέσων είτε σε επίπεδο πολιτικού λόγου θα συναντήσουμε προπαγάνδα και μάλιστα πολλών μορφών καθώς και αντικρουόμενες προπαγανδιστικές προσπάθειες. Άλλη διαπίστωση που έχει επισημανθεί είναι πως η μελέτη της προπαγάνδας παραμένει σχετική με τη σύγχρονη πραγματικότητα και το ίδιο αναγκαία και σε κάθε περίπτωση θα συνεχίσει να παραμένει έτσι και στο μέλλον (Auerbach & Castronovo, 2013, p. 11).

Εκτός από τις ανιχνεύσιμες μορφές προπαγάνδας στο σημειολογικό κομμάτι του μηνύματος και σε επίπεδο λόγου υπάρχουν μορφές που σχετίζονται με την ερμηνεία και το ερμηνευτικό πλαίσιο και είναι σαφώς πολυπλοκότερες των πρώτων, όπως για παράδειγμα η μίξη γεγονότος και ερμηνείας του γεγονότος (Marlin, 2002, p. 110). Τέτοιες τεχνικές δεν είναι δυνατόν να εντοπιστούν με την ίδια ευκολία και κυρίως με τα υπάρχοντα εργαλεία γι' αυτό και δεν επιλέχθηκαν αυτές ή κάποιες άλλες τεχνικές εξίσου περίπλοκες αντί των τεχνασμάτων που ορίζονται σαφώς και είναι ανιχνεύσιμα σε επίπεδο λεξιλογίου ενός κειμένου.

4.4. Τα προς ανάλυση τεχνάσματα

Το 1938, το Ινστιτούτο Ανάλυσης Προπαγάνδας στις Η.Π.Α. έπειτα από αναλύσεις που κράτησαν ένα περίπου χρόνο (Οκτώβριος 1937 – Οκτώβριος 1938) συμπεριέλαβε τα δεδομένα και τις αναλύσεις σε 5 τόμους στους οποίους αναλύεται ενδελεχώς η προπαγάνδα αλλά και γεγονότα που σχετίζονται με αυτή δηλαδή την ιστορία της. Στον πρώτο τόμο ορίζεται για πρώτη φορά μία σειρά από προπαγανδιστικά τεχνάσματα που, όπως αναφέρεται στον τόμο, επικαλούνται το συναίσθημα και σε αυτήν την επίκληση εντοπίζεται η ισχύ τους και με αυτόν τον τρόπο εξαπατούν ένα μεγάλο μέρος της κοινής γνώμης κατά τη διάρκεια που προσλαμβάνει την πληροφορία για το γεγονός για το οποίο ενημερώνεται (Institute for Propaganda Analysis, 1938, p. 19). Τα κύρια αυτά προπαγανδιστικά τεχνάσματα που παρατηρούνται είναι τα εξής:

- 1) Name calling Device (Χαρακτηρισμός)
- 2) Glittering Generalities Device (Απαστράπτουσες Γενικότητες)
- 3) Transfer Device (Μεταφορά)
- 4) Testimonial Device (Μαρτυρία)
- 5) Plain Folks Device (Απλός Άνθρωπος)
- 6) Card Stacking Device («Επιλογή του τραπουλόχαρτου»)
- 7) Band Wagon Device (Φαινόμενο του συρμού)

(Institute for Propaganda Analysis, 1938), (Sproule, 2001), (Auerbach & Castronovo, 2013, p. 370), (US Army, 1979), (Sumanat & Dispanya, 2016)⁶

Ο ορισμός αυτών των τεχνικών δε σημαίνει ότι ορίζεται επακριβώς και αποκωδικοποιείται κάθε τεχνική και στρατηγική προπαγάνδας και

⁶ Οι ελληνικοί όροι αποτελούν ελεύθερη μετάφραση των πρωτοτύπων.

πειθούς, η γνώση αυτών δεν επαρκεί σε καμία περίπτωση για τη δημιουργία ενός «αμυντικού μηχανισμού» στο άτομο μέσω του οποίου θα αντιλαμβάνεται την πλάνη σε κάθε προπαγανδιστική επιχείρηση (Auerbach & Castronovo, 2013, p. 6) και σίγουρα αυτή η αποκωδικοποίηση των τεχνασμάτων αυτών δεν μπορεί να καλύψει, μέχρι σήμερα τουλάχιστον, μία ενδελεχή ανάλυση οποιασδήποτε στρατηγικής προπαγάνδας.

Παρ' όλα αυτά, τα αναγνωρίσιμα κατά μήκος του κειμένου αυτά τεχνάσματα μπορούν να αποτελέσουν μία βασική εκπαίδευση για έναν έλεγχο ενός κομματιού λόγου και μία κριτική επάνω σε αυτό για τυχόν ύπαρξη προπαγάνδας (Auerbach & Castronovo, 2013, p. 370). Η ανάλυση λοιπόν με βάση αυτά θα μπορούσε να οδηγήσει σε ένα είδος φιλτραρίσματος των κειμένων όταν προκύπτει η παρατήρηση ενός ή περισσότερων από τα παραπάνω τεχνάσματα και να αναδεικνύει ανάμεσα σε έναν μεγάλο όγκο κειμένου συγκεκριμένα χωρία που εντοπίζονται αυτές οι τεχνικές.

Τα επτά αυτά προπαγανδιστικά τεχνάσματα μπορούν να εντοπιστούν τόσο σε λεκτικό όσο και σε εικονικό επίπεδο. Αυτό όμως που θα εξεταστεί στην επικείμενη ανάλυση είναι η παρουσία ορισμένων από αυτά αποκλειστικά στο λεκτικό επίπεδο. Ο λόγος για αυτή την επιλογή είναι πως ο σκοπός της ανάλυσης μας είναι να προκύψουν δεδομένα που θα δώσουν στο πρόγραμμα τη δυνατότητα να εντοπίζει προπαγανδιστικά τεχνάσματα αποκλειστικά στο γραπτό λόγο με μία μηχανική αλγοριθμική ανάλυση κειμένου.

Ο «χαρακτηρισμός» (Name calling) είναι μία τεχνική που επιχειρεί να μας οδηγεί στο να κρίνουμε κάτι ή κάποιον χωρίς να δίνουμε σημασία στα

στοιχεία που θα έπρεπε να λαμβάνονται υπόψη για να διαμορφωθεί μία άποψη για το χαρακτηριζόμενο υποκείμενο, αντικείμενο, άποψη, ιδέα ή μία ομάδα από τα προηγούμενα (Institute for Propaganda Analysis, 1938, p. 5). Όταν λαμβάνει χώρα αυτή η τεχνική υποκείμενα, αντικείμενα, απόψεις, πεποιθήσεις ή ομάδες από τέτοια χαρακτηρίζονται με επίθετα ή ονόματα που είναι, συναισθηματικά, αρνητικά φορτισμένα στην κοινή γνώμη γενικότερα ή στο κοινό-στόχο της προπαγάνδας ειδικότερα κατά την πεποίθηση του χαρακτηριζόντος. Ορίζοντας σαφέστερα τι σημαίνει αρνητικά φορτισμένος χαρακτηρισμός από συναισθηματική άποψη, τον ορίζουμε ως ένα επίθετο ή όνομα απέναντι στο οποίο το μεγαλύτερο ή ένα μεγάλο ποσοστό του κοινού – στόχου που δέχεται την προπαγάνδα έχει αρνητική στάση. Πρόκειται για μία τεχνική η οποία όταν εφαρμόζεται κατά ενός φυσικού υποκειμένου αποτελεί ένα είδος επιχειρήματος – λογικής πλάνης που ονομάζεται “ad hominem” στα λατινικά ή αλλιώς «επιχείρημα ενάντια στον άνθρωπο» (Marlin, 2002, p. 110). Ενδέχεται πολλές φορές να είναι ανεπιτυχής και να καταστεί προφανής στο κοινό και να θεωρηθεί το υποκείμενο που τη χρησιμοποιεί αγενές και γενικώς να υπάρξει αρνητική στάση απέναντί του, παρ’ όλα αυτά όταν το χαρακτηριζόμενο υποκείμενο ή άποψη δεν είναι δημοφιλές τότε οι χαρακτηρισμοί εναντίον του έχουν την τάση να γίνονται ευκολότερα αποδεκτοί (Shabo, 2008, p. 49).

Η τεχνική του χαρακτηρισμού μπορεί να είναι άμεση και το όνομα να αποδίδεται άμεσα στον χαρακτηριζόμενο αντίπαλο ή στην αντίπαλη άποψη, ιδέα ή πρακτική. Ενδέχεται επίσης να είναι έμμεση και να επιχειρείται συνήθως μέσω σαρκασμού ή γελοιοποίησης (US Army, 1979, pp. 1-2) έτσι ώστε ο χαρακτηρισμός να επιτυγχάνεται από το κοινό μέσω μιας προβλέψιμης αντίδρασής του.

Παραδείγματα επιθέτων ή κυρίων ονομάτων που μπορούν να αποτελέσουν ευθείς χαρακτηρισμούς είναι π.χ. «φασίστας», «κομμούνι», «ουτοπιστής», «δικτάτορας», «ολιγάρχης», «αντιδημοκρατικός», «ολοκληρωτικό» κ.α. Κάθε επίθετο και όνομα με αρνητική σημασία μπορεί να αποτελέσει χαρακτηρισμό.

Εν τέλει, ο στόχος και η ουσία της τεχνικής είναι να επιτευχθεί ένας στερεοτυπισμός (US Army, 1979, p. 4), με την ευρεία έννοια του όρου, δηλαδή να κατηγοριοποιηθεί ο αντίπαλος ή η αντίπαλη άποψη, ιδέα ή πρακτική στην κοινή γνώμη ή στην πλειοψηφική γνώμη του κοινού – στόχου σε μία «κατηγορία» υποκειμένων, ιδεών ή πρακτικών απέναντι στις οποίες υπάρχει γενικώς αρνητική στάση και να ανακαλείται με βάση αυτή.

Οι «απαστράπτουσες γενικότητες» (Glittering Generalities) ως μέθοδος προπαγάνδας συναντάται όταν λέξεις με αόριστο και ασαφές περιεχόμενο που επιδέχονται πολλών και ποικίλων ερμηνειών χρησιμοποιούνται για την επίτευξη της αποδοχής και της θετικής στάσης απέναντι στις απόψεις αυτού που τις χρησιμοποιεί. Αυτές οι λέξεις αποτελούν συνηθέστερα κοινώς αποδεκτά και διαχρονικά ιδανικά (Institute for Propaganda Analysis, 1938, p. 6) ή αλλιώς «ευχάριστοι όροι με ασαφές περιεχόμενο» (Shabo, 2008, p. 30). Τέτοια παραδείγματα είναι «δημοκρατία», «ανθρωπισμός», «πατρίδα», «κοινωνική δικαιοσύνη», «ανάπτυξη», «ελευθερία» κ.α. Άλλο παράδειγμα αποτελεί η παρακάτω φράση στην αγγλική γλώσσα (οι λέξεις με το ασαφές περιεχόμενο είναι οι υπογραμμισμένες) :

I stand for freedom: for a strong nation, unrivaled in the world. My opponent believes we must compromise on these ideals, but I believe they are our birthright.

Η αυτοματοποίηση της εύρεσης τέτοιων λέξεων ή ομάδων λέξεων, που εδώ ταυτοποιήθηκαν με ποιοτική ανάλυση, με λεξικό θα μπορούσε να αποτελέσει μία τεχνική αυτόματης εύρεσης τεχνικών προπαγάνδας. Το να ψάχνει κάποιος για «απαστράπτουσες γενικότητες» συνίσταται απλά στην έρευνα για ασαφείς και θετικά συναισθηματικά φορτισμένες λέξεις η ιδιαίτερη σημασία των οποίων δεν εξηγείται μέσα σε ένα μήνυμα (Shabo, 2008, p. 32).

Η απόδοση του θετικού ή αρνητικού συναισθήματος σε κάποιον ή κάτι, η οποία περιγράφεται στις δύο προηγούμενες τεχνικές συνοψίζεται σε μία γενικότερη τεχνική: την «μεταφορά» (transfer) (Institute for Propaganda Analysis, 1938, p. 6). Χαρακτηριστικά και συναισθήματα μεταφέρονται σε κάποιον ή κάτι μέσω μίας σύνδεσης μέσα στο λόγο, με την ευρεία έννοια του όρου (Cole, 1988, p. 5). Η σύνδεση θετικά ή αρνητικά φορτισμένων εννοιών, προσώπων ή ομάδων που γίνεται άμεσα ή έμμεσα στον γραπτό, προφορικό και εικονικό λόγο με έννοιες ή πρόσωπα ή ομάδες προσώπων και απόψεων και αποτελεί μία «μεταφορά» συναισθημάτων από τα πρώτα στα δεύτερα συνοψίζει την τεχνική. Σε αυτήν την τεχνική είναι σημαντικά τα στερεότυπα, οι απόψεις και οι συναισθηματικές στάσεις, τα οποία είναι διαδεδομένα και αποδεκτά εκ των προτέρων στο κοινό – στόχο, για την επίτευξη του στόχου της προπαγάνδας από αυτόν που την διενεργεί. Στο εγχειρίδιο προπαγάνδας του πανεπιστημίου της Οξφόρδης αναφέρεται χαρακτηριστικά, είναι προϋπόθεση για την αποτελεσματικότητα της προπαγάνδας να χρησιμοποιεί ένα πλούσιο εύρος συναισθημάτων πέρα από τα αρνητικά συναισθήματα όπως το

φόβο, το μίσος και το φθόνο και να συμπεριλαμβάνει συναισθήματα ευχαρίστησης, χαράς, υπερηφάνειας και γενικότερα και θετικά συναισθήματα (Auerbach & Castronovo, 2013, p. 10). Επομένως η τεχνική της μεταφοράς συναντάται και πρέπει να χρησιμοποιείται με διάφορες μορφές και τα συναισθήματα που μεταφέρονται μπορεί να ποικίλουν ακόμη και από φράση σε φράση.

Η τεχνική της Μαρτυρίας αφορά την ισχυροποίηση της θετικής ή αρνητικής στάσης απέναντι σε μία γνώμη και εμφανίζεται όταν κάποιος αναγνωρίσιμο είτε θετικά είτε αρνητικά στον αναγνώστη άτομο παραθέτει μία άποψη ή ισχυρίζεται πως είναι σωστή ή λανθασμένη και αυτό αναφέρεται σε ένα κείμενο. Ο προπαγανδιστής δηλαδή εκμεταλλεύεται με αυτό το τέχνασμα μαρτυρίες που συναινούν ή εναντιώνονται σε μία οποιαδήποτε λογής άποψη (Institute for Propaganda Analysis, 1938, p. 6). Είναι παρόμοια με το είδος του επιχειρήματος που ονομάζεται επίκληση στην αυθεντία και για να μην δημιουργηθεί σύγχυση σε θεωρητικό επίπεδο στην ανάλυση η κατηγορία προπαγάνδας που θα εντοπίζεται θα είναι είτε Μαρτυρία είτε επίκληση στην αυθεντία (Sproule, 2001, p. 136) και θα καταγραφεί στο αρχείο δεδομένων ως εξής : “Testimonial/Appeal to Authority”. Το άτομο που μαρτυρά υπέρ ή κατά μίας άποψης ενδέχεται και να γνωρίζει πολύ λίγα για αυτήν ή να μην είναι ο ειδήμων να τοποθετηθεί σχετικά με ένα θέμα και όμως επηρεάζει (Marlin, 2002, p. 105). Ενδέχεται ακόμη και να μην έχει πει την φράση που του αποδίδεται, οπότε και πρόκειται για εξαπάτηση και προπαγάνδα παραπληροφόρησης (Cole, 1988, p. 328).

Η τεχνική του «απλού ανθρώπου»(Plain Folks) είναι η παρουσίαση ενός προσώπου ή ομάδας προσώπων ως κομμάτι του «απλού λαού» (Institute for Propaganda Analysis, 1938, pp. 6-7) που κάνει απλές καθημερινές

εργασίες, ή δραστηριότητες της καθημερινής ζωής του μέσου ανθρώπου και σκέφτεται σαν αυτόν και με την απλότητα που σκέπτεται αυτός. Η τεχνική αυτή μπορεί να συνδεθεί και με το λαϊκισμό ως φαινόμενο και από την έρευνα που ακολουθεί προέκυψε ως παρατήρηση ότι εντοπίζεται έμμεσα περισσότερο και λιγότερο με άμεσα ρητορικά λεκτικά σχήματα και ως επί το πλείστον στον εικονικό λόγο, δηλαδή σε εικόνες και βίντεο. Αυτή είναι η αιτία που δεν θα συμπεριληφθεί και δε θα αναλυθεί στην παρούσα έρευνα, χωρίς αυτό να σημαίνει ότι σε μελλοντικές δε θα μπορούσε με κανέναν τρόπο να βρεθεί μία μέθοδος που να εντοπίζει αυτό το τέχνασμα προπαγάνδας αυτόματα.

Το ίδιο ισχύει και για το τέχνασμα της «επιλογής του τραπουλόχαρτου» (Card Stacking). Ο λόγος για την εξαίρεση του όμως διαφέρει. Το τέχνασμα αυτό επιστρατεύει ποικίλες τεχνικές εξαπάτησης. Παραλείπονται γεγονότα (Sproule, 2001, p. 136) και στοιχεία εντέχνως και δίνει λιγότερη ή περισσότερη έμφαση σε γεγονότα (Institute for Propaganda Analysis, 1938, p. 7). Αποτελεί την έμφαση και τη μεροληπτική ανάδειξη ορισμένων εκφάνσεων και πλευρών της πραγματικότητας με ταυτόχρονη έντεχνη μείωση της σημασίας ή παντελή αποσιώπηση άλλων με σκοπό τη διαστρέβλωση της εικόνας της πραγματικότητας στο νου του δέκτη της πληροφορίας για κάποιον συγκεκριμένο απώτερο σκοπό.

Τελευταία τεχνική, η οποία πρόκειται να συμπεριληφθεί στην ανάλυση είναι το «φαινόμενο του συρμού» (Bandwagon effect) (Institute for Propaganda Analysis, 1938, p. 7). Είναι η τεχνική κατά την οποία κάποιος προσπαθεί να προπαγανδίσει υπέρ μιας άποψης διαδίδοντας με κάποιον τρόπο ότι αυτή η άποψη υποστηρίζεται από πολλούς ή από την πλειοψηφία ή από σχεδόν όλους (Marlin, 2002, p. 106). Αντίστροφα ότι προπαγανδίζει κατά μιας άποψης υποστηρίζει με κάποιον τρόπο ότι

αυτήν την υποστηρίζουν και την πιστεύουν λίγοι ή κανείς (Institute for Propaganda Analysis, 1938, p. 7). Η συγκεκριμένη τεχνική βασίζεται σε ένα είδος αυθόρμητης κρίσης που έχει αναπτύξει ο άνθρωπος κατά την εξέλιξη του και στο φαινόμενο του κοινωνικού κομφορμισμού⁷ και μάλιστα η ύπαρξή του αποδεικνύεται και πειραματικά (Nadeau, et al., 1993).

Συνοψίζοντας, τα τεχνάσματα που θα αναλυθούν και θα εντοπιστούν είναι τα εξής:

- 1) Name calling Device (Χαρακτηρισμός)
- 2) Glittering Generalities Device (Απαστράπτουσες Γενικότητες)
- 3) Transfer Device (Μεταφορά)
- 4) Testimonial/Appeal to Authority Device (Μαρτυρία/Επίκληση στην Αυθεντία)
- 5) Band Wagon Device (Φαινόμενο του συρμού)

Οι τεχνικές “name calling”, “transfer”, “testimonial” και “bandwagon” αποτελούν αυτό που ο Smith όρισε ως «προτάσεις» (suggestions) (Silverman, 2011, p. 95) και αυτή η παρατήρηση έχει σημασία διότι τα άλλα δύο τεχνάσματα που εξαιρέθηκαν δεν ανήκουν στην κατηγορία αυτή και ίσως αυτό να σημαίνει πως οι «προτάσεις» να είναι ανιχνεύσιμες σε επίπεδο μηνύματος κάτι έχει ενδιαφέρον και για τη συγκεκριμένη μελέτη και για μελλοντικές παρόμοιες μελέτες.

Πρέπει να αναφερθεί σε αυτό το σημείο πως τα τεχνάσματα αυτά όταν βρίσκονται σε μορφή κειμένου έχει συμπεραθεί πως χρησιμοποιούνται εντέχνως γενικές αρχές της σημασιολογίας του λόγου (Fleming, 1995, p. 12). Οι αρχές αυτές που ορίστηκαν από τον Wendell Johnson ως οι έξι

⁷ <https://www.youtube.com/watch?v=o8BkzvP19v4>

βασικές αρχές της γενικής σημασιολογίας (Fleming, 1995, p. 4) είναι οι εξής:

- ❖ **Μη ταυτοποίηση:** Η λέξη που με την οποία χαρακτηρίζεται κάτι δεν το ταυτοποιεί αναγκαστικά και απόλυτα. Ο χαρακτηρισμός δεν είναι ένα και το αυτό με το χαρακτηριζόμενο. Ο προπαγανδιστής όμως ελπίζει το κοινό να δράσει και να σκεφθεί σα να πιστεύει πως αυτό ισχύει.
- ❖ **Μη καθολικότητα:** Ένας δεν μπορεί να ξέρει ή να λέει όλες τις λεπτομέρειες σχετικά με κάτι. Και αυτή η αρχή όπως και οι υπόλοιπες μπορεί να συνεπικουρήσει σε μία προσπάθεια αντίληψης του παραλογισμού και της γενίκευσης ενός ισχυρισμού.
- ❖ **Αυτο – αντανακλαστικότητα:** Οι άνθρωποι δημιουργούν νοητικά αόριστες σκέψεις από αόριστες λέξεις και εκφράσεις (από λέξεις και εκφράσεις με αόριστο νόημα). Οι έννοιες με αόριστη σημασία χρησιμοποιούνται από τον προπαγανδιστή που εκμεταλλεύεται ακριβώς αυτό το φαινόμενο.
- ❖ **Η πιθανοκρατική αρχή:** Μία λέξη πιθανόν να χαρακτηρίζει κάποιον ή κάτι αλλά και αυτό υπό συνθήκες και ενδέχεται να αλλάξει. Ο προπαγανδιστής θέλει το κοινό – στόχος του να πιστεύει πως ο λόγος του εκφράζει μία αναλλοίωτη αλήθεια που δεν θα αλλάξει καθώς αυτή η έλλογη αμφισβήτηση δεν ωφελεί την προπαγανδιστική του προσπάθεια.
- ❖ **Αντιδράσεις σε σύμβολο – σύνθημα:** Οι αντιδράσεις όταν περιγράφεται μία άποψη που συμβολίζει κάτι έπονται μίας διαδικασίας σκέψης. Ο προπαγανδιστής επιθυμεί πολύ περισσότερο την αυθόρμητη και αυτόματη αντίδραση του κοινού –

στόχου χωρίς την κριτική ανάλυση της άποψής του, όσο αυτό είναι δυνατό.

- ❖ **Επέκταση (ανάλυσης του όρου):** Εδώ περιγράφεται η επιστημονική μέθοδος : Υπόθεση, παρατήρηση, αξιολόγηση και επανεκτίμηση. Αυτή η αποδεικτική διαδικασία για να υποστηριχθεί μία άποψη δε συμβαίνει πάντα στο λόγο και εξαιτίας της δυσκολίας να συμβαίνει πάντα ο προπαγανδιστής το εκμεταλλεύεται ελπίζοντας πως το κάθε άτομο στο οποίο φθάνει η προπαγάνδα του δεν θα προβεί σε τέτοια αξιολόγηση.

(Fleming, 1995)

Η παράθεση τους είναι επικουρική στην ποιοτική ανάλυση για την εξακρίβωση της ύπαρξης ενός τεχνάσματος εκ των πέντε σε ένα κομμάτι λόγου.

Η ανάλυση κειμένου είναι η τεχνική με την οποία θα αναλυθούν τα τεχνάσματα και λαμβάνει χώρα σε επίπεδο μηνύματος, όσον αφορά τα στάδια επικοινωνίας (Gulseth, 2004, p. 14). Αυτό ισχύει και για την αυτοματοποιημένη ανάλυση που θα γίνει κατά την εφαρμογή του αλγορίθμου, που θα ταξινομεί ένα κείμενο ανάλογα με την ύπαρξη ή μη των πέντε τεχνασμάτων προπαγάνδας σε αυτό, στα δεδομένα.

5. Παρόμοια εγχειρήματα

Πριν προχωρήσω στη διαδικασία της ποιοτικής ανάλυσης και των ευρημάτων και των συμπερασμάτων που προέκυψαν από αυτή, θα παραθέσω μία σειρά παρόμοιων εγχειρημάτων που περιγράφονται σε επιστημονικά συγγράμματα και άρθρα. Η καθοδήγηση από τα συμπεράσματα και τη μεθοδολογία των παρόμοιων εγχειρημάτων που εξετάστηκαν ήταν μία σημαντική συμβολή στον εντοπισμό των κατάλληλων μεθόδων και αλγορίθμων για την επεξεργασία των δεδομένων που προέκυψαν από την ποιοτική ανάλυση.

Μία παρόμοια επιστημονική προσπάθεια που διεξήχθη στη Ολλανδία πλησιάζει στη μεθοδολογία την παρούσα έρευνα. Πρόκειται για μία προσπάθεια αυτόματης ανάλυσης κοινοποιήσεων στο “Twitter” προκειμένου να βρεθεί σε ποια χρησιμοποιείται «λαϊκιστικός λόγος» (van Brekel, 2016). Η μέθοδος έρευνας ξεκινά με ένα θεωρητικό πλαίσιο σχετικά με τον λαϊκισμό και τα χαρακτηριστικά του, συμπεριλαμβανομένου ενός λεξιλογίου που σχετίζεται με αυτόν. Συνεχίζεται με την ανάθεση της προσημείωσης των κοινοποιήσεων από εθελοντές που τις βαθμολογούν σε μία πενταβάθμια κλίμακα, όπου «1» σημαίνει απουσία λαϊκιστικής γλώσσας και «5» πολύ έντονη παρουσία του χαρακτηριστικού. Προσημειώνονται 280 “tweets” και κατόπιν το σύνολο των 280 χωρίζεται σε δύο υποσύνολα, ένα με 200 κοινοποιήσεις που συντελεί στην εκμάθηση του αλγορίθμου και ένα υποσύνολο με 80 όπου θα δοκιμαστεί η αποτελεσματικότητά του. Ίδια μέθοδος στο διαχωρισμό των δεδομένων και στην εξακρίβωση της αποτελεσματικότητάς του εφαρμόζεται και στην προκείμενη έρευνα. Ο Van Brekel δοκιμάζει κατόπιν δύο μοντέλα ανάλυσης, το «K» του Fleiss και

την εφαρμογή μία λεξιλογικής ανάλυσης που ονομάζεται LIWC και αποτελεί μία ανάλυση με λεξικό στην ολλανδική γλώσσα το οποίο περιέχει 68 κατηγορίες λέξεων. Το πρώτο μοντέλο δεν ήταν αποτελεσματικό σε αντίθεση με την ανάλυση LIWC που είχε μία σχετική επιτυχία που προσέγγιζε το 70% ακρίβειας χωρίς όμως αυτό να σημαίνει εγγυημένη αποτελεσματικότητα του μοντέλου. Ένα από τα συμπεράσματά της ήταν ότι μπορεί να εντοπιστεί η λαϊκίστικη γλώσσα αυτόματα μέσω ενός εργαλείου και μεθόδου αλλά βέβαια έως έναν βαθμό. Είναι άξιο να υπογραμμιστεί επίσης πως στα συμπεράσματα αναφέρεται ότι όλη η έρευνα βασίστηκε στον θεωρητικό ορισμό του λαϊκισμού, ορίζοντας ως καταλυτικό το ρόλο του θεωρητικού πλαισίου για παρόμοιες έρευνες. Ο σημαντικότερος λόγος που επεξηγήθηκε ενδελεχώς η παραπάνω ερευνητική προσπάθεια είναι ότι ο σκοπός της ομοιάζει σε μεγάλο βαθμό με τον σκοπό της δικής μου.

Μία ακόμη ανάλυση ποσοτικών χαρακτηριστικών κειμένου παρουσιάζεται σε ένα επιστημονικό άρθρο όπου γίνεται προσπάθεια αυτόματης ανάλυσης και εντοπισμού εξτρεμιστικού λόγου. Και εδώ κατηγοριοποιούνται βάσει ενός θεωρητικού πλαισίου είδη επιχειρημάτων στο λόγο τα οποία χαρακτηρίζονται ως εκφράσεις εξτρεμιστικού λόγου και μεταγενέστερα λαμβάνει χώρα η ανάλυση. Η ανάλυση γίνεται βάσει συχνότητας όρων αλλά μέσω λογισμικού και όχι γλώσσας προγραμματισμού. Μεταξύ άλλων, δίνονται κατευθυντήριες γραμμές προκειμένου να χρησιμοποιηθεί μοντέλο ανάλυσης που δεν λαμβάνει υπόψη μόνο συχνότητα μεμονωμένων λέξεων αλλά και συχνότητα ομάδων λέξεων (N-gram modeling, Stolcke, et al. 2000 όπως αναφέρεται σε (Prentice, et al., 2011)). Είναι σημαντικό διότι στην έρευνα χρησιμοποιήθηκε εργαλείο που πραγματοποιεί κατάτμηση του κειμένου

σε όρους και μεταξύ άλλων μπορεί να το ομαδοποιεί και τις λέξεις κατά την κατάτμηση σε ομάδες των δύο, τριών, τεσσάρων και περισσότερων λέξεων. Στην έρευνα του Prentice αναφέρεται επίσης πως ένα μικτό μοντέλο ανάλυσης παρουσιάζει προτερήματα και είναι αποτελεσματικό να χρησιμοποιούνται συμπληρωματικά μοντέλα ακόμη και αν υπάρχει ένα κύριο (Prentice, et al., 2011, p. 71).

Σε ένα άλλο σύγγραμμα για την φυσική γλωσσική επεξεργασία διαδικτυακής προπαγάνδας σημειώνεται πως αποτελεί δημοφιλή τεχνική η χρήση του TF – IDF vectorizer (scikit-learn developers, 2007 - 2018) για την κατάτμηση του κειμένου σε λέξεις ή ομάδες λέξεων και κατόπιν η εφαρμογή του αλγορίθμου (Holm, 2017). Σε αυτήν την ερευνητική προσπάθεια εφαρμόζεται το στατιστικό μοντέλο “K-means” καθώς και το μοντέλο LDA. Τα μοντέλα εφαρμόζονται με υποσύνολα εκμάθησης και δοκιμής όπως ακριβώς και στην περίπτωση του Van Brekel. Επιπρόσθετα, η επεξεργασία των δεδομένων γίνεται εδώ αφού παρατίθενται ως ταξινομημένα και προσημειωμένα κείμενα, τα οποία προσημειώνονται κατόπιν ποιοτικής ανάλυσης. Στο ίδιο σύγγραμμα οι συγγραφείς αποφαίνονται πως η ιδεολογία είναι ανιχνεύσιμη στο κείμενο (Holm, 2017, p. 56), γεγονός που μας προϋδεάζει θετικά για την παρούσα προσπάθεια καθώς η ιδεολογία διαδίδεται μέσω της προπαγάνδας, και αυτό είναι δυνατό να συμβεί σε όλες τις μορφές λόγου. Τέλος, προτείνεται η βιβλιοθήκη «sk-learn» (scikit-learn developers, 2007 - 2018) στην Python και μερικά μοντέλα της για μελλοντικές παρόμοιες εργασίες καθώς και τεχνικές «βαθιάς μάθησης» όπως η χρήση νευρικών δικτύων την οποία δεν μπόρεσα να εφαρμόσω λόγω έλλειψης τεχνικής γνώσης πάνω σε αυτήν την τεχνική.

Σε επιστημονικό άρθρο που μελετά τον αυτόματο εντοπισμό ψευδών ειδήσεων χρησιμοποιήθηκαν οι αλγόριθμοι «Support Vector Machines (SVM)» και μοντέλα τύπου «Naïve Bayes» για την ταξινόμηση των κειμένων μαζί με την προσέγγιση “bag of words” καθώς και με την κατάτμηση σε ομάδες λέξεων, δηλαδή “n-grams”. Ορίζονται μάλιστα από το άρθρο ως οι ενδεικνυόμενα αλγοριθμικά μοντέλα για γλωσσολογικού τύπου αναλύσεις (Conroy, et al., 2015), ιδιαίτερα όταν υποστηρίζονται ή και συμπληρώνονται από άλλου είδους μεθόδους.

Σε άλλο project χρησιμοποιήθηκε ο αλγόριθμος SVM (Support Vector Machines) για τον εντοπισμό σατυρικών στοιχείων στο κείμενο που θα μπορούσαν να αναδείξουν εν δυνάμει παραπλανητικές ειδήσεις (Rubin, et al., 2016.). Γίνεται επίσης κανονικοποίηση της συχνότητας των όρων, η οποία ενδείκνυται για την καλύτερη δυνατή αποτελεσματικότητα του αλγορίθμου και της προβλεπτικής του ικανότητας. Σε αυτό αναφέρεται μία ακόμη έρευνα για την ανίχνευση στοιχείων εξαπάτησης στις ειδήσεις στη ρωσική γλώσσα, όπου χρησιμοποιήθηκαν και συγκρίθηκαν οι αλγόριθμοι επιβλεπόμενης μάθησης Support Vector Machines⁸ και Random Forest Classifier⁹ (Pisarevskaya, 2017).

Τέλος, μία προσπάθεια αυτόματης κατηγοριοποίησης των ειδήσεων σε σατυρικές, ψευδείς, έμπιστες ή προπαγανδιστικές έχει πραγματοποιηθεί και με τη σύγκριση αλγοριθμικών μοντέλων (Rashkin, et al., 2017). Σε αυτήν την έρευνα ενώ δόθηκε εξαιρετική σημασία στους αλγορίθμους μηχανικής μάθησης, το θεωρητικό πλαίσιο δεν ήταν πλήρες, γεγονός που μπορεί να καταστήσει και μη έγκυρη την έρευνα παρά την

⁸ <https://scikit-learn.org/stable/modules/svm.html>

⁹ <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

αποτελεσματικότητα πρόβλεψης και ταξινόμησης των αλγοριθμικών μοντέλων.

Η επισκόπηση των παρόμοιων εγχειρημάτων βοήθησε αποφασιστικά στην επιλογή και των εργαλείων ανάλυσης και της μεθοδολογίας που διαμορφώθηκε οριστικά κατόπιν αυτής της επισκόπησης.

6. Ποιοτική ανάλυση

Το θεωρητικό πλαίσιο αποτελεί τη βάση της ποιοτικής ανάλυσης προπαγάνδας, από την οποία τα αποτελέσματα και τα συμπεράσματα θα σημειωθούν και θα σχολιασθούν για την μετάβαση στο επόμενο βήμα που είναι η αυτοματοποίηση της ανάλυσης.

6.1. Αρχικές διευκρινίσεις για την ανάλυση των τεχνικών

Η ανάλυση θα γίνει στην αγγλική γλώσσα και θα γίνει σε επίπεδο προτάσεων ή παραγράφων με βάση το θεωρητικό πλαίσιο που παρατέθηκε. Ο λόγος που επελέγη η αγγλική γλώσσα ως γλώσσα ανάλυσης είναι το πλήθος διαθέσιμων προγραμματιστικών εργαλείων επεξεργασίας και ανάλυσης κειμένου σε σύγκριση με όλες τις άλλες. Η εύρεση προπαγάνδας θα αφορά τα πέντε τεχνάσματα. Επίσης όταν εντοπίζεται ένα ή περισσότερα τεχνάσματα σε ένα κομμάτι κειμένου αυτό θα κατατάσσεται στα κείμενα που εμπεριέχουν προπαγάνδα. Αν δεν εντοπίζεται κάποιο από αυτά τα πέντε τεχνάσματα θα κατατάσσεται σε αυτά που δεν εμπεριέχουν προπαγάνδα. Είναι άξιο αναφοράς πως η ταξινόμηση των κειμένων σε εμπεριέχοντα και μη εμπεριέχοντα προπαγάνδα θα γίνει με κριτήριο μόνο αυτά τα πέντε τεχνάσματα. Ακόμη και αν εντοπίζεται κάποιο άλλο τέχνασμα, στην παρούσα φάση της έρευνας, δε θα λαμβάνεται υπόψη στην ταξινόμηση. Παρ' όλα αυτά κατά την ποιοτική ανάλυση των προτάσεων και παραγράφων υπήρξε η μέριμνα από μέρους μου σε όσα κομμάτια κειμένου δεν εντοπίζεται ούτε ένα από τα πέντε τεχνάσματα να μην εντοπίζεται και γενικότερα κάποιο άλλο για προληπτικούς λόγους εγκυρότητας της διαδικασίας ανάλυσης. Ακόμη όμως και αν παρατηρείται κάποιο άλλο εκτός των πέντε, αυτό δεν αφορά την παρούσα έρευνα.

6.2. Δεδομένα

Τα δεδομένα και ο τρόπος που θα καταγραφούν και θα εισαχθούν στο προγραμματιστικό περιβάλλον γνωρίζουμε εξ αρχής πως έχουν κομβικό ρόλο στο όλο εγχείρημα. Για αυτό το λόγο η επιλογή των δεδομένων ήταν τυχαία καθ' όλη τη διάρκεια της προσπάθειας εύρεσης και ανάλυσης τους. Αποτελούν φράσεις, προτάσεις ή παραγράφους καταγεγραμμένες από ποικίλα μέσα μαζικής ενημέρωσης, που διαθέτουν διεθνείς ιστοσελίδες στην αγγλική γλώσσα, από άρθρα ιστοσελίδων πολιτικών κομμάτων καθώς και από βιβλία. Από τα Μέσα της Αμερικής δόθηκε ιδιαίτερη προσοχή να αναλυθούν κομμάτια λόγου που ήταν και υπέρ και εναντίον του Τράμπ και εναντίον και υπέρ της Χίλαρι Κλίντον. Επιπλέον, υπήρξε η επιμέλεια να παρθούν κομμάτια κειμένου από κάθε είδος Μέσου, όσον αφορά την ιδεολογική κατεύθυνση και την κομματική ή κρατική¹⁰ στράτευση. Όλα τα κείμενα από τα οποία πάρθηκαν τα κομμάτια του κειμένου που αναλύθηκαν ως προς τα πέντε τεχνάσματα προπαγάνδας που έχουν ορισθεί είναι από το 2000 έως και σήμερα. Η μέθοδος ήταν δηλαδή μία τυχαία επιλογή με προσοχή από τη μεριά μου να υπάρχει στα δεδομένα ένας πλουραλισμός ως την ιδεολογική τοποθέτηση και τη θεματολογία, προκειμένου να μην αναγνωριστούν ως σημαντικοί παράγοντες ταξινόμησης από τον αλγόριθμο κειμενικά στοιχεία που σχετίζονται με το θέμα και με την ιδεολογία που διέπει τον πολιτικό λόγο αλλά με τα λεκτικά σχήματα και τις λέξεις που δηλώνουν την ύπαρξη μίας, κάποιων ή και όλων των πέντε ορισθέντων προπαγανδιστικών τεχνικών ή την απουσία αυτών.

¹⁰ Όταν προπαγανδίζει ως επί το πλείστον μία κυρίαρχη ιδεολογία στο κράτος που δραστηριοποιείται

Η προέλευση των δεδομένων ήταν τυχαία διότι, με εξαίρεση τα προαναφερθέντα χαρακτηριστικά τους που έπρεπε να προσεχθούν, η ανάλυση και ο εντοπισμός των τεχνασμάτων προπαγάνδας στο εκάστοτε κομμάτι κειμένου προς επεξεργασία βάσει λεκτικών σημείων έχει την κυριότερη σημασία για την μεταγενέστερη αλγοριθμική αυτόματη ταξινόμηση που θα επιχειρηθεί και θα βασιστεί σε αυτήν την ανάλυση. Το τελικό σύνολο των δεδομένων αποτελείται από 320 καταχωρήσεις, κομματιών από κείμενα και τις επισημειώσεις τους. Στο 58% των καταχωρήσεων εντοπίζεται προπαγάνδα ενώ στο 42% δεν εντοπίζεται (Γράφημα 1). Τα δεδομένα είναι σχετικά ισοσκελισμένα, γεγονός που υποβοηθά τον αλγόριθμο στην διαδικασία εκμάθησης για την ορθή ταξινόμηση νέων δεδομένων καθώς και η απουσία και η παρουσία του χαρακτηριστικού ως προς το οποίο ταξινομείται μία μεταβλητή είναι εξίσου σημαντικά ώστε να εντοπιστούν ακριβέστερα και αποτελεσματικότερα τα στοιχεία που λαμβάνει υπόψη ο αλγόριθμος για την ταξινόμηση.



Γράφημα 1 - Παρουσία ή απουσία προπαγάνδας επί του συνόλου των δεδομένων (Ποσοστό)

6.3. Ταξινόμηση

Η ταξινόμηση πραγματοποιήθηκε με βάση το θεωρητικό πλαίσιο για την προπαγάνδα γενικότερα και ειδικότερα με βάση τη θεωρία που ορίζει τα χαρακτηριστικά των πέντε τεχνασμάτων προπαγάνδας, Name calling Device (Χαρακτηρισμός), Glittering Generalities Device (Απαστράπτουσες Γενικότητες), Transfer Device (Μεταφορά), Testimonial/Appeal to Authority Device (Μαρτυρία/Επίκληση στην Αυθεντία), Band Wagon Device (Φαινόμενο του συρμού). Τα κείμενα εισάγονταν στην πρώτη στήλη ενός υπολογιστικού φύλλου και στη συνέχεια υπήρχε μία στήλη για κάθε ένα από τα τεχνάσματα. Στο αντίστοιχο ευθυγραμμισμένο με το κείμενο κελί των στηλών των τεχνασμάτων εισαγόταν η ένδειξη “Yes” αν το τέχνασμα ήταν ανιχνεύσιμο με λεκτικά σημάδια στο κείμενο και η ένδειξη “No” αν δεν ανιχνευόταν η τεχνική. Από το σύνολο των ενδείξεων στις πέντε στήλες προέκυπτε μία ακόμη ένδειξη στην επόμενη στήλη, η οποία έλαβε την ονομασία «How many “No”», που ήταν ένας αριθμός από το 0 έως το 5. Η ένδειξη «5» σημαίνει απουσία και των πέντε προπαγανδιστικών τεχνασμάτων το οποίο σημαίνει ότι στο συγκεκριμένο κομμάτι κειμένου δεν υπάρχει προπαγάνδα. Οι ενδείξεις από «0» έως και «4» σημαίνει σημαίνουν πως υπάρχει ένα ή περισσότερα από τα πέντε τεχνάσματα και άρα υπάρχει προπαγάνδα. Στο σημείο αυτό είναι απαραίτητο να αναφερθεί πως υπήρχε κατά τη διάρκεια της ανάλυσης η πρόθεση από μέρους μου το κείμενο που χαρακτηριζόταν με την ένδειξη «5» να μην περιέχει όχι μόνο αυτά τα πέντε τεχνάσματα προπαγάνδας αλλά και κανένα άλλο. Διαφορετικά δεν εισαγόταν στα δεδομένα. Αυτή η ιδιαίτερη μέριμνα υπήρξε με σκοπό την αποφυγή οποιασδήποτε σύγχυσης, στη συνέχεια της ανάλυσης, στην αξιολόγηση της αποτελεσματικότητας του

αλγοριθμικού μοντέλου μηχανικής μάθησης καθώς και στα τελικά συμπεράσματα.

Η ταξινόμηση των δεδομένων έγινε με αυτή τη μέθοδο διότι, αφού λήφθηκε υπόψη η μεθοδολογία των προηγούμενων εγχειρημάτων, κρίθηκε ως κατάλληλη για τη ευκολότερη επεξεργασία των δεδομένων στο προγραμματιστικό περιβάλλον Jupiter με την Python και για την «εκπαίδευση» του αλγορίθμου. Τα δεδομένα από αυτού του είδους την ταξινόμηση των κειμένων θα χρησιμοποιηθούν για την αυτόματη ταξινόμηση μετά από την εφαρμογή του αλγορίθμου. Η διαδικασία αυτή ονομάζεται «επιβλεπόμενη μάθηση», η οποία ενδείκνυται όταν οι κατηγορίες προϋπάρχουν και δεν δημιουργούνται αυτόματα από τον αλγόριθμο (González-Bailón & Petchler, 2015) όπως εν προκειμένω.

Η μορφή των ταξινομημένων δεδομένων όπως αυτά επρόκειτο να καταχωρηθούν προγραμματιστικά απεικονίζεται στον πίνακα 3.

Phrases	Bandwagon Effect	Glittering Generalities	Name Calling	Testimonial/Appeal to Authority	Transfer	How many "No"
Convincing Cramer -- largely viewed as the best candidate they had to beat Heitkamp -- was an important part of a GOP effort to grow their slim Senate majority.	Yes	No	No	No	No	4

Πίνακας 3 – Παράδειγμα καταγραφής των αποτελεσμάτων της ποιοτικής ανάλυσης

6.4. Παρατηρήσεις και συμπεράσματα από την ποιοτική ανάλυση

Η ποιοτική ανάλυση για την εξεύρεση τεχνασμάτων πραγματοποιήθηκε στηριζόμενη στο θεωρητικό πλαίσιο για τα προς ανάλυση τεχνάσματα (βλ. σελ. 13-19). Κομμάτια κειμένου από άρθρα ή γενικότερα πολιτικής θεματολογίας κείμενα εισάγονταν στην πρώτη στήλη και ακολουθούσε η ανάλυση για το καθένα. Για να χαρακτηριστεί ένα κείμενο ως εμπειρέχον ένα συγκεκριμένο προπαγανδιστικό τέχνασμα έπρεπε σε κάποια λέξη ή ομάδα λέξεων να παρατηρούνταν χαρακτηριστικά αυτού του τεχνάσματος. Αυτό ελήφθη υπόψη κατά την ανάλυση καθώς ο αλγόριθμος κάνει τρόπον τινά μία περίπλοκη μέτρηση συχνότητας λέξεων και από αυτή πραγματοποιεί την ταξινόμηση. Επιπλέον από τις λέξεις που παρατηρούσα σε κάθε κομμάτι κειμένου που το επισημαίωνα ως εμπειρέχον ένα ή περισσότερα τεχνάσματα δημιούργησα μικρής έκτασης λεξικά με λέξεις πως συνέβαλαν στο να με οδηγήσουν στο συμπέρασμα ότι υπάρχει προπαγανδιστικό τέχνασμα στο συγκεκριμένο κομμάτι προπαγάνδας.

Αυτό όχι μόνο θα συνεπικουρήσει στην καλύτερη αποτύπωση της αποτελεσματικότητας του αλγορίθμου αλλά επίσης είναι μία αρχή για τη δημιουργία πληρέστερων λεξικών βάσει τα οποία μεταγενέστερα θα μπορούσαν να χρησιμοποιηθούν ως αυτόνομοι τρόποι ανίχνευσης των τεχνασμάτων αυτών.

Τα παρακάτω λεξικά είναι ημιτελή, με την έννοια ότι δεν καλύπτουν όλο ούτε το μεγαλύτερο κομμάτι των λέξεων που σε ένα πολιτικό άρθρο θα μπορούσαν να αποτελέσουν έναν δείκτη πως υπάρχει το αντίστοιχο τέχνασμα προπαγάνδας. Είναι όμως κειμενικά σημάδια τα οποία κατά την ποιοτική ανάλυση οδήγησαν μαζί με άλλους παράγοντες στην

επισημείωση “Yes”, δηλαδή στο συμπέρασμα ότι υπάρχει αυτού του είδους η προπαγάνδα σε αυτή τη φράση.

Name calling lexicon example – Παράδειγμα λεξικού για τον «χαρακτηρισμό»:

Fascist demagogue, dictator, red, oligarchy, communist, alien, agitator, royalist, utopian, rabble-rouser, troublemaker, Tory, bigot, national security, threat, illegal, alleged, surreal, hard-line, unacceptable, Trumpy-ish, militaristic, atrocious, horrible, pushers, hell-raisers, drunkards, who raise hell, foreign, inflammatory, witch hunt, rigged, fake news, isolationist, outrageous, controversial, violent, leftist, leftists, uninformed, mob, bizarre, globalist, stunt, disastrous, tribal, typical, populist, mythical, Putin’s Chef, white working-class, Trump-like, Rust Belt, blue-collar, media reported assassins, rogue, relationship-altering, enemy, cover-up, absurd, dangerous, Blue Dog, Terrorist, terrorists, Evil, The worst, The death of America, Armageddon, Hoax, xenophobic...

Testimonial lexicon example – Παράδειγμα λεξικού για τη «μαρτυρία»:

Polls, analysts, experts, environmental groups, based on, analysts, experts, sources...

Bandwagon lexicon example – Παράδειγμα λεξικού για το «φαινόμενο του συρμού»:

plenty of, largely viewed, that earned nearly, outrage, people, is believed to, said to be, polls, 53% of those, international outcry, 50%<, outcry, most people, majority, more than, widely seen, percent, for millions...

Glittering Generalities lexicon example – Παράδειγμα λεξικού για τις
«απαστράπτουσες γενικότητες»:

civilization, Christianity, good, proper, right, democracy, patriotism, motherhood, fatherhood, scientific method, love, peace, perfect, different, differently, safer, hopeful, prosperity, reform, progress, future, freedom, amazing, tremendous, strong, phenomenal, fabulous, best, beautiful, truth, freedom, honor, liberty, social justice, public service, loyalty, the American way, generosity, brotherhood, glory, family values, honesty, patriotic, transparency, the right thing, the American dream, all options, fantastic thing, great, the biggest change, funny, tremendous step, feel much differently , beautiful, law, strength, conspiracy, division, not right, economic miracles, durable politics, take our country back, America works, Security, power, special, inability, innocence, order, fairness, justice, discord, great, strongest, socialism, mostly true, progressive, no small thing, vision, visions, everyday Americans...

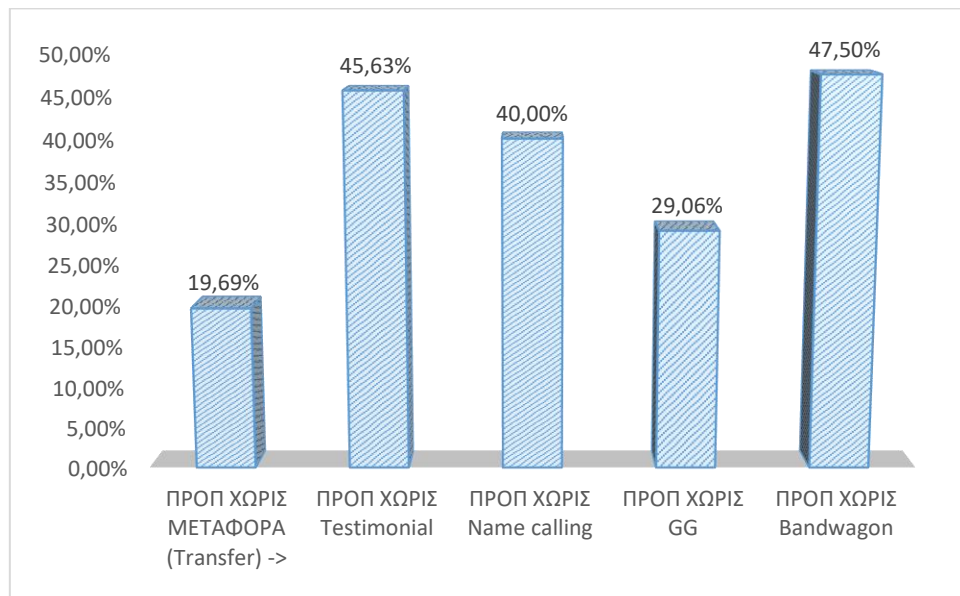
Τα σύνολα λέξεων αυτά εξήχθησαν από παρατηρήσεις στα προς ανάλυση κείμενα και υπάρχουν σε ορισμένα από αυτά. Τα λεξικά είναι ατελή και γι' αυτό και δεν μπορούν σε αυτό το στάδιο να χρησιμοποιηθούν σαν μία αυτοτελής μέθοδος ανίχνευσης των αντίστοιχων προπαγανδιστικών τεχνικών.

Άλλη παρατήρηση κατά την ποιοτική αξιολόγηση που είναι αξιοσημείωτη είναι η συχνότητα που παρατηρείται το τέχνασμα της «μεταφοράς» (Transfer) καθώς και η συχνή συνύπαρξή του με τον χαρακτηρισμό και τις απαστράπτουσες γενικότητες.

	Transfer - Bandwagon	Transfer - Glittering Generalities	Transfer - Name Calling	Transfer - Testimonial
"Yes" και στα δύο	17	60	45	25
Transfer - Συνολικός αριθμός	123			
Ποσοστά επί του συνόλου	14%	49%	37%	20%

Πίνακας 4 – Συσχέτιση του τεχνάσματος της «μεταφοράς» με τα υπόλοιπα

Πιο συγκεκριμένα, η μεταφορά απουσιάζει ως τέχνασμα μόνο στο 19,69% των κειμένων εντοπίστηκε προπαγανδιστικό τέχνασμα χωρίς να εντοπίζεται μεταφορά (Γράφημα 2).



Γράφημα 2 - Ποσοστό απουσίας κάθε τεχνάσματος επί του συνόλου των δεδομένων

Αυτό το στοιχείο μας δείχνει πως, τουλάχιστον όπως συνάγεται από το δείγμα των 320 αναλυθέντων κομματιών κειμένου, η «μεταφορά» είναι η συχνότερη τεχνική που συναντάται. Επίσης, παρατηρήθηκε πως συσχετίζεται περισσότερο με τις τεχνικές «απαστράπτουσες γενικότητες» και τον «χαρακτηρισμό», αφού συνυπάρχει σε ποσοστό 49% με την πρώτη

και 37% με την δεύτερη επί του συνόλου των κομματιών κειμένου που συμπεράθηκε ότι περιείχαν την τεχνική της μεταφοράς (Πίνακας 4 – Συσχέτιση του τεχνάσματος της «μεταφοράς» με τα υπόλοιπα Πίνακας 4).

Άλλη σημαντική παρατήρηση υπήρξε το γεγονός που στα 2/3 των κειμένων, που βρέθηκαν να περιέχουν προπαγάνδα κάποιας μορφής, ανιχνεύτηκαν περισσότερα από ένα τεχνάσματα. Αυτό είναι μία ένδειξη ότι όταν υπάρχει προπαγάνδα σε μία μορφή στο κείμενο πιθανόν αυξάνεται η πιθανότητα να υπάρχει και σε άλλες. Οδηγήθηκα σε αυτό το συμπέρασμα λαμβάνοντας υπόψη επίσης πως τα χωρία κειμένου προς ανάλυση ήταν σχετικά μικρά προς την ολότητα του κειμένου από το οποίο εξήχθησαν. Το σκεπτικό είναι ότι πιθανότατα ένα άλλο τέχνασμα προπαγάνδας ή και το ίδιο να μπορούσε να εντοπιστεί σε ένα άλλο σημείο του κειμένου.

Κάθε τεχνική προπαγάνδας που ορίστηκε ήταν ως επί το πλείστον ανιχνεύσιμη από μία λέξη ή συνδυασμό λέξεων που βρίσκονταν στο κείμενο ή τουλάχιστον αυτά τα λεκτικά σημεία αποτελούσαν μία ένδειξη ότι υπάρχει. Η παρατήρηση αυτή, αν και παρατίθεται με επιφύλαξη γιατί είναι υποκειμενικό συμπέρασμα και μένει να επιβεβαιωθεί επαναλαμβανόμενα, είναι ευοίωνη για της έκβαση της παρούσας προσπάθειας. Ένας από τους σκοπούς της, άλλωστε, είναι και να συμπεραθεί αν υπάρχει η δυνατότητα της αυτοματοποίησης εξεύρεσης κάποιου είδους προπαγάνδας ή αν αυτό φαίνεται να μην είναι ακόμη δυνατό.

7. Έρευνα

7.1. Ερευνητικά ερωτήματα

Κύριο ερευνητικό ερώτημα της επιστημονικής προσπάθειας μας είναι αυτό που αναφέρθηκε και στην εισαγωγή: Υπάρχει η δυνατότητα να αυτοματοποιηθεί μία διαδικασία ανάλυσης προπαγάνδας σε επίπεδο κειμένου και να χρησιμοποιηθεί σε δεδομένα μεγάλου όγκου;

Το ερώτημα γεννά από μόνο του και άλλα υπό – ερωτήματα:

Μπορεί να δημιουργηθεί ειδικό λογισμικό που θα διαχωρίζει επιτυχώς κείμενα σε όσα χαρακτηρίζονται και σε όσα δε χαρακτηρίζονται ως προπαγανδιστικά; Μπορεί να γίνει μία λεπτομερέστερη ταξινόμηση αυτών των κειμένων και ανάλογα με το τέχνασμα προπαγάνδας που παρατηρείται σε αυτά; Με τι ποσοστό επιτυχίας μπορούν να συμβούν όλα αυτά και με ποιες μεθόδους; Ποιες μέθοδοι είναι προτιμότερο να δοκιμαστούν σε παρόμοιες ερευνητικές προσπάθειες στο μέλλον;

7.2. Ερευνητικές υποθέσεις

Το εγχείρημα είναι πρωτότυπο και εν προκειμένω απόλυτες απαντήσεις δεν αναμένεται να δοθούν, αλλά μπορούν να γίνουν παρατηρήσεις και να δοθούν ενδείξεις για το αν και πως είναι δυνατόν να επιτευχθεί ο σκοπός του παρόντος εγχειρήματος. Αναμένεται να εξαχθούν επίσης συμπεράσματα για το ποιος είναι ο βέλτιστος τρόπος αυτόματης επεξεργασίας των δεδομένων με κριτήριο την ακρίβεια και την εγκυρότητα των αποτελεσμάτων. Η διαρκής και σταδιακή προσέγγιση του σκοπού είναι δεδομένη καθώς ακόμη το εγχείρημα βρίσκεται σε πειραματική διαδικασία.

Μία εκ των υποθέσεων που μπορούν να γίνουν είναι ότι θα χρειαστούν μικτές μέθοδοι ανάλυσης των δεδομένων για την επιτυχήστερη ταξινόμησή τους. Η επιβλεπόμενη μάθηση πιθανόν να μην αρκεί λόγω της πολυπλοκότητας του εγχειρήματος.

Επιπρόσθετα, τα προς ανάλυση τεχνάσματα σχετίζονται με το συναίσθημα και η τεχνική της μεταφοράς, η οποία εντοπίζεται συχνότερα απ' όλες τις υπόλοιπες, σχετίζεται αποκλειστικά με αυτό. Επομένως συμπεραίνουμε πως στην έρευνα θα πρέπει να εμπλέκονται και τρόποι αξιολόγησης των συναισθημάτων που διέπουν το εκάστοτε κείμενο προς ανάλυση, ώστε τα αποτελέσματα αυτής της αξιολόγησης να λαμβάνονται υπόψη στη διαδικασία της ταξινόμησης ως ένας από τους παράγοντες που οδηγούν σε αυτή. Υπάρχουν λεξικά τα οποία μπορούν να χρησιμοποιηθούν ως εργαλείο ώστε να γίνει αυτόματα αυτή η διεργασία. Αυτό θα επιχειρηθεί σε μία από τις δοκιμές.

7.3. Κριτήρια αποτελεσματικότητας

Οι παράγοντες που θα κρίνουν το βαθμό επιτυχίας της παρούσας ερευνητικής προσπάθειας είναι οι εξής:

- 1) Το ποσοστό ακρίβειας πρόβλεψης του αλγορίθμου. Η προβλεπτική ικανότητα είναι αναγκαίο κριτήριο κάθε επιστημονικής μεθόδου. Το ποσοστό ακρίβειας θα παρατίθεται μετά ως αποτέλεσμα του κώδικα με την εξής μορφή στο κεφάλαιο των δοκιμών:

Test Accuracy ::

- 2) Τα στοιχεία και οι παράγοντες που συντελούν στην κατηγοριοποίηση των κειμένων από το μοντέλο μηχανικής μάθησης. Αυτά θα παρατίθενται με μορφή γραφήματος μετά από το ποσοστό ακρίβειας σε κάθε δοκιμή.

Ως προς το δεύτερο κριτήριο πρέπει να σημειωθεί πως τα στοιχεία αυτά δεν είναι παρά λέξεις, ομάδες λέξεων ή χαρακτηριστικά με διαφορετική όμως σημαντικότητα το καθένα για τη διαδικασία της ταξινόμησης. Αν τα στοιχεία αυτά αντιστοιχούν σε όσα αναφέρονται στο θεωρητικό πλαίσιο της ανάλυσης ως χαρακτηριστικά της προπαγάνδας ή των επιμέρους τεχνασμάτων της τότε θα συμπεραίνεται πως υπάρχει μία ακόμη ένδειξη πως το εγχείρημα μας είναι εφικτό. Αν τα στοιχεία που αναδεικνύονται ως σημαντικά για την ταξινόμηση από τον αλγόριθμο δεν σχετίζονται κατά κανένα τρόπο με την ύπαρξη ή τη μη ύπαρξη προπαγάνδας τότε το γεγονός αυτό θα αποτελεί αντένδειξη. Η αξιολόγηση των στοιχείων θα είναι αναγκαστικά μη αυτόματη, ποιοτική ή ποσοτική, και όχι σε υπολογιστικό περιβάλλον.

8. Χαρακτηριστικά της Μεθόδου

Η διαδικασία με την οποία θα επιχειρήσουμε να αυτοματοποιήσουμε την ανάλυση προπαγάνδας έχει ορισμένα χαρακτηριστικά άλλα θετικά και άλλα αρνητικά που είναι σημαντικό να παρατεθούν και προς χάρη των μελλοντικών ερευνών πάνω στο ίδιο ζήτημα.

8.1. Πλεονεκτήματα

Το κυριότερο πλεονέκτημα αυτής της μεθόδου ανάλυσης (επιβλεπόμενη μάθηση) είναι η εφαρμογή της σε δεδομένα πολύ μεγάλου όγκου με ταχύτητα μη επιτεύξιμη σε ανάλυση από άνθρωπο. Εξάλειψη του λάθους που οφείλεται σε απροσεξία ή παράλειψη

Ένα από τα σημαντικότερα θετικά στοιχεία αυτής της μεθόδου είναι ότι μπορούν ανά πάσα στιγμή και πολύ εύκολα να αφαιρούνται, αν θεωρηθεί απαραίτητο, και να προστίθενται δεδομένα ώστε να βελτιστοποιείται διαρκώς η αποτελεσματικότητα της ανάλυσης.

Επίσης για τον ίδιο σκοπό είναι δυνατή και το ίδιο εύκολη η αλλαγή παραμέτρων στον αλγόριθμο ανάλυσης των δεδομένων ή και του ίδιου του αλγορίθμου.

Επιπλέον υπάρχει ευχέρεια να χρησιμοποιούνται και άλλες μέθοδοι ανάλυσης με την ήδη εφαρμοσμένη προς βελτίωση της αποτελεσματικότητας, όπως ειδικά λεξικά και συνδυασμοί αναλυτικών διαδικασιών.

Αυτή η άμεση και εύκολη μεταβολή των παραμέτρων είναι αναγκαία σε αυτοματοποιημένες αναλύσεις της γλώσσας καθώς η γλώσσα έχει το μόνιμο χαρακτηριστικό να μεταβάλλεται και επομένως κάποια γλωσσικά

στοιχεία που αποτελούν ένδειξη ή απόδειξη ύπαρξης προπαγάνδας ενδεχομένως να αλλάξουν.

Ένα άλλο πλεονέκτημα της μεθόδου της «επιβλεπόμενης μάθησης» είναι ότι μπορεί να αναδείξει λέξεις και στοιχεία που σχετίζονται με την ύπαρξη μορφών προπαγάνδας σε ένα κείμενο τα οποία δεν είχε παρατηρήσει κάποιος αναλυτής. Κατά αυτήν την έννοια μπορεί όχι μόνο να αυτοματοποιηθεί η ανάλυση των κειμένων για την ανίχνευση προπαγάνδας αλλά και να αναδεικνύονται νέα στοιχεία που θα βοηθούν τους αναλυτές κειμένων στην ποιοτική ανάλυση.

Τέλος, ο συνδυασμός μίας ποιοτικής ανάλυσης και μίας αυτοματοποιημένης μειώνουν τις πιθανότητες να γίνουν σημαντικά λάθη στην πρώτη και η δεύτερη τη συμπληρώνει και την υποβοηθά.

8.2. Περιορισμοί και προκλήσεις

Οι προκλήσεις και τα μειονεκτήματα είναι πολλά και περίπλοκα καθώς το εγχείρημα είναι δύσκολο και οι παρόμοιες προσπάθειες που έχουν γίνει λίγες. Επίσης, η τεχνητή νοημοσύνη δεν είναι γνωστό πόσες και ποιες δυνατότητες έχει σε σχέση με εκείνες του ανθρώπου ούτε είναι γνωστό αν αυτή τεχνολογία είναι σίγουρα σε τέτοιο στάδιο ώστε να φέρει σε πέρας με επιτυχία μία περίπλοκη ανάλυση όπως αυτή των τεχνασμάτων προπαγάνδας. Η έρευνα είναι σε πειραματικό στάδιο.

Αρχικά, σε επίπεδο λεξιλογικού περιεχομένου των κειμένων υποθέτουμε πως ίσως κάποιες λέξεις που δεν σχετίζονται με την προπαγάνδα να επηρεάσουν, αν όχι την ακρίβεια, την αποτελεσματικότητα του αλγοριθμικού μοντέλου να ταξινομή τα κείμενα. Αυτό είναι μία πρόκληση που πιθανόν να ξεπεραστεί με συνεχόμενες πειραματικές εφαρμογές και κατόπιν ανάλογη τροποποίηση στα δεδομένα.

Πρέπει να αναφερθεί πως η ανάλυση δεν πραγματοποιήθηκε για κάθε τέχνασμα ξεχωριστά λόγω ανεπάρκειας ικανοποιητικού όγκου δεδομένων και ιεραρχίας των ερωτημάτων της έρευνάς μας. Αρχικά τα δεδομένα αναλύθηκαν σε ενιαία μορφή διότι δόθηκε προτεραιότητα στην απάντηση του κύριου ερευνητικού ερωτήματος.

Σημειώνεται ότι αυτές οι λέξεις που πιστεύουμε ότι θα προκαλέσουν σύγχυση στην ανάλυση δεν περιλαμβάνονται στις γενικά συχνές λέξεις που παρατηρούνται σε κάθε είδους κειμένου (“stopwords”), οι οποίες υπάρχουν σε λεξικό και είναι δυνατόν να αφαιρεθούν. Από την άλλη, η αφαίρεση αυτών των λέξεων μπορεί και αυτή να προκαλέσει προβλήματα καθώς μία από αυτές τις λέξεις σε κάποιες περιπτώσεις να αποτελεί ένδειξη ύπαρξης τεχνάσματος προπαγάνδας.

Επιπλέον, θεωρούμε πως μπορεί να υπάρξει «προπαγάνδα εκ παραδρομής». Ορίζουμε αυτό το είδος ως ακούσια προπαγάνδα.

Ενδεχομένως δηλαδή να υπάρχει μία μεροληπτική παρουσίαση των γεγονότων υπέρ ή κατά οποιουδήποτε. Με την παρούσα μέθοδο σε καμία περίπτωση δεν μπορούμε να ανιχνεύσουμε το σκοπό της προπαγάνδας και αν είναι ακούσια ή εκούσια καθώς αυτό καθορίζεται, μεταξύ άλλων, και από εξωκειμενικά στοιχεία και άλλους παράγοντες, που δεν υπάρχει νόημα να εξηγηθούν επί του παρόντος.

Με την παρούσα μέθοδο, επίσης, δεν υπάρχει δυνατότητα να γνωρίζουμε αν η προπαγάνδα διενεργείται από τον γράφοντα το κείμενο ή από αυτόν που παραθέτει ως πηγή ο γράφων. Ίσως σε αυτήν την πρόκληση υπάρχουν λύσεις αλλά δεν θα δοθούν σε αυτήν την έρευνα διότι κάτι τέτοιο απαιτεί επιπρόσθετη θεωρία και ανάλυση σε γλωσσολογικά στοιχεία εκτός του θέματος που αφορά την έρευνα.

Μία σειρά περιορισμών που πρέπει να υπογραμμιστούν είναι και αυτοί που προκύπτουν από τη φύση των δεδομένων. Τα δεδομένα εντοπίστηκαν σε κατηγορίες πολιτικών άρθρων και σε κείμενα από ιστοσελίδες πολιτικών κομμάτων. Τα δεδομένα στα οποία μπορεί να εφαρμοστεί η ανάλυση δεν μπορεί να είναι μη πολιτικά κείμενα. Επομένως υπάρχει ο περιορισμός να εισάγονται κείμενα που είναι πολιτικού χαρακτήρα ώστε να εφαρμόζεται επιτυχώς η αλγοριθμική ανάλυση και ταξινόμηση τους, εφόσον βέβαια το εγχείρημά μας έχει κάποια επιτυχία.

Ένας άλλος περιορισμός, που αφορά πάλι τη φύση των δεδομένων, σχετίζεται με την ποσότητα τους. Χρειάζονται περισσότερα πειράματα με μεγαλύτερο όγκο δεδομένων και διαφορετικής προέλευσης για να μπορούμε να υποθέσουμε με σιγουριά πως μπορούμε να δημιουργήσουμε ένα λογισμικό εντοπισμού προπαγάνδας το οποίο θα μαθαίνει από τα δεδομένα που θα αναλύει και θα τελειοποιείται.

Μία τελευταία πρόκληση που αφορά το τεχνικό κομμάτι είναι ότι ορισμένοι αλγόριθμοι, συμπεριλαμβανομένου αυτού των αρχικών δοκιμών (Random Forest Classifier), δίνουν διαφορετικά αποτελέσματα με τις ίδιες ακριβώς παραμέτρους σε όλο τον κώδικα. Θα εξεταστεί ενδεχόμενη αντικατάστασή του αρχικού αλγορίθμου εξαιτίας αυτού.

Ως εκ τούτου, και εξαιτίας όλων των περιορισμών και των προκλήσεων που περιεγράφηκαν, το αν είναι δυνατόν να δημιουργηθεί ένα τέτοιο πρόγραμμα δεν είναι σίγουρο ότι θα απαντηθεί στην παρούσα έρευνα.

9. Διαδικασία αυτοματοποιημένης ανάλυσης

Πριν ξεκινήσει η διαδικασία τα δεδομένα βρίσκονταν σε μορφή πίνακα σε αρχείο τύπου «.xlsx». Με αυτή τη μορφή εισήχθησαν και στο προγραμματιστικό περιβάλλον Jupiter ώστε να γίνει η επεξεργασία τους με τη γλώσσα Python.

Κατά τη διαδικασία παράθεσης του κώδικα, η οποία θα ακολουθήσει, θα εισάγονται εντός αυτού υποσημειώσεις με υπερσυνδέσμους των επισήμων εγχειριδίων χρήσης της κάθε «βιβλιοθήκης» (library) στην ηλεκτρονική τους μορφή.

Εξ αρχής είχαν εισαχθεί κάποιες βασικές «βιβλιοθήκες» (pandas¹¹, string¹², re¹³, matplotlib¹⁴, numpy¹⁵) προκειμένου να γίνει δυνατή η επεξεργασία και η απεικόνιση των δεδομένων:

```
In [1]: import pandas as pd
import string
import re
import matplotlib
import numpy as np

%matplotlib inline
print("hello")
```

hello

¹¹ <https://pandas.pydata.org/pandas-docs/stable/>

¹² <https://docs.python.org/3/library/string.html>

¹³ <https://docs.python.org/3/library/re.html>

¹⁴ <https://matplotlib.org/>

¹⁵ <https://docs.scipy.org/doc/>

Έπειτα ακολούθησε ο κώδικας για την εισαγωγή του αρχείου δεδομένων, το οποίο ορίστηκε ως μεταβλητή:

```
In [2]: propaganda_df = pd.read_excel('propaganda.xlsx', sheet_name='Analysis')
```

Κατόπιν δημιουργήθηκε νέα στήλη για να κατηγοριοποιηθούν τα κείμενα ως εμπειρέχοντα προπαγάνδα με τη μορφή ενός ή περισσότερων από τα πέντε ορισμένα τεχνάσματα ή μη εμπειρέχοντα:

```
In [5]: propaganda_df['Prop'] = np.where((propaganda_df['Bandwagon Effect']=='Yes') | (propaganda_df['Glittering Generalities']=(propaganda_df['Name Calling']=='Yes') | (propaganda_df['Testimonial/Appeal to Authority'] | (propaganda_df['Transfer']=='Yes') , 1, 0)
```

Όποιο κεφαλαίο γράμμα υπήρχε στα προς ανάλυση κείμενα μετατράπηκε σε μικρό, και δημιουργήθηκε νέα στήλη με τη νέα αυτή μορφή των κειμένων που ορίστηκε με κώδικα ως εξής:

```
In [7]: propaganda_df['Text'] = propaganda_df[propaganda_df['Phrases'].notnull()].apply(lambda row: row['Phrases'].lower().repl
```

Μετά εισήχθησαν «βιβλιοθήκες»(nltk¹⁶, pattern¹⁷,ssl¹⁸) με σκοπό όλες οι λέξεις μέσα στα δεδομένα να αντικατασταθούν από το λήμμα τους, σε οποιαδήποτε από τις γραμματικές μορφές και αν αυτό βρίσκεται: επίθετο , ρήμα, αντωνυμία, επίρρημα. Επίσης, σε αυτό το στάδιο, διαγράφηκαν από τα προς ανάλυση χωρία κειμένου όλες οι λέξεις που εντοπίζονται

¹⁶ <https://www.nltk.org/>

¹⁷ <https://pypi.org/project/pattern3/>

¹⁸ <https://docs.python.org/3/library/ssl.html>

συχνά σε οποιοδήποτε κείμενο στο αγγλικά (“stopwords”) για να μην επηρεαστεί ο αλγόριθμος από τη συχνή παρουσία τους, και μειωθεί η αποτελεσματικότητά του. Ο ορισμός και εφαρμογή των ανωτέρω βημάτων πραγματοποιήθηκαν με τον παρακάτω κώδικα:


```

import nltk
from nltk.corpus import stopwords
nltk.download('punkt')
nltk.download('words')
nltk.download('stopwords')

from nltk.stem import WordNetLemmatizer

stopword_list = nltk.corpus.stopwords.words('english')
wnl = WordNetLemmatizer()

from nltk.corpus import wordnet as wn
import ssl

from pattern.en import tag, parse, pos_tag
from nltk.corpus import wordnet as wn
# Annotate text tokens with POS tags

def pos_tag_text(text):
    # convert Penn treebank tag to wordnet tag

    def penn_to_wn_tags(pos_tag):
        if pos_tag.startswith('J'):
            return wn.ADJ
        elif pos_tag.startswith('V'):
            return wn.VERB
        elif pos_tag.startswith('N'):
            return wn.NOUN
        elif pos_tag.startswith('R'):
            return wn.ADV
        else:
            return None

    tagged_text = tag(text)
    tagged_lower_text = [(word.lower(), penn_to_wn_tags(pos_tag))
                        for word, pos_tag in
                        tagged_text]
    return tagged_lower_text
# lemmatize text based on POS tags

def lemmatize_text(text):
    pos_tagged_text = pos_tag_text(text)
    lemmatized_tokens = [wnl.lemmatize(word, pos_tag) if pos_tag
                        else word
                        for word, pos_tag in pos_tagged_text]
    lemmatized_text = ' '.join(lemmatized_tokens)
    return lemmatized_text

```

```

propaganda_df['Text'] = propaganda_df['Text'].apply(lambda x:lemmatize_text(x))

```

Έπειτα, κατασκευάστηκε προγραμματιστικά ένας πίνακας που κάθε στήλη του ήταν και μία λέξη του τροποποιημένου πλέον συνόλου των κειμένων. Κάθε λέξη που αντιπροσωπεύεται από καθεμία στήλη λαμβάνει μία τιμή για κάθε κείμενο. Αυτή η τιμή υπολογίζεται από μία εξίσωση που προσμετρά συχνότητα όρων αλλά και αντίστροφη συχνότητα όρων μέσα σε έναν αριθμό εγγράφων (Manning, et al., 2008).

Το μοντέλο αυτό της συχνότητας όρων και της αντεστραμμένης συχνότητας του στα έγγραφα (Term Frequency – Inversed Document Frequency) χρησιμοποιήθηκε μέσω ειδικής βιβλιοθήκης που υπάρχει στην Python 3 (Tf-idf Vectorizer¹⁹).

```
from sklearn.feature_extraction.text import TfidfVectorizer
#CountVectorizer(stop_words='english', binary=True)
# I only want you to look for words in the emotional lexicon
# because we don't know what's up with the other words
vec = TfidfVectorizer(sublinear_tf=True, min_df=4, norm='l1', encoding='latin-1', ngram_range=(1, 4), stop_words='english')

#TfidfVectorizer(stop_words = 'english', #RareWords.word,
#                use_idf=True,
#                norm= 'l2' ) #'l1') # ELL - ONE
matrix = vec.fit_transform(propaganda_df['Text'])
vocab = vec.get_feature_names()
wordcount_df = pd.DataFrame(matrix.toarray(), columns=vocab)
wordcount_df.head()
```

Ακολούθησε η σύνδεση και συνένωση του πρωτεύοντος πίνακα δεδομένων με τον πίνακα των χαρακτηριστικών που δημιουργήθηκε από την ανάλυση συχνότητας όρων:

```
vec_propaganda_df = pd.concat([propaganda_df, wordcount_df], axis=1, join='inner')
```

```
vec_propaganda_df.head()
```

¹⁹ 4.2.3. Text feature extraction - https://scikit-learn.org/stable/modules/feature_extraction.html

Σε αυτό το σημείο πραγματοποιείται η εισαγωγή του αλγορίθμου Random Forest Classifier²⁰ καθώς και βιβλιοθηκών που θα διαχωρίσουν τα δεδομένα σε αυτά που θα εκπαιδεύσουν τον αλγόριθμο και σε αυτά που θα δοκιμαστεί η αποτελεσματικότητά του:

```
# Machine Learning
from sklearn.ensemble import RandomForestClassifier, RandomForestRegressor
from sklearn.naive_bayes import MultinomialNB
from sklearn.svm import LinearSVC
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import log_loss, confusion_matrix
from sklearn.metrics import accuracy_score
```

```
X = sent[wordcount_df.columns]
y = sent['Prop']
```

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.25, random_state=0)
```

Ο αλγόριθμος που πρόκειται να χρησιμοποιήσουμε ορίζεται σε μεταβλητή και κατόπιν εφαρμόζεται μέσω μίας βιβλιοθήκης (GridSearchCV²¹) μία διαδικασία αυτόματης ανίχνευσης των βέλτιστων ρυθμίσεων και παραμέτρων για τον αλγόριθμο, η οποία καθορίζει τις βέλτιστες παραμέτρους με κριτήριο την ακρίβεια:

```
rf = RandomForestClassifier()
```

```
grid = {'min_samples_leaf' : [ 5, 10, 20, 40], 'max_features' : ['auto', 'log2', 'sqrt'], 'min_samples_split' : [5, 15,
]}
clf = GridSearchCV(rf, grid, cv=5)
```

```
clf.fit(X_train, y_train)
clf.best_estimator
```

²⁰ <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

²¹ https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html

Έπεται η εφαρμογή του αλγορίθμου στα δεδομένα, αφού έχουν καθοριστεί οι βέλτιστες παράμετροι. Αυτό είναι το στάδιο της εκπαίδευσης του αλγορίθμου κατά το οποίο λαμβάνει υπόψη για τη ταξινόμηση που πρόκειται να κάνει στοιχεία από το 75% των δεδομένων, όπως αυτό ορίστηκε νωρίτερα (`test_size=0.25`). Τα στοιχεία που θα λάβει υπόψη για μεταγενέστερες ταξινομήσεις, μετά από αυτή τη διαδικασία μηχανικής μάθησης, έχουν για εμάς σημασία ίδια με αυτή της ακρίβειας του. Είναι σημαντικό λοιπόν μέχρι αυτό το σημείο ο ορισμός των παραμέτρων και η επεξεργασία των δεδομένων να γίνει με τον βέλτιστο δυνατό τρόπο, και για να επιτευχθεί αυτό θα γίνουν αρκετές δοκιμές αλλάζοντας σημεία του κώδικα έως και αυτό το σημείο (συμπεριλαμβανομένου του αλγοριθμικού μοντέλου):

```
rf = RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
    max_depth=None, max_features='auto', max_leaf_nodes=None,
    min_impurity_decrease=0.0, min_impurity_split=None,
    min_samples_leaf=3, min_samples_split=13,
    min_weight_fraction_leaf=0.0, n_estimators=10, n_jobs=1,
    oob_score=False, random_state=None, verbose=0,
    warm_start=False)
#rf = LinearSVC()
rf.fit(X_train, y_train)
#val_pred_probs = rf.predict_proba(X_val)

val_pred = rf.predict(X_val)
```

```
train_pred_probs = rf.predict_proba(X_train)
train_pred = rf.predict(X_train)
```

```
log_loss(y_val, val_pred_probs, labels=rf.classes_)
```

Σε αυτό σημείο ρυθμίζονται παράμετροι σχετικά με το γράφημα που απεικονίζει τις λέξεις με βάση τις οποίες τα κείμενα, όπως αυτό προκύπτει μετά την επεξεργασία των δεδομένων από τον αλγόριθμο:

```
fi = pd.DataFrame(rf.feature_importances_, X_train.columns)
fi.columns = ['Importance']
fi.sort_values(by = 'Importance', ascending=False)[0:20]
```

Αφού ολοκληρωθεί και η προηγούμενη διαδικασία θα δοθούν κάποιες εντολές που σχετίζονται με το σχεδιασμό του πίνακα που απεικονίζει τις σωστές και τις λάθος προβλέψεις του αλγορίθμου, όταν γίνει η εφαρμογή στα δεδομένα που διαχωρίστηκαν με σκοπό τη δοκιμή και αποτελούν το 25% των συνολικών δεδομένων (στην πρώτη τουλάχιστον δοκιμή). Πιο συγκεκριμένα, ο πίνακας αυτός (confusion matrix) απεικονίζει τις σωστές προβλέψεις και για την ύπαρξη και για τη μη ύπαρξη προπαγάνδας, καθώς και τις λανθασμένες προβλέψεις, όπου είτε ταξινομείται το εκάστοτε σε αυτά που έχουν προπαγάνδα ενώ είχε βρεθεί να μην εμπεριέχει κανένα τέχνασμα είτε σε αυτά που δεν εμπεριέχουν προπαγανδιστική τεχνική ενώ είχε βρεθεί ότι εμπεριείχε. Τα χαρακτηριστικά, λοιπόν, του πίνακα αυτού ορίζονται παρακάτω:

```

print(__doc__)

import itertools
import numpy as np
import matplotlib.pyplot as plt

from sklearn import svm, datasets
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix, f1_score

def plot_confusion_matrix(cm, classes,
                          normalize=False,
                          title='Confusion matrix',
                          cmap=plt.cm.Blues):
    """
    This function prints and plots the confusion matrix.
    Normalization can be applied by setting `normalize=True`.
    """
    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
        print("Normalized confusion matrix")
    else:
        print('Confusion matrix, without normalization')

    print(cm)

    plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar()
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)

    fmt = '.2f' if normalize else 'd'
    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, format(cm[i, j], fmt),
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")

    plt.tight_layout()
    plt.ylabel('True label')
    plt.xlabel('Predicted label')

```

Ο κώδικας για την οπτικοποίηση των εξαχθέντων αποτελεσμάτων δεν παρατίθεται εδώ καθώς θα παρατεθεί στο κεφάλαιο των αποτελεσμάτων.

10. Αποτελέσματα

Σύμφωνα με τα κριτήρια αποτελεσματικότητας που έχουν τεθεί τα αποτελέσματα των δοκιμών θα πρέπει να απεικονίζουν τουλάχιστον το ποσοστό και τα ακριβή στοιχεία της αποτελεσματικότητας της πρόβλεψης του αλγορίθμου καθώς και τα χαρακτηριστικά του κάθε κειμένου που λήφθηκαν υπόψη για την ταξινόμηση.

Για να εξαχθούν και να απεικονιστούν τα αποτελέσματα χρησιμοποιήθηκε ο παρακάτω κώδικας:

Για να σχεδιαστεί η απεικόνιση της σημαντικότητας των λέξεων στην ταξινόμηση:

```
import matplotlib.pyplot as plt

plot = fi.sort_values(by = 'Importance', ascending=True).tail(45).plot.barh(figsize = (14,8), legend = False, title = 'E
fig = plot.get_figure()
fig.savefig('importance.png')
```

Για να σχεδιαστεί ο πίνακας με την απεικόνιση του ακριβούς αριθμού των προβλέψεων:

```
# Compute confusion matrix
conf_matrix = confusion_matrix(y_train, train_pred)

cnf_matrix = confusion_matrix(y_val, val_pred)
np.set_printoptions(precision=2)

# Plot non-normalised confusion matrix
plt.figure()
plot_confusion_matrix(cnf_matrix, classes=rf.classes_,
                      title='Confusion matrix')
```

Για να καθοριστεί αναλυτικότερα η εμφάνιση του πίνακα γράφτηκε επίσης ο παρακάτω κώδικας:

```

print(__doc__)

import itertools
import numpy as np
import matplotlib.pyplot as plt

from sklearn import svm, datasets
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix, f1_score

def plot_confusion_matrix(cm, classes,
                          normalize=False,
                          title='Confusion matrix',
                          cmap=plt.cm.Blues):
    """
    This function prints and plots the confusion matrix.
    Normalization can be applied by setting `normalize=True`.
    """
    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
        print("Normalized confusion matrix")
    else:
        print('Confusion matrix, without normalization')

    print(cm)

    plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar()
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)

    fmt = '.2f' if normalize else 'd'
    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, format(cm[i, j], fmt),
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")

    plt.tight_layout()
    plt.ylabel('True label')
    plt.xlabel('Predicted label')

```

Automatically created module for IPython interactive environment

Τέλος παρατέθηκε η τελευταία γραμμή του κώδικα για να υπολογιστεί η ακρίβεια πρόβλεψής του. Οι τιμές που μπορεί να λάβει είναι από 0 έως 1, με το 0 να σημαίνει την απόλυτη αποτυχία να προβλέψει έστω και μία ταξινόμηση σωστά και το 1 να αποτελεί την απόλυτη επιτυχία στις ταξινομήσεις:

```
print("Test Accuracy ::", accuracy_score(y_val, val_pred))
```


10.1. Αρχικές δοκιμές

Το πρώτο πείραμα στην πρώτη δοκιμή αφορά την ποσότητα των συνεχόμενων λέξεων με την οποία σχετίζεται ο αλγόριθμος (n-grams). Θα αυξομειωθεί η παράμετρος που την ορίζει (n-gram range) προκειμένου να συμπεραθεί η καλύτερη δυνατή για να βελτιωθούν τα αποτελέσματα της προσπάθειάς μας. Σε αυτήν την περίπτωση δοκιμάζουμε 1 με τρεις συνεχόμενες λέξεις ως παράμετρο.

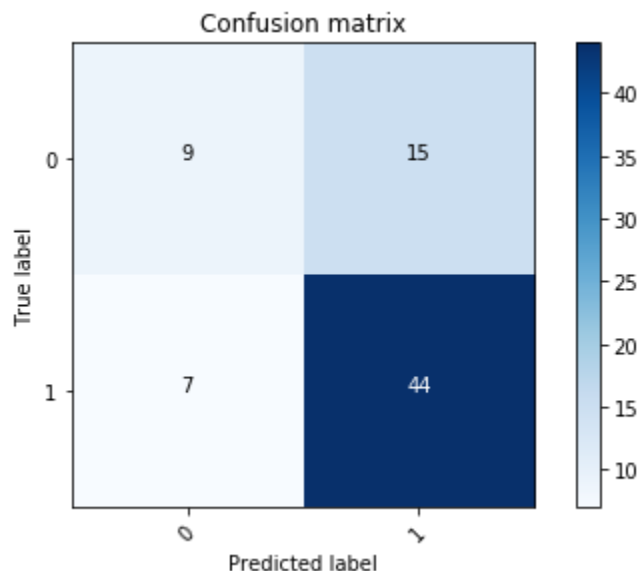
Παράμετροι στον αναλυτή συχνότητας όρων όπως καταγράφηκαν:

```
TfidfVectorizer(sublinear_tf=True, min_df=4, norm='l2', encoding='latin-1',  
ngram_range=(1, 3), stop_words='english')
```

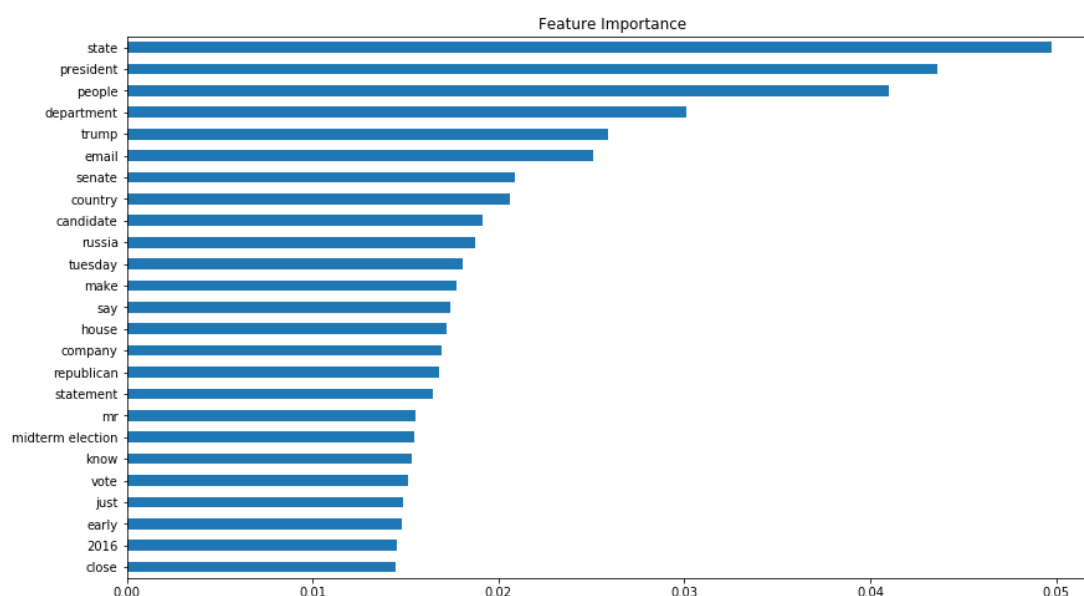
Ακρίβεια του αλγορίθμου αλλάζοντας μόνο την παράμετρο που ορίστηκε παραπάνω:

Test Accuracy :: 0.7066666666666667 - (Ποσοστό ακρίβειας 70%)

Σωστές και λάθος προβλέψεις του αλγοριθμικού μοντέλου σε αυτή τη δοκιμή:



Λέξεις (έχοντας μετατραπεί στο λήμμα τους από την αρχική τους μορφή) τις οποίες έλαβε υπόψη ο αλγόριθμος για την ταξινόμηση



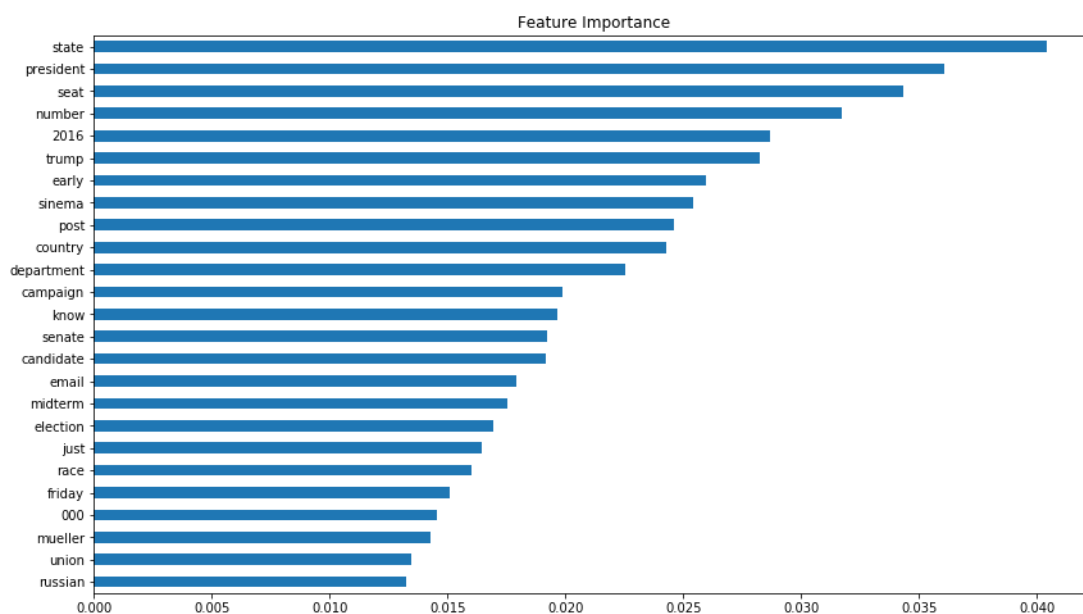
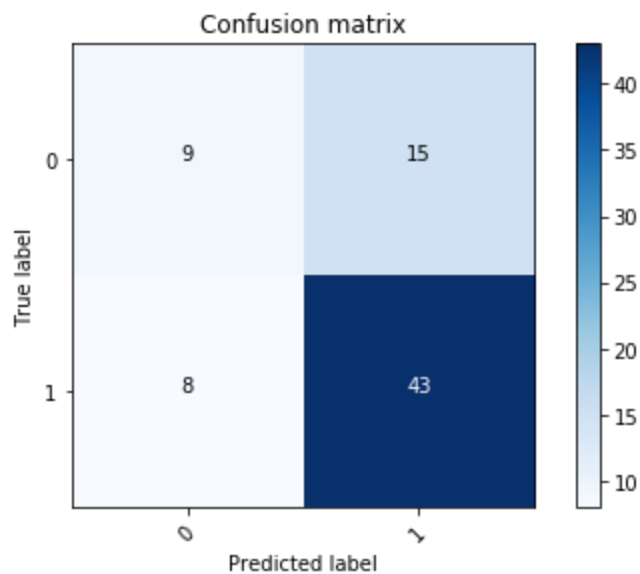
Σε αυτή τη δοκιμή παρά το θετικό ποσοστό ακρίβειας λήφθηκαν αρκετοί όροι υπόψη οι οποίοι δεν δικαιολογούν, σύμφωνα με τη θεωρία ύπαρξης προπαγάνδας. Για παράδειγμα, οι λέξεις “email”, “candidate”, “house”, “department”, “mr”, “early”, “2016”, “tuesday” κατά κανένα τρόπο δεν μπορούν να αποτελέσουν κριτήριο για την ύπαρξη ή μη του χαρακτηριστικού της προπαγάνδας. Η έρευνα βέβαια είναι σε πρώιμο στάδιο και η πειραματισμοί θα συνεχιστούν.

Κρατώντας τις υπόλοιπες παραμέτρους σταθερές αλλάζουμε ξανά το εύρος των n-grams. Αυτή τη φορά θα είναι μία έως τέσσερις συνεχόμενες λέξεις ομαδοποιημένες:

```
vec = TfidfVectorizer(sublinear_tf=True, min_df=4, norm='l2',  
encoding='latin-1', ngram_range=(1, 4), stop_words='english')
```

Σωστές και λανθασμένες προβλέψεις του αλγοριθμικού μοντέλου σε αυτή τη δοκιμή:

Test Accuracy :: 0.72 (72% σωστές προβλέψεις επί του συνόλου των πραγματικών ταξινομήσεων)



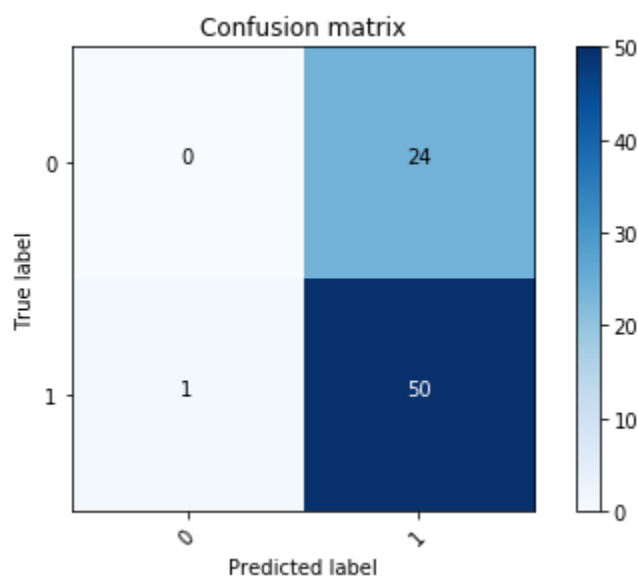
Πάρα την ύπαρξη λέξεων όπως “know”, “country”, “president”, “Russian”(επειδή τα κείμενα προέρχονται σε μεγάλο βαθμό από

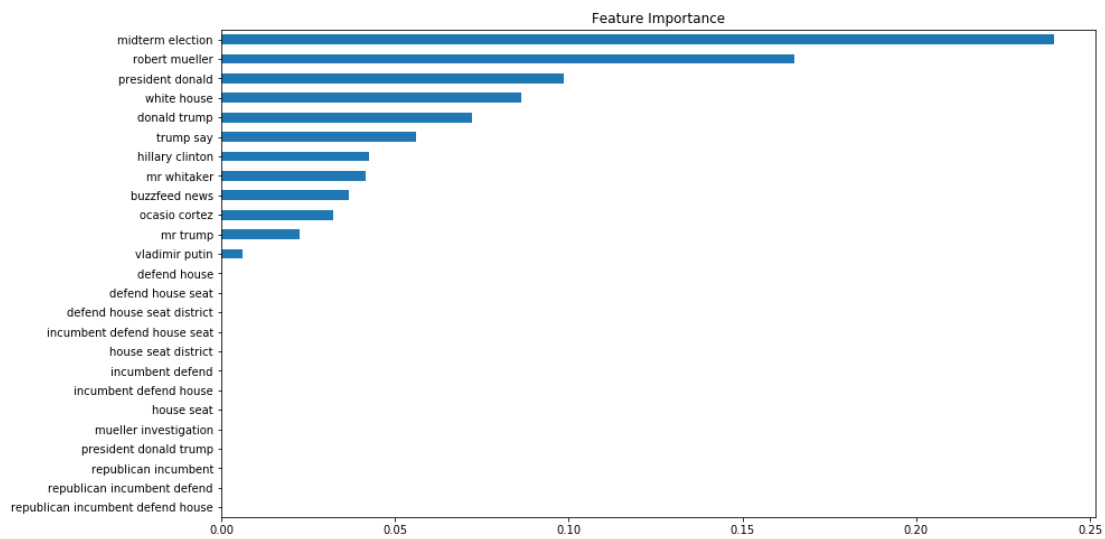
αμερικανικά ΜΜΕ είναι σημαντικός δείκτης) που σε πολιτικά κείμενα και άρθρα ενδέχεται να αποτελούν δείκτη ύπαρξης ή μη τεχνάσματα προπαγάνδας.

Κρατώντας τις υπόλοιπες παραμέτρους και πάλι σταθερές αλλάζουμε ξανά το εύρος των n-grams. Αυτή τη φορά όμως θα είναι από δύο έως τέσσερις συνεχόμενες λέξεις ομαδοποιημένες. Αυτή η παράμετρος αφαιρεί δεδομένα από την ανάλυση με ιδιαίτερη σημασία για την εξεύρεση προπαγάνδας, τα οποία είναι απομονωμένες λέξεις, που είναι σημαντικά και αυτό έχει παρατηρηθεί ήδη από την ποιοτική ανάλυση. Δοκιμάζεται όμως πειραματικά γιατί υποτέθηκε ότι υπήρχε ένα ενδεχόμενο να αναδείξει σημαντικά χαρακτηριστικά για την εξεύρεση.

```
vec = TfidfVectorizer(sublinear_tf=True, min_df=4, norm='l2',  
encoding='latin-1', ngram_range=(2, 4), stop_words='english')
```

Test Accuracy :: 0.6666666666666666 (66% σωστές προβλέψεις επί του συνόλου των πραγματικών ταξινομήσεων)





Από τα παραπάνω είναι φανερό πως η ανάλυση που λαμβάνει υπόψη μία απομονωμένη έως τέσσερις λέξεις μαζί είναι όσο αποτελεσματική χρειάζεται για να δίνει το μέγιστο δυνατό αποτέλεσμα στην ακρίβεια του αλγορίθμου καθώς σε μικρότερο εύρος επηρεάζεται η ακρίβεια πρόβλεψης ενώ αν το εύρος ξεκινά από δύο λέξεις και περισσότερες ομαδοποιημένες χάνονται πολύτιμες πληροφορίες, οι οποίες είναι μεμονωμένοι όροι που ενδέχεται να δηλώνουν την ύπαρξη προπαγάνδας. Για να επιβεβαιωθεί αυτός ο ισχυρισμός ξαναδοκιμάστηκε η ανάλυση με εύρος ομαδοποίησης λέξεων 1 έως 4 με τις εξής παραμέτρους στη βιβλιοθήκη κατάτμησης κειμένου:

```
TfidfVectorizer(sublinear_tf=True, min_df=4, norm='l2', encoding='latin-1', ngram_range=(1, 4), stop_words='english')
```

Και με το εξής αποτέλεσμα στην ακρίβεια:

Test Accuracy :: 0.7222222222222222 (72% σωστές προβλέψεις επί του συνόλου των πραγματικών ταξινομήσεων)

Από αυτό καταλαβαίνουμε πως ο τυχαίος χωρισμός των δεδομένων που συμβαίνει κάθε φορά κατά των χωρισμό του συνόλου δεδομένων εκμάθησης και του συνόλου δεδομένων δοκιμής δεν επηρεάζει τη σταθερά θετική απόδοση του αλγορίθμου με αυτήν την παράμετρο (`ngram_range=(1, 4)`) σε σχέση με άλλες τιμές σε αυτή.

Επιπλέον πρέπει να σημειωθεί ότι και σε μία πειραματική δοκιμή με τον αλγόριθμο Naïve Bayes²² με τον οποίο αναλύθηκαν τα δεδομένα, το εύρος των n-grams ισούταν με μία έως δύο λέξεις μαζί και η ακρίβεια πρόβλεψης του αλγορίθμου ήταν 66% (`Test Accuracy :: 0.6666666666666666`), μικρότερη από κάθε άλλη δοκιμή που έγινε με αυτό τον αλγόριθμο και έγινε με εύρος ομαδοποιημένων λέξεων από 1 έως 4.

Η δοκιμή αυτή αναφέρθηκε προς επίρρωση της θέσης πως οι ιδανικές παράμετροι για την κατάτμηση κειμένου με αυτά τα δεδομένα είναι αυτές που προαναφέρθηκαν. Ο συγκεκριμένος αλγόριθμος προστέθηκε μεταγενέστερα και ο αντίστοιχος κώδικας θα παρατεθεί στο επόμενο κεφάλαιο όπου και νέες δοκιμές και από αποτελέσματα από το βελτιωμένο μοντέλο ανάλυσης θα παρουσιαστούν.

10.2. Βελτίωση και επαναξιολόγηση

Οι πρώτες δοκιμές ουσιαστικά δεν τελεσφόρησαν διότι παρά το θετικό ποσοστό ακρίβειας τα κριτήρια ταξινόμησης δεν ήταν ανάλογα ούτε καν σε έναν σχετικό βαθμό με αυτά που λάβαμε υπόψη κατά την ποιοτική ανάλυση. Δεν μπορούμε σε καμία περίπτωση από τη μέχρι τώρα ανάλυση να πούμε ότι λάβαμε την ένδειξη ότι είναι δυνατόν να δημιουργηθεί πρόγραμμα αυτοματοποιημένης ανάλυσης προπαγάνδας.

²² https://scikit-learn.org/stable/modules/naive_bayes.html

Παρατηρώντας αυτά, αναγκαστήκαμε να βελτιώσουμε τον κώδικα και να προσθέσουμε νέες διαδικασίες ανάλυσης και να αποδώσουμε νέα χαρακτηριστικά στα δεδομένα.

Αρχικά από την βιβλιοθήκη nltk²³ αποφασίσαμε να χρησιμοποιήσουμε ένα λεξικό εξόρυξης συναισθήματος προκειμένου να αποδώσουμε νέα χαρακτηριστικά στα κείμενα και να τροφοδοτήσουμε με νέα και σημαντικά δεδομένα τα αλγοριθμικά μοντέλα (εκτός από το Random Forest Classifier πρόκειται να χρησιμοποιήσουμε και το αλγοριθμικό μοντέλο Naïve Bayes) με στόχο να λάβουμε καλύτερα αποτελέσματα. Εδώ αναφέρουμε επίσης πως το συναίσθημα θεωρούμε πως σχετίζεται με την προπαγάνδα δεδομένου και ότι το συχνότερο τέχνασμα που συναντήθηκε στα δεδομένα όπως προαναφέρθηκε, ήταν η «μεταφορά», τεχνική η οποία αφορά σχεδόν μόνο αυτό. Το λεξικό που χρησιμοποιήθηκε ήταν το “vader lexicon” (Hutto & Gilbert., 2014) το οποίο κατά τη δοκιμή του είχε τις μεγαλύτερες αποδόσεις σε σύγκριση με έντεκα άλλα λεξικά, με ποσοστά τα οποία κυμάνθηκαν από 96% έως και 99% στα ποσοστά ακρίβειας καθορισμού των συναισθημάτων σε κείμενα από κοινωνικά δίκτυα (Hutto & Gilbert., 2014, p. 225), έκτασης ανάλογης με αυτών που αναλύονται σε αυτήν την περίπτωση. Η ανάλυση των λημματοποιημένων φράσεων καθώς και η τοποθέτηση των αποτελεσμάτων αυτής σε στήλες έγινε με τον κώδικα που ακολουθεί:

²³ <https://www.nltk.org/>

```

nltk.download('vader_lexicon')
from nltk.sentiment.vader import SentimentIntensityAnalyzer as sia
sid = sia()

ed=[(sid.polarity_scores(i)['neg']) for i in vec_propaganda_df['Phrases']]
er=[(sid.polarity_scores(i)['neu']) for i in vec_propaganda_df['Phrases']]
et=[(sid.polarity_scores(i)['pos']) for i in vec_propaganda_df['Phrases']]
ut=[(sid.polarity_scores(i)['compound']) for i in vec_propaganda_df['Phrases']]
neg=pd.DataFrame(ed,columns=["neg"])
neu=pd.DataFrame(er,columns=["neu"])
pos=pd.DataFrame(et,columns=["pos"])
compound=pd.DataFrame(ut,columns=["compound"])

ola= pd.concat([neg,neu,pos,compound], axis=1, join='inner')
wordsent= pd.concat([wordcount_df,ola], axis=1, join='inner')
sent=pd.concat([vec_propaganda_df,ola], axis=1, join='inner')

```

```

[nltk_data] Downloading package vader_lexicon to
[nltk_data] /home/maltezos/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!

```

Έπειτα, τοποθετήθηκε στο τέλος ο παρακάτω κώδικας με τον οποίο εισάχθηκε το αλγοριθμικό μοντέλο. Περιλαμβάνονται γραμμές κώδικα για την κατάτμηση των δεδομένων σε δεδομένα εκπαίδευσης του αλγοριθμικού μοντέλου και σε δεδομένα δοκιμής του, όπως έγινε και με το προηγούμενο, καθώς επίσης και ο κώδικας για το γράφημα αναπαράστασης των αποτελεσμάτων των προβλέψεων του αλγορίθμου:

```

nv=MultinomialNB(alpha=0.1)

```

```

nv.fit(X_train, y_train)
valp= nv.predict(X_val)
tbp=nv.predict_proba(X_train)
tp=nv.predict(X_train)

```

```

cnf_matrix = confusion_matrix(y_val, valp)
np.set_printoptions(precision=2)

plt.figure()
plot_confusion_matrix(cnf_matrix, classes=nv.classes_,
                      title='Confusion matrix')

```

Με τον παρακάτω κώδικα απεικονίζονται διαδοχικά το ποσοστό ακρίβειας πρόβλεψης μετά την εφαρμογή του αλγοριθμικού μοντέλου και τα χαρακτηριστικά του κειμένου που λήφθηκαν υπόψη για την

ταξινόμηση των κειμένων. Με αυτόν τον κώδικα λαμβάνονται απευθείας σε μορφή κειμένου και όχι απεικονισμένα:

```
print("Test Accuracy ::", accuracy_score(y_val, valp))
print('-----')
neg_class_prob_sorted = nv.feature_log_prob_[0, :].argsort()
pos_class_prob_sorted = nv.feature_log_prob_[1, :].argsort()

print(np.take(vec.get_feature_names(), neg_class_prob_sorted[:40]))
print(np.take(vec.get_feature_names(), pos_class_prob_sorted[:40]))
```

10.2.1. Δοκιμές και αποτελέσματα

Με το λεξικό εξόρυξης συναισθήματος να έχει προστεθεί υπάρχουν επιπλέον χαρακτηριστικά που βοηθούν την ταξινόμηση. Οι νέες στήλες με τα αποτελέσματα μετά από την ανάλυση με το λεξικό συναισθήματος που προστίθενται έχουν την εξής μορφή:

neg	neu	pos	compound
0.000	0.742	0.258	0.8271

Η στήλη “neg” αντιστοιχεί στην αξιολόγηση του εκάστοτε κειμένου με βάση το αρνητικό συναίσθημα που εντοπίζεται συνολικά σε αυτό, η στήλη “neu” αντιστοιχεί στο βαθμό συναισθηματικής ουδετερότητας, η στήλη “pos” στην αξιολόγηση του θετικού συναισθήματος και η στήλη “compound” σε ένα αποτέλεσμα που εξάγεται από τον συνυπολογισμό των προηγούμενων τριών αξιολογήσεων. Οι τιμές που μπορούν να λάβουν οι τρεις πρώτες στήλες είναι από 0 έως και 1, ενώ η στήλη “compound” μπορεί να λάβει τιμές από -1 έως και 1, με τις τιμές που

βρίσκονται κάτω του 0 να δείχνουν ύπαρξη αρνητικού συναισθήματος και τις τιμές άνω του 0 θετικού συναισθήματος συνολικά.

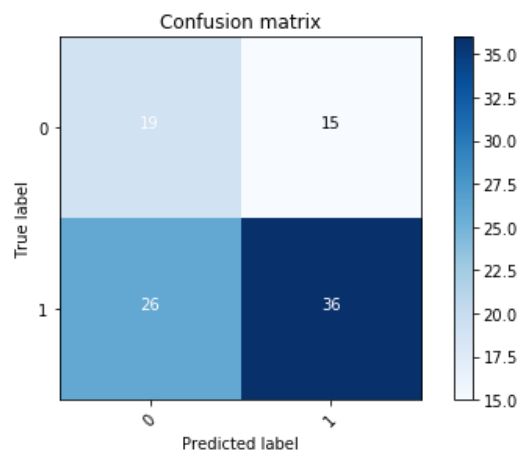
Τα χαρακτηριστικά αυτά δεν μπόρεσαν να αξιοποιηθούν όμως με τον αλγόριθμο "MultinomialNB" (Naïve Bayes) καθώς οι τιμές που λάμβανε μία παράμετρος κατά τις δοκιμές ήταν αρνητική και αυτό δημιουργούσε λάθος στην ανάλυση του αλγορίθμου και ο κώδικας δεν παρήγαγε αποτέλεσμα. Ως εκ τούτου στον εν λόγω αλγόριθμο χρησιμοποιήθηκαν ως δεδομένα οι μεμονωμένες και ομαδοποιημένες λέξεις που είχαν χρησιμοποιηθεί κατά τις αρχικές δοκιμές.

Δοκιμές με το αλγοριθμικό μοντέλο "Random Forest Classifier"

Τα δεδομένα που αφορούν τα συναισθήματα προστέθηκαν στο αρχικό αρχείο των δεδομένων με τις απομονωμένες και ομαδοποιημένες λέξεις. Έπειτα τέθηκε σε εφαρμογή εκ νέου η διαδικασία ανάλυσης με τον αρχικό αλγόριθμο (Random Forest Classifier) με τις βέλτιστες παραμέτρους. Χρησιμοποιώντας τον ίδιο κώδικα με τις αρχικές δοκιμές λάβαμε τα εξής αποτελέσματα:

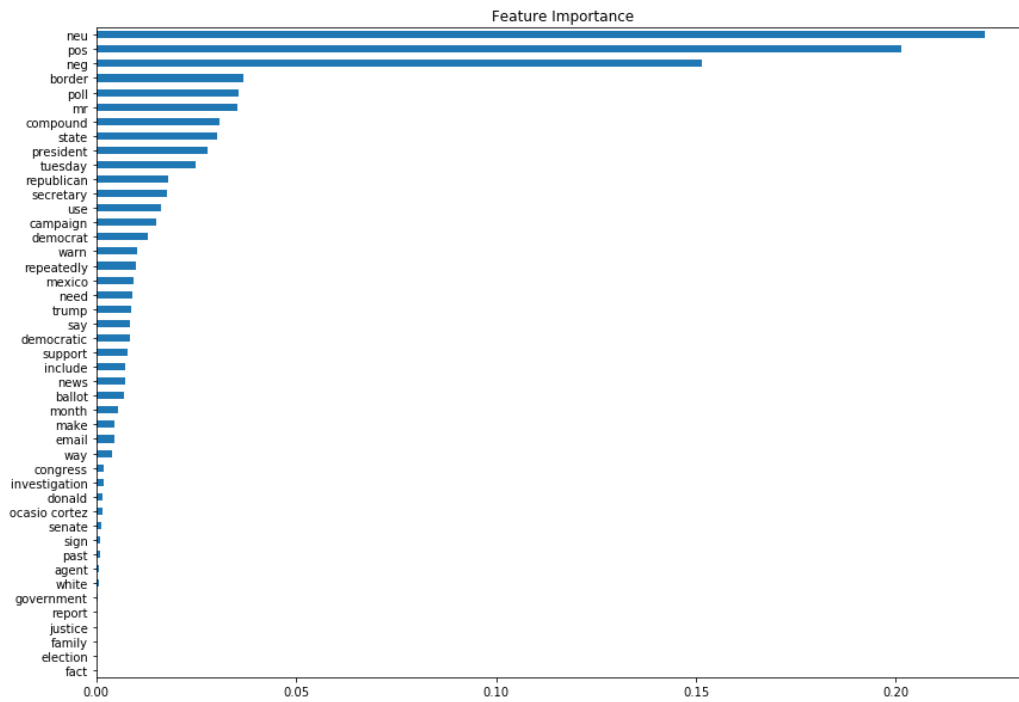
Δοκιμή 1^η

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.30,  
random_state=0)
```



Ποσοστό ακρίβειας:

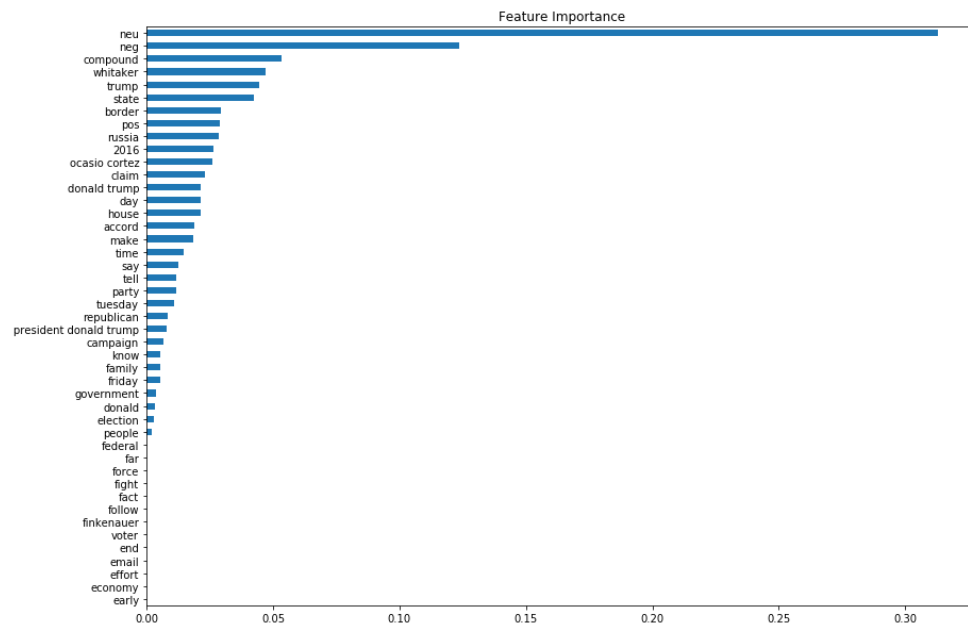
Test Accuracy :: 0.5729166666666666 (57%)



Δοκιμή 2^η

`X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.30,
random_state=0)`

/



Ποσοστό ακρίβειας:
Test Accuracy :: 0.53125 (53%)

Σχόλια

Το ποσοστό ακρίβειας του αλγορίθμου παρατηρήθηκε επαναλαμβανόμενα ότι έπεσε σε σχέση με τις καταγραφές των αρχικών δοκιμών.

Επιπλέον το αλγοριθμικό αυτό μοντέλο, όπως αναφέρθηκε και στο κεφάλαιο των περιορισμών της μεθόδου που χρησιμοποιούμε δε μας δίνει σταθερά αποτελέσματα με σταθερές παραμέτρους.

Δεν δόθηκαν επίσης ενδείξεις, βάσει των χαρακτηριστικών που θεωρήθηκαν σημαντικές κατά τη διαδικασία της ανάλυσης των δεδομένων από τον αλγόριθμο, ότι μπορούμε με αυτή τη μέθοδο να εντοπίζουμε με αυτόματο τρόπο ποια κείμενα εμπεριέχουν προπαγανδιστικά τεχνάσματα. Τα λεξιλογικά στοιχεία με τα οποία πραγματοποιήθηκε η ταξινόμηση από το αλγοριθμικό μοντέλο δεν υποδηλώνουν την ύπαρξη ή όχι κάποιου ή περισσότερων από τα πέντε τεχνάσματα προπαγάνδας, σύμφωνα με τη θεωρία και τα ευρήματα της ποιοτικής ανάλυσης.

Η μέθοδος και η προσπάθεια αυτή θεωρούμε πως απέτυχε.

Παρ' όλα αυτά, η επαναλαμβανόμενα υψηλή σημασία που λάμβαναν οι αξιολογήσεις συναισθήματος για την ταξινόμηση των κειμένων αποτέλεσε μία σημαντική ένδειξη, την οποία λάβαμε από αυτοματοποιημένη ανάλυση και επομένως αντικειμενικά, ότι η προπαγάνδα συνδέεται άρρηκτα με το συναίσθημα και τα προπαγανδιστικά τεχνάσματα επικαλούνται αυτό ως επί το πλείστον.

Σημειωτέο πως αυτή η παρατήρηση δεν έρχεται σε αντιπαράθεση με όσα γνωρίζουμε από τη θεωρία. Την ισχυροποιεί, λειτουργώντας

επιβεβαιωτικά και ενισχυτικά προς αυτήν. Το πιο αξιοσημείωτο όμως είναι πως αυτή η ένδειξη μπορεί να βοηθήσει παρόμοιες μελλοντικές έρευνες.

Δοκιμές με το αλγοριθμικό μοντέλο “MultinomialNB”

Οι παρακάτω καταγραφές των δοκιμών θα έχουν διαδοχικά την εξής μορφή για κάθε δοκιμή: κώδικας για την κατάτμηση με αλλαγμένες παραμέτρους, Γράφημα αναλυτικής παράθεσης των προβλέψεων (confusion matrix), ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση, σχόλια επί των αποτελεσμάτων.

Για τον κώδικα της κατάτμησης των δεδομένων υπογραμμίζονται οι παράμετροι που άλλαξαν και επίσης οι λέξεις που σχετίζονται σύμφωνα με τη θεωρία με οποιοδήποτε από τα πέντε τεχνάσματα υπογραμμίζονται στα στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση. Για αυτόν τον λόγο δεν παρατίθενται με μορφή εικόνας τα αποτελέσματα.

Πριν παραθέσουμε τις δοκιμές πρέπει να αναφέρουμε πως με το αλγοριθμικό μοντέλο αυτό οι δοκιμές έφεραν θετικά αποτελέσματα γι’ αυτό και παρατίθενται περισσότερες από όσες παρατέθηκαν προηγουμένως.

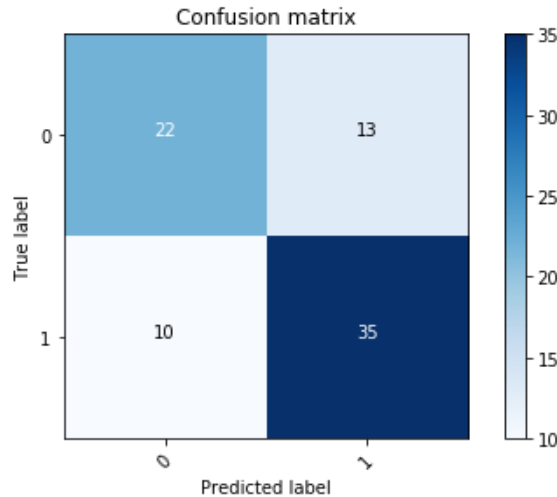
Η πρώτη λίστα λέξεων μετά το ποσοστό επιτυχίας αφορά τα λιγότερο σημαντικά χαρακτηριστικά – κριτήρια ταξινόμησης για την κατάταξη σε μη εμπειριέχοντα προπαγάνδα κείμενα ενώ η δεύτερη τα λιγότερο σημαντικά χαρακτηριστικά – κριτήρια ταξινόμησης για τα εμπειριέχοντα προπαγάνδα κείμενα. Συγκεκριμένα αν στην πρώτη λίστα δούμε κάποιες από τις πρώτες λέξεις να είναι σύμφωνα με τη θεωρία δείκτες για την ύπαρξη μίας τουλάχιστον εκ των προπαγανδιστικών τεχνικών τότε αυτό είναι θετική ένδειξη για την προσπάθειά μας.

Δοκιμή 1^η

Κώδικας για την κατάτμηση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.25,  
random_state=1)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

Test Accuracy :: 0.7125

```
-----  
['far' 'want' 'include' 'immigration' 'child' 'hillary clinton'  
'hillary' 'health' 'grow' 'great' 'goal' 'saudi arabia' 'continue'  
'fact' 'think' 'end' 'session' 'strong' 'business' 'warn' 'cast'  
'liberal' 'ask' 'protect' 'political' 'midterm' 'recent' 'midterm  
election' 'arabia' 'play' '2018' 'mueller investigation' 'believe'  
'best' 'job' 'blue' 'defend' 'district' 'seat district' 'defend house  
seat district' 'republican incumbent' 'defend house' 'incumbent  
defend house seat' 'incumbent defend house' 'incumbent defend'  
'defend house seat' 'house seat' 'republican incumbent defend'  
'republican incumbent defend house' 'house seat district']  
['000' 'company' 'mexico' 'incumbent defend house seat', 'incumbent  
defend house' 'incumbent defend' 'incumbent', 'house seat district'  
'ocasio' 'house seat' 'finkenauer' 'early' 'district' 'defend house  
seat district' 'defend house seat' 'defend house' 'country' 'governor'  
'ocasio cortez' 'nation' 'record' '20' 'washington' 'trump win'  
'personal' 'secretary' 'seat district' 'robert mueller' 'speak'  
'republican incumbent defend house' 'republican incumbent defend'  
'republican incumbent' 'repeatedly' 'robert' 'cortez' 'special'  
'friday' 'counsel' 'mr whitaker' 'special counsel' 'moscow'  
'statement' 'wednesday' 'lose' 'number' 'run' 'vladimir putin'  
'vladimir' 'hour' 'ballot']
```

Δοκιμή 2^η

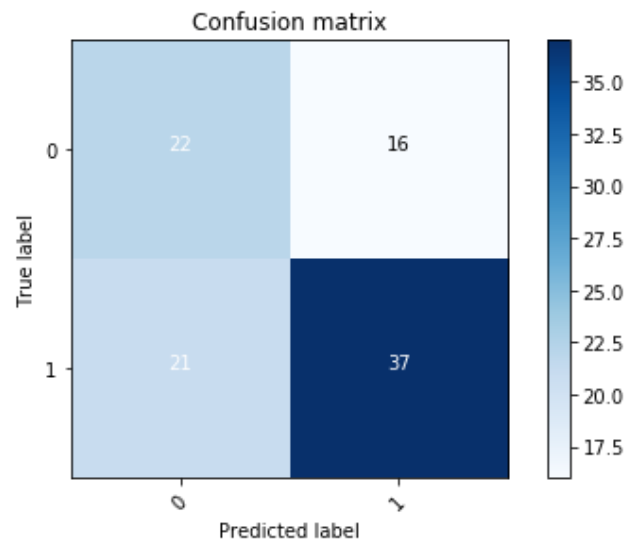
Κώδικας για την κατάρτιση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,  
random_state=2)
```

```
nv=MultinomialNB(alpha=1)
```

```
nv=MultinomialNB(alpha=0.1)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

```
Test Accuracy :: 0.6770833333333334
```

```
-----  
['democratic' 'business' 'buzzfeed' 'buzzfeed news' 'present'  
'political' 'cast' 'play' 'grow' 'charge' 'child' 'nearly' 'nation'  
'health' 'midterm election' 'midterm' 'immigration' 'liberal' 'bring'  
'job' 'putin' 'best' 'warn' 'vladimir putin' 'vladimir' 'trade'  
'think' 'far' 'support' 'strong' 'blue' 'force' 'reporter' 'goal'  
'recent' 'great' 'believe' 'include' 'end' 'right' 'control']
```

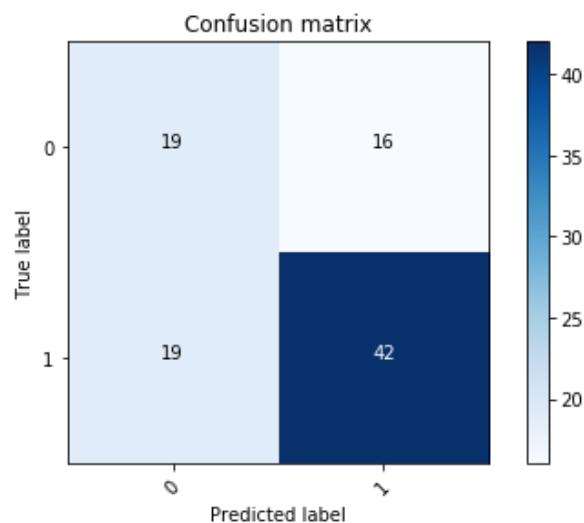
```
'attorney' 'program' 'ask' 'base' 'mr cohen' 'attorney general'
'meeting' 'saudi' 'year']
['000' 'minority' 'mexico' 'incumbent defend house seat' 'incumbent
defend house' 'incumbent defend' 'incumbent' 'house seat district'
'defend house' 'house seat' 'good' 'come' 'early' 'committee'
'company' 'defend house seat district' 'defend house seat' 'governor'
'project' 'decide' 'republican incumbent' '20' 'washington'
'republican incumbent defend' 'republican incumbent defend house'
'run' 'russia' 'seat district' 'secretary' 'trump win' 'standard'
'tuesday' 'believe' 'group' 'meeting' 'friday' 'mueller' 'win' 'mr
whitaker' 'personal' 'regulation' 'wednesday' 'statement' 'robert'
'speak' 'community' 'lose' 'mueller investigation' 'repeatedly'
'robert mueller']
```

Δοκιμή 3^η

Κώδικας για την κατάτμηση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,
random_state=30)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

Test Accuracy :: 0.6354166666666666

```
['best' 'immigration' 'strong' 'business' 'seat need' 'follow'
'liberal' 'need' 'right' 'reuters' 'special counsel' 'goal' 'control'
'child' 'warn' 'great' 'blue' 'grow' 'health' 'include' 'job'
'midterm election' 'percent' 'think' 'far' 'american' 'recent']
```



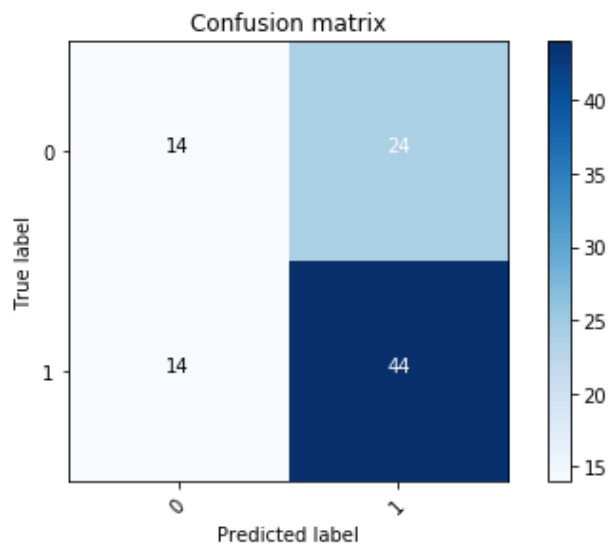
```
'majority' 'believe' 'play' 'political' 'midterm' 'end' 'law'
'program' 'ask' 'saudi' 'trump campaign' 'lose' 'case' 'pick' 'bring'
'increase' 'cast' 'nation' 'democratic' 'arabia' 'saudi arabia'
'session' 'counsel']
['000' 'mexico' 'incumbent defend house seat' 'incumbent defend
house' 'incumbent defend' 'increase' 'house seat district' 'house
seat' 'hour' 'governor' 'forward' 'early' 'drop' 'district' 'defend
house seat district' 'defend house seat' 'defend house' 'post'
'bring' 'sign' '20' 'wednesday' 'washington' 'trump win' 'statement'
'project' 'secretary' 'seat district' 'defend' 'result' 'republican
incumbent defend house' 'republican incumbent defend' 'republican
incumbent' 'tuesday' 'email' 'white house' 'friday' 'mueller' 'mr
whitaker' 'moscow' 'personal' 'robert' 'speak' 'lose' 'robert
mueller' 'mueller investigation' 'repeatedly' 'incumbent' 'vladimir
putin' 'vladimir']
```

Δοκιμή 4^η

Κώδικας για την κατάτμηση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,
random_state=32)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

```
Test Accuracy :: 0.6041666666666666
```

```

['liberal' 'business' 'include' 'immigration' 'decision' 'health'
'grow' 'great' 'midterm election' 'goal' 'play' 'political' 'think'
'far' 'strong' 'present' 'economy' 'child' 'blue' 'bring' 'believe'
'like' 'know' 'khashoggi' 'warn' 'want' 'john' 'agency' 'job' 'best'
'issue' 'world' 'recent' 'mexico' 'midterm' 'conspiracy' 'hillary'
'hillary clinton' 'claim' 'end' 'right' 'putin' 'control' 'protect'
'mr cohen' 'vladimir' 'vladimir putin' 'ask' 'continue' 'base']
['000' 'result' 'republican incumbent defend house' 'republican
incumbent defend' 'charge' 'republican incumbent' 'repeatedly'
'release' 'prosecutor' 'company' 'deal' 'robert' 'defend house'
'defend house seat district' 'month' 'mexico' 'early' 'incumbent
defend house seat' 'incumbent defend house' 'incumbent defend'
'finkenauer' 'house seat district' 'house seat' 'defend house seat'
'robert mueller' 'governor' 'seat district' '20' 'wednesday'
'washington' 'use' 'trump win' 'trump say' 'statement' 'secretary'
'tuesday' 'special' 'meeting' 'attorney general' 'friday' 'counsel'
'mr whitaker' 'federal' 'moscow' 'special counsel' 'personal'
'regulation' 'speak' 'decision' 'thing']

```

Δοκιμή 5^η

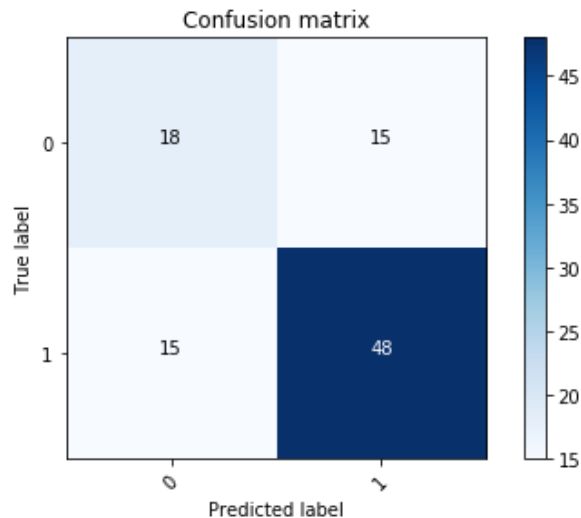
Κώδικας για την κατάτμηση:

```

X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,
random_state=50)

```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

Test Accuracy :: 0.6875

```

-----
['goal' 'warn' 'bring' 'business' 'buzzfeed' 'buzzfeed news' 'job'
 'recent' 'midterm election' 'midterm' 'liberal' 'child' 'follow'
 'right' 'special counsel' 'saudi arabia' 'best' 'blue' 'strong'
 'immigration' '2018' 'include' 'believe' 'political' 'agent' 'health'
 'grow' 'arabia' 'ask' 'play' 'think' 'great' 'policy' 'far' 'hillary'
 'hillary clinton' 'end' 'putin' 'control' 'mr cohen' 'vladimir putin'
 'vladimir' 'just' 'continue' 'trump campaign' 'case' 'million'
 'party' 'need' 'cast']
['000' 'defend house seat' 'mr whitaker' 'mexico' 'meeting' 'john'
 'incumbent defend house seat' 'incumbent defend house' 'incumbent
 defend' 'house seat district' 'house seat' 'hour' 'governor' 'friday'
 'fact' 'email' 'deal' 'early' 'drop' 'defend house' 'number' 'obama'
 'chamber' 'defend house seat district' 'statement' 'release' 'trump
 win' '23' 'seat district' 'republican incumbent defend house'
 'result' 'republican incumbent' 'washington' 'wednesday' '20'
 'republican incumbent defend' 'secretary' 'month' 'general'
 'attorney' 'attorney general' 'moscow' 'personal' 'robert' 'record'
 'speak' 'robert mueller' 'repeatedly' 'turn' 'incumbent']

```

Δοκιμή 6^η

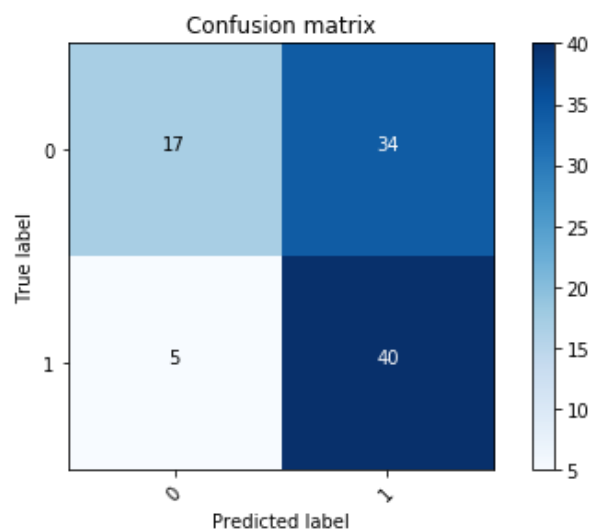
Κώδικας για την κατάτμηση:

```

X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,
random_state=53)

```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

```

Test Accuracy :: 0.625

```

```

['year' 'cast' 'include' 'immigration' 'hold' 'child' 'health'
'think' 'grow' 'recent' 'economy' 'great' 'record' 'goal' 'standard'
'conspiracy' 'follow' 'far' 'nearly' 'end' 'session' 'strong' 'job'
'case' 'midterm' 'warn' 'military' 'agent' 'want' 'liberal' 'play'
'mueller investigation' 'attorney' 'nation' 'attorney general'
'political' 'believe' 'best' 'blue' 'business' 'midterm election'
'claim' 'poll' 'right' 'agency' 'continue' 'control' 'protect' 'deny'
'meeting']
['000' 'personal' 'defend house seat' 'mexico' 'incumbent defend
house seat' 'incumbent defend house' 'incumbent defend' 'early'
'incumbent' 'house seat district' 'house seat' 'governor' 'defend
house seat district' 'forward' 'fact' 'increase' 'committee' 'record'
'speak' 'republican incumbent defend' 'republican incumbent defend
house' '20' 'defend house' 'seat district' 'secretary' 'single'
'republican incumbent' 'defend' 'trump campaign' 'trump win'
'washington' 'deal' 'week' 'congressional' 'ballot' 'deny' 'moscow'
'month' 'friday' 'release' 'mr whitaker' 'statement' 'wednesday'
'robert' 'counsel' 'lose' 'repeatedly' 'special counsel' 'robert
mueller' 'run']

```

Δοκιμή 7^η

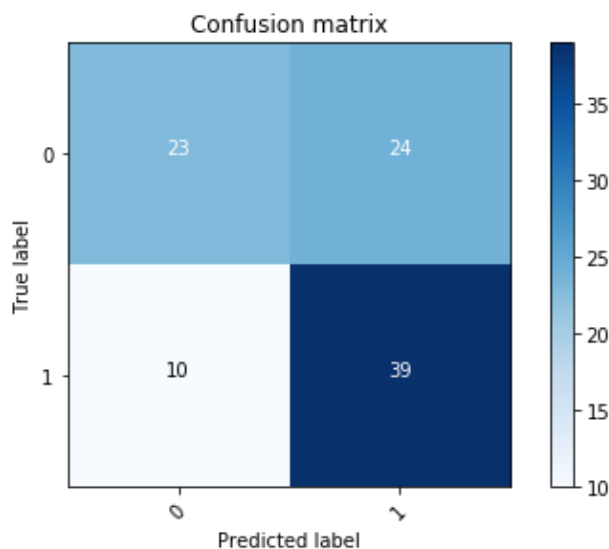
Κώδικας για την κατάτμηση:

```

X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,
random_state=73)

```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

Test Accuracy :: 0.6458333333333334

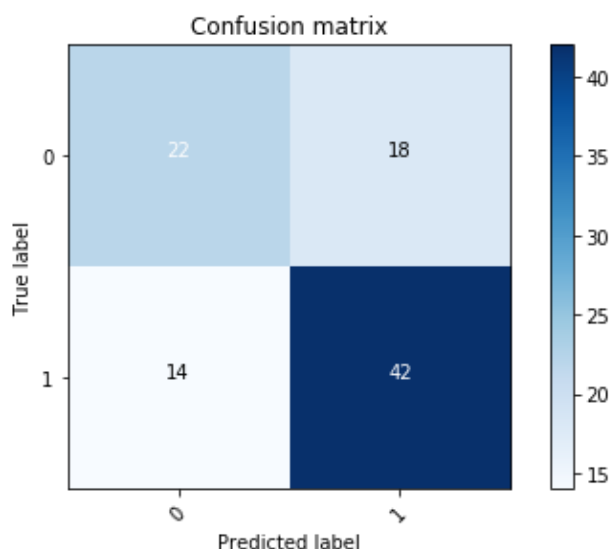
['justice' 'nation' 'case' 'great' 'chamber' 'strong' 'goal' 'child'
'political' 'warn' 'midterm' 'think' 'play' 'percent' 'far' 'fact'
'obama' 'midterm election' 'economy' 'democratic' 'country' 'just'
'health' 'buzzfeed news' 'grow' 'business' 'single' 'recent' 'job'
'agent' 'law' 'buzzfeed' 'include' 'immigration' 'raise' 'national'
'liberal' 'hold' 'hillary clinton' 'believe' 'best' 'hillary' 'blue'
'incumbent' 'district' 'defend house seat' 'defend house seat'
district' 'house seat district' 'seat district' 'incumbent defend']
['000' 'republican incumbent defend' 'republican incumbent' 'record'
'prosecutor' 'governor' 'seat district' 'secretary' 'speak'
'personal' 'defend house' 'trump win' 'result' 'defend house seat'
district' 'defend house seat' 'month' 'early' 'mexico' 'incumbent'
defend house seat' 'incumbent defend house' 'washington' 'incumbent'
defend' 'increase' 'house seat district' 'house seat' '20' 'district'
'republican incumbent defend house' 'tuesday' 'reuters' 'charge'
'conspiracy' 'want' 'meeting' 'friday' 'mr whitaker' 'federal'
'wednesday' 'statement' 'robert' 'repeatedly' 'robert mueller'
'incumbent' 'nearly' 'hour' 'good' 'company' 'drop' 'defend' 'union']

Δοκιμή 8η

Κώδικας για την κατάτμηση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,  
random_state=78)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

Test Accuracy :: 0.6145833333333334

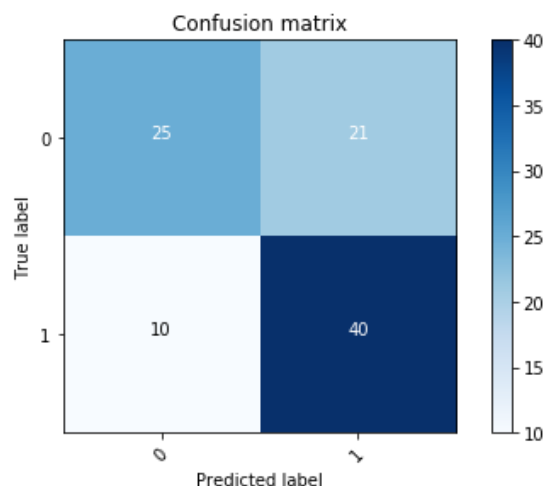
```
-----  
['midterm' 'grow' 'turkish' 'play' 'great' 'political' 'chamber'  
 'saudi arabia' 'goal' 'health' 'child' 'think' 'support' 'strong'  
 'raise' 'seat need' 'community' 'single' 'follow' 'fight' 'session'  
 'business' 'pick' 'midterm election' 'warn' 'liberal' 'military'  
 'mueller investigation' 'majority' 'job' 'need' 'end' 'high'  
 'conspiracy' 'arabia' 'far' 'recent' 'saudi' 'include' 'immigration'  
 'believe' 'best' 'blue' 'american' 'claim' 'poll' 'right' 'continue'  
 'country' 'putin']  
['000' 'republican incumbent defend' 'republican incumbent' 'post'  
 'personal' 'defend house' 'defend house seat' 'defend house seat  
 district' 'obama' 'district' 'republican incumbent defend house'  
 'month' 'mexico' 'men' 'world' 'incumbent defend house seat'  
 'incumbent defend house' 'incumbent defend' 'incumbent' 'house seat  
 district' 'house seat' 'early' 'result' 'governor' 'sinema' 'ballot'  
 'seat district' 'secretary' 'trump win' 'washington' '20' 'speak'  
 'friday' 'hillary' 'hillary clinton' 'mr whitaker' 'statement'  
 'wednesday' 'robert' 'repeatedly' 'robert mueller' 'run' 'vladimir  
 putin' 'vladimir' 'bring' 'hour' 'good' 'company' 'drop' 'defend']
```

Δοκιμή 9^η

Κώδικας για την κατάτμηση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.3,  
random_state=95)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Ποσοστό ακρίβειας και λέξεις – στοιχεία που λήφθηκαν υπόψη για την ταξινόμηση:

Test Accuracy :: 0.6770833333333334

```
-----  
['president donald' 'present' 'political' 'child' 'play' 'past'  
'nearly' 'national' 'midterm election' 'midterm' 'community' 'like'  
'donald trump' 'liberal' 'job' 'include' 'control' 'immigration'  
'health' 'grow' 'great' 'goal' 'force' 'fight' 'far' 'justice'  
'president donald trump' 'donald' 'want' 'think' 'strong' 'saudi  
arabia' 'arabia' '2018' 'ballot' 'trump say' 'believe' 'turkish'  
'blue' 'warn' 'saudi' 'bring' 'business' 'recent' 'best' 'end'  
'candidate' 'right' 'special' 'general']  
['000' 'republican incumbent defend house' 'close' 'republican  
incumbent defend' 'republican incumbent' 'release' 'defend house'  
'defend house seat' 'defend house seat district' 'number' 'mr  
whitaker' 'early' 'mexico' 'john' 'incumbent defend house seat'  
'incumbent defend house' 'friday' 'incumbent defend' 'incumbent'  
'house seat district' 'governor' 'house seat' 'seat district'  
'secretary' 'prosecutor' '20' 'wednesday' 'washington' 'trump win'  
'statement' 'charge' 'deny' 'month' 'attorney general' 'personal'  
'robert' 'nearly' 'record' 'speak' 'robert mueller' 'repeatedly'  
'run' 'hour' 'good' 'company' 'clinton' 'drop' 'defend' 'present'  
'2016']
```

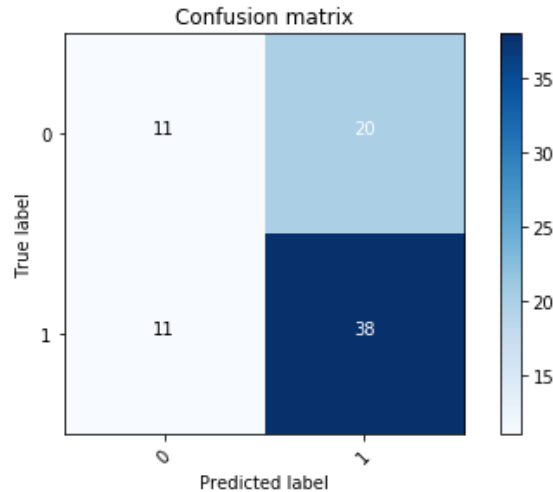
Μέχρι αυτό το σημείο λήφθηκαν από τον κώδικα 50 λέξεις για την ταξινόμηση σε προπαγανδιστικά κείμενα και 50 για την ταξινόμηση σε μη προπαγανδιστικά. Επειδή τα αποτελέσματα που λάβαμε ήταν ενθαρρυντικά σε κάθε δοκιμή αυξήσαμε τη λήψη των χαρακτηριστικών σε 200 προκειμένου να λάβουμε περισσότερες πληροφορίες για τα κριτήρια ταξινόμησης και για τη συμφωνία τους ή όχι με τα κριτήρια ταξινόμησης που λήφθηκαν υπόψη κατά την ποιοτική ανάλυση.

Δοκιμή 10^η

Κώδικας για την κατάτμηση:

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.25,  
random_state=15)
```

Γράφημα αναλυτικής παράθεσης των προβλέψεων:



Test Accuracy :: 0.6125

['liberal' 'child' 'job' 'democratic' 'include' 'immigration' 'want' 'health' 'trade' 'business' 'bring' 'warn' 'blue' 'best' 'grow' 'believe' 'midterm election' 'support' 'strong' 'recent' 'single' 'midterm' 'far' 'political' 'goal' 'great' 'play' 'military' 'think' 'voter' 'hillary' 'hillary clinton' 'end' 'right' 'control' 'deny' 'trump campaign' 'ask' 'mr cohen' 'case' 'party' 'need' 'cast' 'nation' 'putin' 'arabia' 'saudi arabia' 'world' 'vladimir putin' 'session' 'past' 'seat need' 'majority' 'obama' 'mueller' 'investigation' 'like' 'agent' 'week' 'national' 'turkish' 'buzzfeed' 'buzzfeed news' 'security' 'conspiracy' 'economy' 'result' 'change' 'hour' 'justice' 'agency' 'kill' 'minority' 'fight' 'chamber' 'charge' 'nearly' 'congress' '2018' 'lose' 'later' 'team' 'decision' 'run' 'speak' 'candidate' 'community' 'just' 'continue' 'white' 'counsel' 'follow' 'law' 'raise' 'claim' 'reuters' 'special counsel' 'vladimir' 'million' 'report' 'white house' 'district' 'moscow' 'release' 'republican incumbent defend' 'republican incumbent defend house' 'defend house seat' 'defend house' 'republican incumbent' 'seat district' 'house seat district' 'incumbent defend house' 'incumbent defend' 'defend house seat district' 'house seat' 'incumbent defend house seat' 'turn' 'fact' 'attorney' 'repeatedly' 'saudi' 'attorney general' 'pick' 'poll' 'mr trump' 'leader' 'present' 'know' 'time' 'country' 'increase' 'high' 'percent' 'reporter' 'county' 'program' 'accord' 'drop' 'issue' 'administration' 'way' 'force' 'policy' 'sign' 'khashoggi' 'thing' 'standard' 'meeting' 'day' 'record' 'ballot' 'ocasio' 'ocasio cortez' 'cortez' 'leave' 'incumbent' 'men' 'trump say' 'president donald trump' 'president donald' 'base' 'regulation' 'hold' 'american' 'general' 'defend' 'official' 'russia' 'secretary' 'year' 'robert mueller' 'special' 'mr whitaker' 'russian' 'protect' 'wednesday' 'intelligence' 'allow' 'deal' 'new' 'people' 'donald trump' 'donald' 'comment' 'month' 'whitaker' '000' 'effort' 'clinton' 'win' 'union' 'sinema' 'use' 'trump win' 'statement' 'work' 'robert' 'investigation' 'congressional' 'finkenauer' 'tuesday']

['000' 'result' 'republican incumbent' 'company' 'governor' 'mr whitaker' 'decide' 'seat district' 'moscow' 'secretary' 'defend house' 'defend house seat' 'defend house seat district' 'mexico' 'republican incumbent defend house' 'early' 'republican incumbent defend' 'incumbent defend house seat' 'incumbent defend house' 'incumbent defend' 'increase' 'trump win' 'allow' 'house seat

district' 'use' 'washington' '20' '23' 'house seat' 'friday' 'issue'
'personal' 'wednesday' 'statement' 'counsel' 'robert' 'record'
'speak' 'reuters' 'repeatedly' 'special counsel' 'robert mueller'
'incumbent' 'hour' 'case' 'ballot' 'good' 'drop' 'reporter' 'defend'
'minority' 'saudi arabia' 'arabia' 'forward' 'district' 'committee'
'white house' 'week' 'world' 'pick' 'tuesday' 'chamber' 'team' 'seat'
need' 'post' 'project' 'prosecutor' 'finkenauer' 'ask' 'county'
'standard' 'force' 'trump campaign' 'midterm' 'midterm election'
'attorney general' 'number' 'john' 'release' 'run' 'hillary clinton'
'hillary' 'month' 'sinema' 'conspiracy' 'charge' 'lose' 'health'
'percent' 'present' 'agent' 'deny' 'khashoggi' 'email' 'congressional'
'blue' 'raise' 'vladimir' 'vladimir putin' 'law' 'cast' 'border'
'bring' 'security' 'ocasio cortez' 'ocasio' 'cortez' 'accord'
'member' 'grow' 'men' 'want' 'sign' 'high' 'program' 'special' 'come'
'regulation' 'agency' 'russia' 'job' 'thing' 'clinton' 'mr cohen'
'win' 'attorney' 'union' 'mueller' 'meeting' 'mr trump' 'like'
'comment' 'nearly' 'deal' 'seat' 'buzzfeed news' 'buzzfeed' 'putin'
'department' 'business' 'general' 'new' 'fight' 'political' 'mueller'
investigation' 'turn' 'close' 'justice' 'intelligence' 'majority'
'official' 'hold' 'federal' 'turkish' 'obama' 'lead' 'effort' 'trade'
'economy' 'administration' 'later' 'believe' 'goal' 'decision'
'follow' 'million' 'cohen' '2016' '2018' 'fact' 'trump say'
'president donald trump' 'president donald' 'community' 'saudi'
'kill' 'control' 'voter' 'end' 'best' 'whitaker' 'include' 'recent'
'continue' 'news' 'tell' 'time' 'single' 'nation' 'military' 'leave'
'report' 'think' 'strong' 'white' 'leader' 'work' 'far' 'immigration'
'state']

Σχόλια επί των αποτελεσμάτων

Και στην τελευταία δοκιμή που παρατέθηκαν περισσότερες λέξεις – κριτήρια ταξινόμησης και στις προηγούμενες έγινε φανερό πως με τον αλγόριθμο “Naïve Bayes” λάβαμε επανειλημμένα θετικά ποσοστά ακρίβειας και θετική συσχέτιση μεταξύ κριτηρίων ταξινόμησης και την αυτοματοποιημένη και την ποιοτική ανάλυση. Οι παρατηρήσεις αυτές έχουν κομβική σημασία για το όλο εγχείρημα, η οποία θα αναλυθεί στα συμπεράσματα.

Σύγκριση αποτελεσμάτων βάσει δεδομένων της ποιοτικής ανάλυσης

Τα ενδεικτικά λεξικά που εξήχθησαν από την προγενέστερη ποιοτική έρευνα για το κάθε τέχνασμα προπαγάνδας ξεχωριστά αποτελούν το αντικειμενικότερο κριτήριο αξιολόγησης της ορθότητας των στοιχείων που λαμβάνει υπόψη ο αλγόριθμος “MultinomialNB” (Naïve Bayes) και θα ενωθούν στο παρακάτω μεγάλο λεξικό. Οι λέξεις που είναι

υπογραμμισμένες σε αυτό βρέθηκαν ανάμεσα στις λέξεις που έλαβε υπόψη ο αλγόριθμος για την ταξινόμηση:

“Fascist demagogue, dictator, red, oligarchy, communist, alien, agitator, royalist, utopian, rabble-rouser, troublemaker, Tory, bigot, national security, threat, illegal, alleged, surreal, hard-line, unacceptable, Trumpy-ish, militaristic, atrocious, horrible, pushers, hell-raisers, drunkards, who raise hell, foreign, inflammatory, witch hunt, rigged, fake news, isolationist, outrageous, controversial, violent, leftist, leftists, uninformed, mob, bizarre, globalist, stunt, disastrous, tribal, typical, populist, mythical, Putin’s Chef, white working-class, Trump-like, Rust Belt, blue-collar, media reported assassins, rogue, relationship-altering, enemy, cover-up, absurd, dangerous, Blue Dog, Terrorist, terrorists, Evil, The worst, The death of America, Armageddon, Hoax, xenophobic, polls, analysts, experts, environmental groups, based on, analysts, experts, sources, plenty of, largely viewed, that earned nearly, outrage, people, is believed to, said to be, polls, 53% of those, international outcry, 50%<, outcry, most people, majority, more than, widely seen, percent, for millions, civilization, christianity, good, proper, right, democracy, patriotism, motherhood, fatherhood, scientific method, love, peace, perfect, different, differently, safer, hopeful, prosperity, reform, progress, future, freedom, amazing, tremendous, strong, phenomenal, fabulous, best, beautiful, truth, freedom, honor, liberty, social justice, public service, loyalty, the American way, generosity, brotherhood, glory, family values, honesty, patriotic, transparency, the right thing, the American dream, all options, fantastic thing, great, the biggest change, funny, tremendous step, feel much differently , beautiful, law, strength, conspiracy, division, not right, economic miracles, durable politics, take our country back, America works, Security, power, special, inability, innocence, order, fairness, justice, discord,

great, strongest, socialism, mostly true, progressive, no small thing, vision, visions, everyday Americans”

Πρέπει να σημειωθεί αρχικά πως στο λεξικό αρκεί να βρεθούν κάποιες λέξεις και όχι όλες καθώς δημιουργήθηκε από δεδομένα που βρέθηκαν κατά την ποιοτική έρευνα ως μία ενδεικτική προσπάθεια να δημιουργηθεί μία ακόμη μέθοδος με την οποία θα επιχειρήσουμε να βρούμε προπαγάνδα στο μέλλον. Το δεύτερο και σημαντικότερο γεγονός που πρέπει να σημειωθεί είναι πως όταν βρέθηκε το λήμμα μίας εκ των παραπάνω λέξεων ως σημαντικό για την ταξινόμηση χαρακτηριστικό του κειμένου, υπογραμμίστηκε και η λέξη στο παραπάνω λεξικό. Δεν λήφθηκαν υπόψη δηλαδή μόνο αυτούσιες οι λέξεις αλλά και τα λήμματά τους.

11. Συμπεράσματα

Τα αποτελέσματα που λάβαμε τόσο από τις πρώτες δοκιμές όσο και από τις δοκιμές μετά τον εμπλουτισμό των δεδομένων με την αξιολόγηση συναισθήματος, οι οποίες έγιναν με τον Random Forest Classifier, δεν ήταν ενθαρρυντικά και σε καμία περίπτωση επαρκή για να λάβει το θεωρητικό μας ερώτημα μία θετική απάντηση. Παρά τα θετικά ποσοστά ακρίβειας, τα οποία δεν κρίθηκαν ιδιαίτερα υψηλά, τα κριτήρια αξιολόγησης δεν ήταν σε καμία περίπτωση ανάλογα με αυτά που λαμβάνονται υπόψη σε μία ποιοτική έρευνα σύμφωνα με τη μέχρι τώρα θεωρία για την κάθε προπαγανδιστική τεχνική ξεχωριστά και για την προπαγάνδα συλλήβδην.

Αντίθετα, στις τελικές δοκιμές, χρησιμοποιώντας το αλγοριθμικό μοντέλο Naïve Bayes, τα αποτελέσματα ήταν θετικά σε βαθμό που μας οδηγεί να απαντήσουμε θετικά και στο ερευνητικό μας ερώτημα, ότι η προπαγάνδα δηλαδή σε κάποιες τουλάχιστον μορφές της είναι δυνατόν να εντοπίζεται αυτόματα. Η πλειοψηφία των στοιχείων που λαμβάνονταν υπόψη για την ταξινόμηση σε κάθε δοκιμή και μάλιστα με εφαρμογή της ανάλυσης διαφορετικό κομμάτι των δεδομένων σχετίζονταν, σύμφωνα με το θεωρητικό πλαίσιο, με τουλάχιστον ένα από τα πέντε τεχνάσματα προπαγάνδας που θέσαμε προς ανάλυση. Το λογισμικό, που δημιουργήσαμε μέσω της γλώσσας προγραμματισμού Python και με τη χρήση του αλγοριθμικού αυτού μοντέλου, ταξινόμησε κείμενα σε προπαγανδιστικά και μη λαμβάνοντας υπόψη αυτά περίπου τα λεξιλογικά στοιχεία που θα λάμβανε κάποιος και κατά τη διάρκεια μίας ποιοτικής ανάλυσης. Σε αυτό το σημείο πρέπει να σημειώσουμε πως το ποσοστό ακρίβειας, ενώ υπήρξε σταθερά θετικό, ενδέχεται να μην έφτασε

μεγαλύτερα επίπεδα εξαιτίας παραγόντων που δεν αφορούν το πρόγραμμα. Τα λάθη στην ανάλυση δεν είναι απίθανα καθώς αυτή έγινε ποιοτικά. Η δυσαναλογία των δεδομένων σε επίπεδο θεματολογίας που αφορούσαν ή σε σχέση με τη χρονική περίοδο την οποία δημιουργήθηκαν (ως κείμενα) επίσης μπορεί να δημιουργήσει ασυνέπειες στην ταξινόμηση. Η ποσότητα των δεδομένων είναι επίσης ένα ζήτημα που μπορεί να επηρέασε τα κριτήρια ταξινόμησης. Παρ' όλα αυτά και υπό αυτές τις ατελείς συνθήκες τα κριτήρια της αυτόματης ταξινόμησης ήταν επαναλαμβανόμενα και σε όλες τις δοκιμές σύμφωνα με την θεωρία της ποιοτικής ανάλυσης.

Καταλήγοντας η απάντηση στο ερώτημα αν μπορεί να εντοπιστεί αυτοματοποιημένα προπαγάνδα, σε κάποιες τουλάχιστον μορφές της, μετά από ανάλυση μεγάλου όγκου δεδομένων είναι θετική. Η απάντηση στο πώς και σε ποιο βαθμό μπορεί να εντοπιστεί θεωρούμε πως αλλάζει διαρκώς και είναι διαφορετική ανάλογα με την κάθε προπαγανδιστική τεχνική και τέχνασμα.

Επιπρόσθετα πρέπει να πούμε ότι μέσω της διαδικασίας αυτής και άλλου είδους αυτοματοποιημένων αναλύσεων προπαγάνδας μπορεί να προκύψουν όροι σχετιζόμενοι με κάποια τεχνάσματα τους οποίους κάποιος αναλυτής προπαγανδιστικών τεχνικών δεν είχε συσχετίσει με την ανάλυση του και δεν τους λάμβανε υπόψη. Και το δικό μας λογισμικό δηλαδή και άλλα παρόμοια μπορούν να βοηθήσουν και να ενισχύσουν την ποιοτική έρευνα. Αυτή θεωρούμε άλλωστε πως είναι η πρωταρχική χρήση τέτοιων λογισμικών: Να αποτελέσουν ένα εργαλείο για τον αναλυτή ή την ομάδα αναλυτών και όχι να τους αντικαταστήσουν, κάτι που σε κάθε περίπτωση με τη σημερινή τεχνολογία είναι αδύνατο.

Στη διάρκεια της έρευνας εντοπίσαμε άμεση συσχέτιση συναισθημάτων και προπαγάνδας και μάλιστα επανειλημμένα αλλά δεν μπόρεσε να αξιοποιηθεί ιδανικά για την πρόβλεψη και την ταξινόμηση. Η επαναλαμβανόμενη διαγνωσμένη ως υψηλής σημαντικότητας συσχέτιση αυτή από τις δοκιμές με τον αλγόριθμο Random Forest Classifier μας έδειξε πως μπορούμε με σχετική ασφάλεια να βγάλουμε αυτό το συμπέρασμα.

Τέλος οι λέξεις που δίνονται από την αυτοματοποιημένη ανάλυση ως κριτήρια ταξινόμησης μαζί με τις λέξεις, που συμπεραίνεται κατά την ποιοτική ανάλυση πως αποτελούν κριτήρια ύπαρξης προπαγανδιστικής τεχνικής, είναι δυνατόν να οδηγήσουν στη δημιουργία λεξικών – εργαλείων που θα υποβοηθήσουν ίσως και καθοριστικά μελλοντικές προσπάθειες με τον ίδιο ή παρόμοιους σκοπούς.

12. Βελτιώσεις – Οδηγίες για μελλοντικά εγχειρήματα

Η έρευνα που πραγματοποιήθηκε, εκτός από τα συμπεράσματα στα οποία μας οδήγησε, έδωσε και πιθανές κατευθυντήριες γραμμές τόσο για τη βελτίωση του λογισμικού την οποία θεωρούμε απαραίτητη για να είναι λειτουργικό όσο και για μελλοντικές έρευνες και προσπάθειες πάνω στο ίδιο ή σε συναφές θέμα.

Το συμπέρασμα πως η ύπαρξη ορισμένων έστω προπαγανδιστικών τεχνικών σχετίζεται με το συμπέρασμα μας οδηγεί στην άποψη πως η εμπλοκή λεξικών αξιολόγησης συναισθήματος στην ανάλυση (έστω σε κάποιο στάδιο αυτής) ενδείκνυται. Επίσης, λόγω του ότι η ανάλυση περιορίστηκε σε λίγα μόνο τεχνάσματα προπαγάνδας σχετικά με το

εύρος των τεχνικών της και ακόμη και έτσι το εγχείρημα μας δεν είχε απόλυτη αλλά σχετική επιτυχία, όπως και αναμενόταν, θεωρούμε απαραίτητη τη χρήση ποιο εξελιγμένων μορφών ανάλυσης για αυτόν το σκοπό, με γνώμονα τη μέγιστη δυνατή αποτελεσματικότητα της ανάλυσης. Τα «νευρικά δίκτυα» (Neural Networks²⁴) ίσως και γενικότερα τεχνικές «βαθιάς μάθησης» (Deep Learning²⁵) είναι η ανάλυση που μπορεί να οδηγήσει στην ανάπτυξη τεχνητής νοημοσύνης που να εντοπίζει προπαγανδιστικά τεχνάσματα και με ποσοστό επιτυχίας σίγουρα μεγαλύτερο από αυτό που λάβαμε.

13. Χρησιμότητα του λογισμικού

Πιθανές χρήσεις του λογισμικού που δημιουργήσαμε, βελτιωμένων εκδόσεων του και άλλων παρόμοιων λογισμικών είναι ωφέλιμο να αναφερθούν.

Αρχικά μπορεί να χρησιμοποιηθεί ως εργαλείο που θα φιλτράρει μεγάλο όγκο δεδομένων αποτελούμενο από πολιτικές ειδήσεις ή πολιτικό λόγο εν γένει σε γραπτή μορφή και θα εντοπίζει τεχνικές προπαγάνδας, κατηγοριοποιώντας παράλληλα τα κείμενα που τις εμπεριέχουν σε μία ξεχωριστή κατηγορία. Κατόπιν, τα διαχωρισμένα πλέον προπαγανδιστικά κείμενα θα αναλύονται ποιοτικά. Η μέθοδος αυτή μπορεί να διευκολύνει και να ενισχύσει σε αξιοσημείωτο βαθμό προσπάθειες να αναλυθούν οι μέθοδοι προπαγάνδας που χρησιμοποιεί ένα συγκεκριμένο κόμμα ή ένα Μέσο ή πολλά Μέσα εναντίον ή υπέρ ενός προσώπου ή ενός κόμματος.

²⁴ https://en.wikipedia.org/wiki/Artificial_neural_network

²⁵ https://en.wikipedia.org/wiki/Deep_learning

Επικοινωνιολόγοι, πολιτικοί, δημοσιογράφοι, κρατικές υπηρεσίες ή ακόμη και στρατιωτικές υπηρεσίες ψυχολογικών επιχειρήσεων είναι οι κυριότερες ειδικότητες που θα μπορούσαν να χρησιμοποιήσουν αυτό το εργαλείο σε μία βελτιωμένη έκδοσή του για ποικίλους σκοπούς.

14. Αναφορές

Auerbach, J. & Castronovo, R., 2013. *The Oxford handbook of propaganda studies*. New York: Oxford University Press.

Bernays, E., 1928. *Propaganda*. New York: Liveright.

Central Intelligence Agency, 1969. *Propaganda Notes*, USA: CIA.

Cole, R., 1988. *The encyclopedia of propaganda*. New York: Sharpe.

Conroy, N. J., Rubin, V. L. & Chen., Y., 2015. *Automatic deception detection: Methods for finding fake news*. London, Ontario: American Society for Information Science.

Department of the US Army, 2008. *Soldier's Manual & Trainers Guide - Psychological Operation Specialist*. Washington, DC: Department of the US Army.

Ellul, J., 1965. *Propaganda: the formation of men's attitudes*. New York: Vintage Books.

Fairbanks, J., Fitch, N., Knauf, N. & Briscoe, E., 2018. *Credibility Assessment in the News : Do we need to read?*, New York, USA: s.n.

Fleming, C. A., 1995. Understanding propaganda from a general semantics perspective. " *ETC.: A Review of General Semantics*, pp. 3-13.

Floridi, L., 2009. Logical fallacies as informational shortcuts. *Synthese*, pp. 317-325.

González-Bailón, S. & Petchler, R., 2015. Automated Content Analysis of Online Political Communication. Στο: S. Coleman & D. Freelon, επιμ. *Handbook of Digital Politics*. London: Edward Elgar, pp. 433-450.

- Gulseth, H. L., 2004. *The Use of Propaganda in the Rwandan Genocide*. Oslo: Department of Political Science, University of Oslo.
- Herman, E. S. & Chomsky, N., 1994. A Propaganda Model. Στο: *Manufacturing consent : the political economy of the mass media*. London: Vintage, pp. 1-35.
- Holm, R. R., 2017. *Natural language processing of online propaganda as a means of passively monitoring an adversarial ideology*, United States: Naval Postgraduate School Monterey.
- Hutto, C. J. & Gilbert., E. E., 2014. *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. Eighth. International Conference on Weblogs and Social Media (ICWSM-14)., s.n.
- Institute for Propaganda Analysis, 1938. *Propaganda Analysis: Vol 1 of the Publications of the Institute for Propaganda Analysis, inc.*. 1st Edition επιμ. New York: Institute for Propaganda Analysis, inc..
- Lasswell, H. D., 1948. The structure and function of communication in society. *The communication of ideas*, Issue 37, pp. 215-228.
- Manning, C. D., Raghavan, P. & Schütze, H., 2008. Scoring, term weighting and the vector space model.. *Introduction to information retrieval*, 2(4), pp. 109-133.
- Manning, M. & Romerstein , H., 2004. *Historical dictionary of American propaganda*. Westport, Connecticut: Greenwood Publishing Group, Inc..
- Marlin, R., 2002. *Propaganda and the Ethics of Persuasion*. Peterborough, Ontario: Broadview Press, LTD..

- Merriam-Webster, 2018. *Merriam-Webster, Incorporated*. [Ηλεκτρονικό]
Available at: <https://www.merriam-webster.com/dictionary/propaganda>
[Πρόσβαση 21 Μάιος 2018].
- Nadeau, R., Cloutier, E. & Guay, J.-H., 1993. New evidence about the
existence of a bandwagon effect in the opinion formation process. *International
Political Science Review*, 14(2), pp. 203-213.
- Nelson, R. A., 1996. *A Chronology and Glossary of Propaganda in the United
States*. s.l.:Greenwood.
- O'Donnell, V. & Jowett, G. S., 2014. *Propaganda & persuasion*. 6th Edition επιμ.
California: Sage publications.
- Oxford University Press, χ.χ. *Oxford Living Dictionaries*. [Ηλεκτρονικό]
Available at: <https://en.oxforddictionaries.com/definition/propaganda>
[Πρόσβαση 27 12 2018].
- Pisarevskaya, D., 2017. Deception Detection in News Reports in the Russian
Language: Lexics and Discourse. *Proceedings of the 2017 EMNLP Workshop:
Natural Language Processing meets Journalism*, 7 September, p. 74–79.
- Pratkanis, A. & Aronson, E., 2001. *Age of Propaganda: The Everyday Use and
Abuse of Persuasion*. 2nd Edition επιμ. California, Santa Cruz: Macmillan.
- Prentice, S. και συν., 2011. Analyzing the semantic content and persuasive
composition of extremist media: A case study of texts produced during the
Gaza conflict. *Information Systems Frontiers*, 13(1), pp. 61-73.
- Rashkin, H. και συν., 2017. Truth of varying shades: Analyzing language in
fake news and political fact-checking.. *Proceedings of the 2017 Conference on
Empirical Methods in Natural Language Processing*, 7–11 September, p. 2931–
2937.

Ross, S. T., 2002. Understanding Propaganda: The Epistemic Merit Model and Its Application to Art. *Journal of Aesthetic Education*, Issue 36, pp. 16-30.

Rubin, V. L., Conroy, N. J., Chen, Y. & Cornwell, S., 2016.. Fake news or truth? using satirical cues to detect potentially misleading news. *Proceedings of the Second Workshop on Computational Approaches to Deception Detection*, 12-17 June, p. 7-17.

scikit-learn developers, 2007 - 2018. *scikit-learn*. [Ηλεκτρονικό]

Available at: [https://scikit-](https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html)

[learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html](https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html)

[Πρόσβαση 28 January 2019].

Shabo, M., 2008. *Techniques of propaganda and persuasion*. s.l.:Prestwick House Inc.

Silverman, H. I., 2011. Reuters: Principles Of Trust Or Propaganda?. *Journal of Applied Business Research*, November/December, pp. 93-115.

Sproule, J. M., 2001. Authorship and origins of the seven propaganda devices: A research note. *Rhetoric & Public Affairs*, pp. 135-143..

Sumanat, W. & Dispanya, E.-o., 2016. Propaganda Techniques in the Political Discourses of Aung San Suu Kyi. *ELT VIBES: International E-Journal for Research in ELT*, 10 Απριλίου, pp. 137-148.

Tacchini, E. και συν., 2017. *Some Like it Hoax: Automated Fake News Detection in Social Networks*, Santa Cruz: School of Engineering, University of California.

The Editors of Encyclopaedia Britannica, 2015. *Encyclopaedia Britannica*.

[Ηλεκτρονικό]

Available at: <https://www.britannica.com/topic/psychological-warfare>

[Πρόσβαση 6 1 2019].

Turing, A. M., 1936. *ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHIEDUNGSPROBLEM*. New Jersey, U.S.A.: Princeton University.

Turing, A. M., 1950. Computing Machinery and Intelligence. *Mind*, October, pp. 433-460.

US Army, 1979. Appendix I: PSYOP Techniques. Στο: J. Ronald, επιμ. *Psychological Operations Field Manual No.33-1*. Washington DC: Headquarters: Department of the Army, pp. 1-6.

US Army, 2005. *Field Manual 3-05.30*. Washington, DC: Department of the Army.

van Brakel, T. R., 2016. *Determining Automatic analysis of Populist Language Tweets in the Dutch Political spectrum*. Twente: s.n.

Γεωργιάδης, Γ. Κ., 1967. *Η Προπαγάνδα*. Αθήνα: Γεωργιάδης.

Киселёв, М., 2011. *Psyfactor*. [Ηλεκτρονικό]
Available at: <https://psyfactor.org/propaganda5.htm>
[Πρόσβαση 27 12 2018].

15. Παράρτημα

Phrases	Bandwagon Effect	Glittering Generalities	Name Calling	Testimonial/Appeal to Authority	Transfer
Convincing Cramer -- largely viewed as the best candidate they had to beat Heitkamp -- was an important part of a GOP effort to grow their slim Senate majority.	Yes	No	No	No	No
Her influence could also span beyond her in-person appearances. Onward Together, a group founded by Clinton aimed at "advancing the progressive vision that earned nearly 66 million votes in the last election," has been quietly giving out money to Democratic candidates and left-wing groups.	Yes	No	No	No	No
Italy's hard-line interior minister and deputy prime minister sparked outrage on Thursday after announcing a proposed amendment to his security decree that would force "little ethnic shops" to close by 9 p.m.	Yes	No	Yes	No	No
includes so many factors, and involves numerous steps involving upward of 100 people signing off on every single admit, that there is little room for bias against Asians to rear its head.	Yes	No	No	No	No
He is believed to have been killed by Saudi agents -- an allegation the nation has denied.	Yes	No	No	No	No
He's said to be afraid his presence would backfire, give Trump a foil, and energize the Republicans Democrats hope will stay home in November.	Yes	No	No	No	No
Saudi Arabia had first insisted that khashoggi left the consulate alive, but after international outcry teamed up with Turkish authorities to investigate the journalist's disappearance.	Yes	No	No	No	No
"This is the 13th week with no orders out of the Pacific Northwest and you know you want to be patriotic and you want to do the right thing but you wonder, why am I suffering so that Apple can grow bigger, so that tech can get the advantage of what we're doing," Heitkamp asked rhetorically.	No	Yes	No	No	Yes
He is playing the patriotic card, and even suggesting things could change after Election Day, despite having no proof at all that could happen.	No	Yes	No	No	Yes
"Mooching? My family never asked for free stuff, just a fair shot. We can't allow the American dream to be crushed by you, Stabenow and socialism.	No	Yes	No	No	No
The administration says "all options are on the table"	No	Yes	No	No	No
He said it would be a "fantastic thing" if the nations improved ties, but warned "it might be just the opposite".	No	Yes	No	No	No
make America great again	No	Yes	No	No	Yes

With his trade policies, Donald Trump has embarked on bringing about the biggest change to the way the US does business with the rest of the world for decades.	No	Yes	No	No	No
Why is the left freaked out by the NPC meme? Because it's funny, effective and mostly true.	No	No	Yes	No	No
President Donald Trump said on Saturday the release of U.S. pastor Andrew Brunson after two years in Turkey's custody was a "tremendous step" toward improved relations with Turkey, but he denied cutting a deal with Ankara.	No	Yes	No	No	No
We feel much differently about Turkey today than we did yesterday, and I think we have a chance of really becoming much closer to Turkey	No	Yes	No	No	No
At a rally in October, Trump said that he and Kim "fell in love" over some "beautiful" and "great" letters the North Korean leader had sent to the White House.	No	Yes	No	No	No
"You can vote for Democrat mob rule or you can vote for a Republican party that stands proudly for law and order, fairness, freedom and justice."	No	Yes	No	No	Yes
It is important for the long-term strength of the institution to have the resources we need	No	Yes	No	No	No
This is the right thing for the government to do	No	Yes	No	No	No
The Conspiracy has a strategic goal, which continues to this day, to sow division and discord in the U.S. political system ... including the upcoming 2018 midterm election.	No	Yes	No	No	No
Federal prosecutors have charged a Russian woman with conspiracy to defraud the United States	No	No	No	No	Yes
He just goes along — even when you get a sense that he knows it's not right.	No	Yes	No	No	No
"When you hear all this talk about 'economic miracles' right now, remember who started it," Obama said.	No	Yes	No	No	Yes
Obama and Democrats here are up against a Republican campaign in Nevada that embodies the surprisingly durable politics of Trump's party.	No	Yes	No	No	Yes
That is not how America works. That is how some tin-pot dictatorship works	No	No	Yes	No	No
The recruiting program was halted in 2016 over national security concerns and then shuttered, but thousands who had already signed enlistment contracts remained in the pipeline, slated for extra vetting that the Pentagon has said it does not have the manpower to do.	No	Yes	No	No	No
Their races, and others like them in Midwest states Trump won, are testing the power of the Trump coalition ahead of 2020, especially on behalf of candidates not named Trump.	No	No	No	No	No
"He's a special man," Trump said. "He's a total winner."	No	Yes	No	No	No

Still, Republicans have been surprised by the inability of the two candidates to make it a closer fight.	No	No	No	No	No
"There is the presumption of innocence until proven guilty. In this case, people have turned it upside down, unfortunately."	No	Yes	No	No	No
The European Union has also said the trade parts of the proposals are unacceptable.	No	No	Yes	No	No
. Mrs May has so far stuck by the plan, saying it is a workable one which takes on board both the UK and EU's red lines, is the best possible one for the UK and EU economies and avoids the need for a visible border on the island of Ireland.	Yes	No	No	No	No
"I don't think you're going to find an overwhelming number of folk, African-Americans in particular, carrying the flag for John James," Anthony said. "John James has presented himself in a way that is somewhat militaristic and Trumpy-ish."	No	No	Yes	No	Yes
In a Facebook post on Monday, Shapiro described her intention to bring her children to "birddog" in the Senate hallways and basement to protest the "horrible atrocious men" there.	No	No	Yes	No	No
Matteo Salvini, leader of Italy's far-right League party	No	No	Yes	No	Yes
foreign-owned grocery shops have become "meeting places for drunkards, pushers, hell-raisers."	No	No	Yes	No	Yes
It is not an initiative against foreign shops, but it is to limit the abuses and irregularities of certain shops -- almost all of which are managed by foreign citizens -- which have become havens for people who raise hell	No	Yes	Yes	No	Yes
He is currently in the UK, from where he once again dismissed the investigation as a "rigged witch hunt".	No	No	Yes	No	Yes
The explosive publication of an unverified dossier alleging that Russia holds compromising material on Mr Trump has also raised prickly questions for him. He has batted them away, dismissing the allegations as "fake news".	No	Yes	Yes	No	Yes
The missile strike was the first time the US had directly targeted the Syrian regime since the conflict began, and viewed as a breathtaking policy shift for a previously isolationist leader.	Yes	No	Yes	No	Yes
Scepticism over Trump's candidacy stemmed not only from his controversial platform on immigration and outrageous campaign style, but from his celebrity past.	Yes	Yes	Yes	No	Yes
they fear a repeat of their crushing loss in the 2016 presidential election.	No	No	No	No	Yes
As we've reported, the left's violent rhetoric against Republicans has continued to ramp up ahead of the midterm elections	No	Yes	Yes	No	No
Leftist Official Suspended Over Calling For Republican Beheadings With Guillotines	No	No	Yes	No	Yes

Uninformed leftists scurry confronted with reality	No	No	Yes	No	Yes
Liberal Mob Gets Defused & Red Pilled	No	No	No	No	No
French President makes more bizarre comments on national identity	No	Yes	No	No	Yes
Macron's globalist rhetoric hasn't done much for his approval ratings, which have been hovering at around 29 per cent – an all time low.	Yes	No	Yes	No	Yes
Benson said that Trump should take a DNA test to see if he has more native American in him than Warren, following her disastrous PR stunt.	No	No	No	No	Yes
"She's appealing, ironically, to tribal politics." he added.	No	No	No	No	Yes
"They become straight-up, typical lifetime politicians that felt that, 'I'm in power, and I'm on the top of the mountain, and I'm gonna look down, and I'm gonna expect everybody to continue to elevate me.'"	No	Yes	Yes	No	No
But what's really striking about his candidacy is how much he's redefined what it means to be a populist candidate raging against the status quo in 2018.	No	No	Yes	No	Yes
That prompted the national media to descend on West Virginia right after the election, eager to explain the archetype of this mythical white working-class voter who had swung so far to the right.	No	No	Yes	No	No
The screener then checks one of three boxes, indicating whether the recruit presents a major, moderate, or minor risk based on those factors.	No	No	No	No	No
their best efforts to present themselves as Trump-like candidates aren't lifting them up high enough to win two Democratic Senate seats.	No	No	Yes	No	Yes
Barletta and Renacci, both loyal Trump allies and, like him, wealthy businessmen turned politicians, are down by double digits against Democratic incumbents. In other Midwest states that Trump narrowly won, including Michigan and Wisconsin, Republican candidates running statewide are struggling as well.	Yes	No	Yes	No	Yes
Also, Paul said, "I think we really need to discontinue our arms sales to Saudi Arabia and have a long and serious discussion about whether or not they want to be an ally, or they want to be an enemy."	No	No	Yes	No	Yes
They were energized by the president's presence in their town, seeing it as a sign that Republicans, for the first time in decades, were scared of losing the seat.	No	No	No	No	No
But when they were asked about their support for Democratic Senate candidate Rep. Kyrsten Sinema, many responded with a sigh.	Yes	No	No	No	No

Nearly all of the 14 young, progressive, Latinx voters and organizers who spoke with BuzzFeed News — most of whom were involved with immigration activist organizations — acknowledged that Sinema’s record on this highly personal issue was not ideal.	Yes	Yes	No	No	Yes
To create a false equivalency there is absurd and dangerous	No	No	Yes	No	No
Sinema is a Blue Dog Democrat.	No	No	Yes	No	Yes
The groups have spent millions of dollars on ads and organized thousands of their supporters on her behalf throughout the campaign.	No	No	No	No	No
she was by far the strongest candidate in the race	No	Yes	No	No	Yes
We’ve been proud to defend people’s health and rights alongside Kyrsten in the House and look forward to continuing that work together in the Senate	No	Yes	No	No	Yes
They told BuzzFeed News they supported Abboud in the primaries, a young Muslim woman with an anti-establishment, pro-immigrant platform, who in many ways fits the trend of young, minority, progressive candidates driving the “blue wave” in other states.	No	No	Yes	No	Yes
She’s part of the minority group, she supports immigrants, she supports reproductive rights, she wanted health care for all, so I believe that we needed someone new, a new face to represent the minority group she’s a part of	No	Yes	No	No	Yes
When they tell you that they almost would have preferred Flake running for Senate as opposed to Sinema?	No	No	No	No	No
Sinema also voted to greatly increase the length and intensity of the vetting process for Syrian and Iraqi asylum-seekers by requiring that the federal government certify that they are not a threat through background investigations before they are allowed to enter the US.	No	No	No	No	No
Just in October, 659 new members joined their canvassing effort and completed over 200 shifts.	No	No	No	No	Yes
So instead of focusing on the negative, the volunteers have chosen to emphasize Sinema’s liberal and loyal stances on abortion, universal health care, and LGBT issues.	No	Yes	No	No	Yes
And maybe in a more friendly Congress, some of the volunteers said, her liberal side will start to take over.	No	Yes	No	No	Yes
Angelica Romero is the vice president of Planned Parenthood Generation’s Arizona State University chapter. An earlier version of this post misstated her title.	No	No	No	No	No

Polls closed in more than half of the 50 U.S. states on Tuesday after a divisive election campaign that will shape the future of Donald Trump's presidency, and most of the key races that will decide control of the U.S. Congress were still too close to call.	No	No	No	No	No
Republican Mike Braun opened an early lead in the crucial Senate showdown in Indiana with incumbent Democrat Joe Donnelly, but the race was too close to call with just 20 percent of the votes in.	No	No	No	No	No
Tightly contested Senate races in Florida and West Virginia also were too close to call, as were high-profile races for governor in Ohio and Georgia.	No	No	No	No	No
A full picture of the voting results will not emerge for hours	No	No	No	No	No
The first national elections since Trump captured the White House in a 2016 upset became a referendum on the polarizing president	No	No	Yes	No	Yes
All 435 seats in the House, 35 seats in the 100-member Senate and 36 of the 50 state governorships were up for grabs.	No	No	No	No	No
If Democrats capture the House, they could block Trump's policy agenda and launch congressional investigations into Trump's administration	No	No	No	No	No
A Republican victory in both chambers of Congress would be a validation for Trump's polarizing style	No	No	Yes	No	Yes
This year is a big date for the Main Directorate of the General Staff of the Armed Forces of the Russian Federation.	No	Yes	No	No	Yes
The Russian president suggested returning the historic name to the military intelligence agency	No	No	No	No	Yes
On Friday, Russian military intelligence officers were celebrating their professional holiday.	No	No	No	No	No
No less important is the fact that we have dealt a blow to the terrorists and are preventing their return to our territory.	No	Yes	Yes	No	Yes
It stands accused of being behind a number of high-profile missions over the past years, including attempting a coup in Montenegro, stealing private emails of the Democratic party and the Hillary Clinton campaign, unleashing computer malware on Ukraine and, of course, attempting to kill a former double agent with chemical weapons in Salisbury.	No	No	No	No	Yes
Seventeen Republican incumbents are defending House seats in districts that Democrat Hillary Clinton won in the 2016 presidential election.	No	No	No	No	No
Some 182 Republican incumbents are defending House seats in districts Trump won in 2016.	No	No	No	No	No

Nine Republican incumbents are defending House seats in districts where Trump's margin of victory in 2016 was less than 5 percentage points.	No	No	No	No	No
A total of 155 Republican incumbents are defending House seats in districts Trump won by double digits in 2016.	No	No	No	No	No
Democrats are defending 10 Senate seats in states that Trump won in 2016.	No	No	No	No	No
Five of those races are widely seen as competitive or leaning Republican.	Yes	No	No	No	No
The poll, conducted online on Tuesday, was based on responses from 36,740 people who voted in 37 states. The poll is ongoing and will be updated as people continue to vote.	No	No	No	No	No
Trump and Republicans touted the strong economy during the campaign	No	Yes	No	No	Yes
Trump also played up his hardline positions on immigration in the final stretch, a strategy aimed at driving his base to turn out.	No	No	Yes	No	Yes
Democrats made healthcare and protecting the Affordable Care Act, popularly known as Obamacare, the central theme of their House and Senate campaigns, warning that people could lose coverage for pre-existing health conditions and other protections if Republicans kept control of Congress.	No	Yes	No	No	Yes
Democrats have been widely expected to pick up the seats needed to wrest control of the chamber from President Donald Trump's Republicans, who were expected to maintain control of the Senate.	Yes	No	No	No	No
It did not project the margin of victory or elaborate immediately on the basis for its projection.	No	No	No	No	No
wage increases are no small thing	No	Yes	No	No	No
For millions, it means getting ahead when for so long they were barely getting by.	Yes	Yes	No	No	No
there is a party that looks at these positive changes with disdain.	No	Yes	No	No	Yes
Comparing policies that helped families save money to rape and the end of the world is not only offensive, it's insulting to the American people's intelligence.	No	Yes	Yes	No	Yes
At the end of the day, it comes down to two visions for this country.	No	Yes	No	No	No
Republicans have a vision for a booming economy, safer communities, and a stronger military, and policies that empower Americans to make the best decisions for themselves, their businesses, and their families.	No	Yes	No	No	Yes
Democrats have a goal of government-dictated programs and policies that will limit choices, bust budgets, and leave middle-class families behind.	No	Yes	No	No	Yes
Why? Because Democrats are summoning spirits of two years past.	No	Yes	No	No	Yes

Two years ago, our country was in a very different place. Our economy was stagnant, and we were told to accept the sluggish two percent growth as the 'new normal.'	No	Yes	Yes	No	No
we were told to accept	No	Yes	No	No	No
Thanks to an absurdly high corporate tax rate, we hemorrhaged companies (and jobs) to other countries, and small businesses credited regulations and the tax code with their inability to grow and hire. Republicans changed that.	No	Yes	No	No	Yes
It means Republicans voted to protect the American people	No	Yes	No	No	Yes
Republicans believe that everyday Americans know better than the federal government how to live their lives, run their businesses, and raise their families.	No	Yes	No	No	Yes
The Democrats? Well, the vast majority did vote NO on all of these regulatory rollbacks....	No	No	Yes	No	Yes
What Americans do have Democrats to thank for is the destruction of community banks across the country.	No	Yes	No	No	Yes
The goal was to protect Americans from bailouts, financial fraud, and abusive banking.	No	Yes	No	No	No
We've all heard it	Yes	Yes	No	No	No
Congress doesn't do anything.	No	Yes	No	No	Yes
Community banks and credit unions don't have the resources to navigate excessive regulations	No	No	No	No	No
it had gotten so bad by the time President Trump came into office that we were losing an average of one community bank or credit union every single day.	No	Yes	Yes	No	Yes
It also created a rogue agency	No	No	Yes	No	Yes
Democrats value the positive contributions of everyone	No	Yes	No	No	Yes
disabled Americans have been able to more fully participate	No	No	No	No	No
the Democratic Party is working to expand equal opportunities	No	Yes	No	No	Yes
We support one fair wage for all workers by ending the sub-minimum wage for people with disabilities under Section 14(c) of the Fair Labor Standards Act.	No	Yes	No	No	No
Democrats believe all children should be able to lead happy, successful lives.	No	Yes	No	No	Yes
President Franklin Roosevelt enacted the G.I. Bill	No	No	No	No	No
In 2010, President Obama signed into law student loan reform	No	No	No	No	No
We must prepare the next generation for success in college and the workforce.	No	Yes	No	No	No
The Department of Transportation and the EPA issued new fuel-economy standards	No	No	No	No	No
We know that climate change is one of our nation's greatest challenges	No	Yes	No	No	No
Democrats will fight to protect our planet for future generations.	No	Yes	No	No	Yes

Donald Trump, on the other hand, has called climate change "a hoax invented by the Chinese," and withdrawn from the Paris Climate Agreement.	No	No	Yes	No	Yes
In a country as great as ours, no child should be afraid to go to school or walk around their neighborhood.	No	Yes	No	No	Yes
World Patent Marketing was founded in 2014	No	No	No	No	No
When prospective customers left their contact information on the company website, an employee would call them back and follow a script	No	No	No	No	No
In joining the board, Mr. Whitaker was quoted in a news release issued by the company as saying that he was honored to be a part of World Patent Marketing	No	No	No	No	No
Mr. Whitaker, using his Iowa law firm's email, told a man who had complained to Mr. Cooper that he was a former federal prosecutor and served on the company's board.	No	No	No	No	No
Record Number of Migrants in Families Were Apprehended at Border Last Month	No	No	No	No	No
The number of migrants traveling in families who were apprehended at the southwest border surged past 20,000 last month	No	No	No	No	No
The 23,121 people traveling in families who either surrendered or were arrested by Border Patrol agents represent a record over a one-month period.	No	No	No	No	No
I can tell you Matt Whitaker's a great guy	No	Yes	No	No	Yes
Mr. Whitaker attended meetings at the White House; in some cases, the president and the attorney general were both there	No	No	No	No	No
The witness, Sam Clovis, is a former Trump campaign aide and a former candidate for Iowa state treasurer.	No	No	No	No	Yes
Senator Chuck Schumer of New York, the chamber's top Democrat, sent a letter to Mr. Trump on Friday	No	No	No	No	No
The group Protect Democracy filed the lawsuit for the plaintiffs.	No	No	No	No	No
Jeff Sessions' Temporary Replacement Has Repeatedly Criticized Robert Mueller's Investigation	No	No	No	No	No
If Whitaker did decide to fire Mueller, he would have to tell Mueller in writing and detail his specific reasons for doing so, under the regulations.	No	No	No	No	No
Whitaker released a statement Wednesday night praising Sessions as a "dedicated public servant" and saying that it was a "true honor that the President has confidence in my ability to lead the Department of Justice as Acting Attorney General."	No	Yes	No	No	Yes
I am committed to leading a fair Department with the highest ethical standards, that upholds the rule of law, and seeks justice for all Americans.	No	Yes	No	No	Yes

Whitaker was rumored as a leading contender for several top legal jobs in the Trump administration in recent months.	Yes	Yes	No	No	Yes
Democrat Kyrsten Sinema wins Arizona Senate seat	No	No	No	No	No
Trump calls on authorities to stop Florida election recounts	No	No	No	No	No
Environmental damage prompts a rethink among US coastal voters	No	Yes	No	No	No
Donald Trump and the global assault on press freedom	No	Yes	No	No	Yes
Perhaps most famously, Sessions was the architect of the "zero tolerance" border policy that led to family separations	Yes	No	Yes	No	Yes
Sessions was as xenophobic	No	No	Yes	No	Yes
US judge blocks construction of Keystone oil pipeline	No	No	No	No	No
Alexandria Ocasio-Cortez campaigned on Medicare for All and abolishing ICE.	No	No	No	No	No
Early in the speech she gave across northeast Iowa, in union halls and living rooms and farmyards, Abby Finkenauer got the thing on many listeners' minds out of the way: She made a joke about her age.	No	No	No	No	No
In all of the most obvious ways, Ocasio-Cortez and Finkenauer couldn't be more different from one another.	Yes	No	No	No	No
Ocasio-Cortez is a Latina from one of the country's most liberal districts, a democratic socialist who beat back the establishment with promises of Medicare for All and abolishing ICE.	No	Yes	No	No	Yes
A few weeks before her stunning primary upset in June, Ocasio-Cortez spoke to a tiny group of voters on a rooftop in Jackson Heights, Queens, the roar of traffic behind her.	No	No	No	No	No
"It doesn't take 100 years to do this," she says in the ad. "It takes political courage."	No	Yes	No	No	No
When she spoke at fundraisers and union hall picnics, Finkenauer cited a single moment as the one where she decided, definitively, to run for Congress.	No	No	No	No	No
In their campaigns against older, wealthier men, both Finkenauer and Ocasio-Cortez leaned heavily on their working-class identities.	No	No	No	No	No
The Democrats' lieutenant governor candidate, Rita Hart, stopped into the union office, then ducked quickly out again after her speech. Finkenauer, in a plaid shirt and quilted vest, stayed. Two Department of Transportation workers came to tell her they had been in the state capital with Finkenauer during the vote over bargaining rights.	No	No	No	No	No
A retired four-star general is President Donald Trump's pick to be U.S. ambassador to Saudi Arabia	No	No	No	No	No

Trump announced Tuesday that he is nominating John Abizaid, the longest-serving commander of the U.S. Central Command, to lead the U.S. Embassy in Riyadh. It's a post that has been empty since former ambassador Joseph Westphal left in January 2017.	No	No	No	No	No
If confirmed by the Senate, Abizaid would become ambassador as the Trump administration is weighing the U.S. response to the killing of Washington Post contributor Jamal Khashoggi at the Saudi Consulate in Istanbul.	No	No	No	No	No
Abizaid, who retired in 2007, served in wars in Grenada, the Persian Gulf, Bosnia, Kosovo, Afghanistan and Iraq.	No	No	No	No	No
The number of Republican women elected to the House dropped from 23 to 13 after the election.	No	No	No	No	No
Elections Wednesday will also determine party leadership in the Senate.	No	No	No	No	No
Democrats press the case that President Donald Trump violated the law and Constitution by making Whitaker the country's chief law enforcement officer.	No	No	No	No	Yes
The rest of the GOP line-up below the Kentucky Republican is expected to shuffle slightly. Sen. John Cornyn of Texas, the GOP whip, is being forced out by term limits. That allows Sen. John Thune of South Dakota to move up to the No. 2 spot. Sen. John Barrasso of Wyoming and Sen.	No	No	No	No	No
Congressional Democrats, meanwhile, have called the appointment unconstitutional and demanded that Whitaker recuse himself from overseeing the Mueller investigation.	No	No	Yes	No	Yes
House Minority Leader Nancy Pelosi is challenging the Democrats who are fighting to stop her from becoming speaker to put forward an alternative.	No	No	No	No	No
Republican Ohio Rep. Jim Jordan and California Rep. Kevin McCarthy are running for House minority leader.	No	No	No	No	No
They say untrained experts should not be allowed to decide	No	No	Yes	No	Yes
Tallying machines in Palm Beach County overheated and caused mismatched results with the recount of 174,000 early voting ballots. The machine recount is supposed to be finished by Thursday.	No	No	No	No	No
Supervisor of Elections Susan Bucher said that the machines are "old" and were not designed to work around the clock.	No	No	No	No	No
Rick Scott is Florida's Republican governor, not a senator-elect.	No	No	No	No	No

The Special Counsel Independence and Integrity Act would make it harder for President Donald Trump to undermine the investigation into Russian meddling and any coordination between Moscow and the Trump campaign.	No	No	No	No	Yes
Trump has repeatedly dismissed Special Counsel Robert Mueller's investigation as a political witch hunt.	No	No	Yes	No	Yes
McConnell came to the Senate floor himself to prevent Flake's bill from moving forward.	No	No	No	No	No
Trump also skipped out of a Veterans Day wreath laying ceremony at Arlington Cemetery in Virginia, across the river from the White House, though VA Secretary Robert Wilkie was there. The President told Fox News on Sunday that he had made a mistake in skipping the event.	No	No	No	No	No
Both Green and other Agriculture Department officials declined to answer questions about the incident, according to the paper.	No	No	No	No	No
The investigation details the creation and eventual removal of the policy, which sought to "ensure LGBT members felt protected by their local 4-H program," by asking the organization to "treat all students consistent with their gender identity and allow them 'equal access,'" according to the paper. It also details the role of Heidi Green, the former chief of staff for Secretary of Agriculture Sonny Perdue, in requesting the policy be rescinded within several days of its publication, the paper said.	No	No	No	No	No
Trump's misleading claim that he warned about Osama bin Laden	No	No	Yes	No	Yes
Mr. Trump's claim that he delivered a prescient warning about Bin Laden is hyperbolic.	No	No	Yes	No	Yes
Polls reflect that the same Rust Belt, blue-collar voters that propelled Trump over Hillary Clinton are not showing the same kind of enthusiasm for those candidates, even after Trump's repeated visits and encouragement.	No	No	Yes	Yes	No
The president's job approval in both states has been underwater, with 53% of those surveyed disapproving in Pennsylvania and 54% in Ohio, based on the most recent polls available.	Yes	No	No	Yes	No
According to the affidavit, the conspiracy with which Khusyaynova is charged "has a strategic goal, which continues to this day, to sow division and discord in the US political system, including by creating social and political polarization, undermining faith in democratic institutions, and influencing US election, including the upcoming 2018 midterm election.	No	Yes	No	Yes	Yes

"If we can't find a way to activate the largest and fastest growing demographic in this country, there's no way we can take our country back," the actress America Ferrera told the crowd.	No	Yes	No	Yes	No
"SHE follows the lawsuit by the MAVNIS and SHE feels like the MAVNIS are being harmed because they are described as a threat to National Security," reads the counterintelligence assessment report shared with BuzzFeed News	No	Yes	Yes	Yes	Yes
This isn't what happened, of course, but make no mistake, what the US Department of Justice is alleging - a cyber-attack by Russian military officers against a US political party and its leaders - is, or at least should be, equally inflammatory.	No	No	Yes	Yes	Yes
Project Lakhta, per the affidavit, involves groups like the Internet Research Agency and Concord Management and Consulting. Yevgeniy Viktorovich Prigozhin, known as "Putin's Chef," controls Concord, the affidavit details, and "spent significant funds to further" Project Lakhta.	No	No	Yes	Yes	Yes
Polls reflect that the same Rust Belt, blue-collar voters that propelled Trump over Hillary Clinton are not showing the same kind of enthusiasm for those candidates	No	Yes	Yes	Yes	No
In the weeks since his disappearance, Turkish media reported that Khashoggi was tortured, killed, and eventually dismembered inside the Saudi Consulate in Istanbul.	No	No	No	Yes	Yes
Reports by the Turkish government also indicate that a team of 15 assassins, reportedly equipped with a bone saw, entered and left the country in a short time to carry out the assassination operation.	No	No	Yes	Yes	No
Saudi Arabia's foreign minister said Sunday that the death of newspaper columnist Jamal Khashoggi was the result of a "rogue operation," and blamed the government's initial denial of his death on a "cover-up."	No	No	Yes	Yes	Yes
Mr Trump claimed his tough talk was causing the "money to pour in", although analysts point out that countries were already increasing their contributions under a 2014 agreement.	No	No	No	Yes	No
However, he has rowed back on a campaign promise to label China a currency manipulator - a move that experts warned could have provoked a trade war.	No	No	Yes	Yes	Yes
But environmental groups warned that undoing the regulations would have serious consequences at home and abroad.	No	Yes	No	Yes	No
based on the most recent polls available.	No	No	No	Yes	No

His reelection team will be keeping a close eye on results in those two races. Win or lose, their performance will likely offer Trump's campaign data crucial to figuring out how to maintain and galvanize its 2016 coalition of voters, according to two sources in Trump's orbit.	No	No	No	No	No
HRC Arizona has been bold and unequivocal in its support for her bid to be our next Senator.	No	Yes	No	Yes	Yes
A spokesperson for EMILY's List told BuzzFeed News in a statement that it endorsed Sinema	No	No	No	No	No
The Democrats are favored by election forecasters to pick up the 23 seats they need to gain a majority in the House	No	No	No	No	No
opinion polls show they have slimmer hopes of picking up the two seats they need to gain control of the Senate.	No	No	No	Yes	No
Trump accused Democrats without offering any evidence of "openly encouraging millions of illegal aliens to break our laws, violate our borders and overrun our country."	No	No	Yes	No	Yes
Russian President Vladimir Putin suggested the internationally-recognized old name of the agency should be returned.	Yes	Yes	No	Yes	Yes
Putin praised the military intelligence for the role they played in the past few years in Syria, where the Russian military was assisting the government in fighting jihadist groups.	No	No	Yes	No	Yes
The litany of accusations is denied by Moscow and seems to be ignored by the Russian people. A recent opinion poll about the Novichok poisoning affair in Britain showed that while almost 80 percent of Russians knew about it, only 7 percent believed that the Russian intelligence was involved and just 4 said it made them think worse of the spooks.	Yes	No	No	Yes	Yes
Just 42 percent of registered voters cast ballots in the 2014 congressional midterm elections, the lowest turnout since at least 1964, according to the U.S. Census Bureau.	Yes	No	No	Yes	Yes
Republicans now hold 33 governorships and 65 legislative chambers, with Democrats controlling 31 and two being tied, according to the nonpartisan National Conference of State Legislatures (NCSL)	No	No	No	Yes	No
By the NCSL's record, Republican dominance of state legislatures is at a historic high.	No	Yes	No	Yes	Yes
U.S. voters were deeply divided about Republican President Donald Trump's job performance and the direction of the country as they cast ballots in Tuesday's congressional midterm elections, according to the Reuters/ipsos Election Day poll.	No	No	No	Yes	No
Election forecasters said Democrats would likely pick up the 23 seats they need to gain a majority in the House	Yes	No	No	Yes	No
Immigration, healthcare and the economy ranked as the top issues for voters, the Reuters/ipsos poll found.	No	Yes	No	Yes	No

The Reuters/Ipsos poll found robust support for changing the nation's gun laws. Seven in 10 voters said they wanted "moderate" or "strong" regulations and restrictions for firearms, the poll found.	Yes	Yes	No	Yes	Yes
Fox News Channel projected on Tuesday that Democrats will win control of the U.S. House of Representatives.	No	No	No	Yes	Yes
The truth is, House Republicans have been hard at work delivering on the American people's priorities.	No	Yes	No	No	Yes
The number of migrants traveling in families who were apprehended at the southwest border surged past 20,000 last month, according to new data released on Friday by the Customs and Border Protection agency.	No	No	No	No	No
Mr. Whitaker attended meetings at the White House; in some cases, the president and the attorney general were both there, officials who were also present said.	No	No	No	Yes	No
Later that month, the New York Times reported that Whitaker would have taken over for Rosenstein as acting deputy attorney general if Rosenstein resigned following a Times report that he previously questioned Trump's fitness as president.	No	Yes	No	Yes	Yes
Sessions cranked through conservative achievements on immigration, drugs, crime, and civil rights reversals that fueled Trump's base	No	Yes	No	No	Yes
Turkish officials claim Khashoggi was killed by a 15-member assassination squad sent from Riyadh on orders from the highest levels of the Saudi government.	No	No	No	Yes	Yes
Analysts say the House Republican side of the aisle will be made up of 90 percent white men in the new Congress.	No	No	No	Yes	Yes
CNN's exit polls show that 60% of veterans backed Trump in 2016 and 63% supported Republicans in that year's congressional races, while 36% backed Democrats. In 2018, the percentage of veterans backing Republicans fell to 58%, while 41% gave their votes to Democrats.	Yes	No	No	Yes	Yes
The day to day business of the U.S. House of Representatives isn't always the flashiest of headlines, so we can't blame everyday Americans for not knowing every bill and every vote.	No	Yes	No	No	Yes
But exit polls showed that 38% of people who cast a vote in House races did so to oppose Trump, compared to 26% who did so to support him.	Yes	No	No	Yes	Yes
I recognize it's boring to go out every single day and remind everybody that unemployment is below 4%	No	Yes	No	No	Yes

Papadopoulos arrived at a minimum-security camp in Oxford, Wisconsin, according to the Federal Bureau of Prisons.	No	No	No	No	No
He was sentenced in September for lying to the FBI about his interactions with Russian intermediaries during the 2016 presidential campaign.	No	No	No	No	No
The former Trump campaign chairman's cooperation agreement with special counsel Robert Mueller unraveled, prosecutors say, after he repeatedly lied to investigators, an allegation Manafort denies.	No	No	No	No	No
In exchange, prosecutors agreed that if he provided "substantial assistance" they could help him get a lighter sentence.	No	No	No	No	No
That probe is examining the conduct of Democratic lobbyist Tony Podesta, former Obama White House counsel Greg Craig and former Republican congressman and lobbyist Vin Weber. None of the three has been charged.	No	No	No	No	No
writer Jonathan Lemire contributed to this report.	No	No	No	No	No
At a roundtable in Mississippi later Monday, Trump seemed to acknowledge that children were affected.	No	No	No	No	Yes
"The agents on scene, in their professional judgment, made the decision to address those assaults using less lethal devices," McAleenan told reporters.	No	Yes	No	No	Yes
Mexico's Interior Ministry said in a statement it would immediately deport the people arrested on its side of the border and would reinforce security.	No	No	No	No	No
a decision quickly met with derision by critics on social media.	Yes	No	No	No	Yes
Mexico's government said on Tuesday it would bestow the country's top honor for foreigners on Jared Kushner, U.S. President Donald Trump's son-in-law	No	No	No	No	No
Cox, who trailed in the vote tally until Monday, increased his lead to 529 votes out of about 113,000 cast, after updated numbers were released by Fresno, Kings and Tulare counties.	No	No	No	No	No
That dynamic played out in Orange County, where the last four Republican seats in what was once a GOP stronghold went to the Democrats.	Yes	No	No	No	Yes
Ivanka Trump defended her use of a private email account as she was moving into an adviser's position in her father's administration, saying that it cannot be compared to the flap over former Secretary of State Hillary Clinton's private email server	No	No	No	No	No
All of my emails are stored and preserved. There were no deletions	No	No	No	No	No
There is no restriction of using personal email	No	No	No	No	No

In fact, we're instructed that if we receive an email to our personal account that could relate to government work, you simply just forward it to your government account so it can be archived.	No	No	No	No	No
Clinton used a personal email account linked to a private server at her home in Chappaqua, a New York City suburb, during her time as the top U.S. diplomat under President Barack Obama.	No	No	No	No	No
Clinton deleted thousands of emails that she and her lawyers decided were personal or unrelated to her work as secretary of state before she turned over thousands of other emails to federal investigators.	No	No	No	No	No
Ivanka Trump said she is not worried about legal exposure for herself, her father or anyone else in her family regarding special counsel Robert Mueller's probe into Russian interference in the 2016 elections.	No	No	No	No	No
I know the facts as they relate to me and my family, and so I have nothing to be concerned about	No	Yes	No	No	Yes
Michael has been a great advocate in many ways. I'm tremendously grateful to him for aggressively representing me in my fight to regain my voice. But in other ways Michael has not treated me with the respect and deference an attorney should show to a client.	No	Yes	No	No	Yes
The pornographic film actress Stephanie Clifford moved on Wednesday to distance herself from the lawyer Michael Avenatti, saying he had filed a defamation suit against President Trump against her wishes and sometimes spoke on her behalf without her consent.	No	No	No	No	No
Two Iranians were behind the ransomware attack that crippled Atlanta's government for days this year, the Justice Department said	No	No	No	Yes	Yes
Cybersecurity researchers said they were surprised to learn that the SamSam group was based in Iran.	No	No	No	Yes	Yes
The Treasury Department also imposed sanctions on two other Iranians, accusing them of changing the Bitcoin obtained by the hackers into Iranian rial.	No	No	No	No	No
Such an outcome is not certain.	No	No	No	No	No
There's a good possibility that we can make a deal, and he is open to it	No	No	No	No	No
U.S. intelligence agencies have concluded that the crown prince must have at least known about the plot to kill Khashoggi.	No	No	No	No	No
The two leaders are meeting at the Group of 20 rich and developing nations in Buenos Aires, Argentina. The meeting is closed to reporters.	No	No	No	No	No
Trade analysts and administration officials acknowledge it won't be easy.	No	No	No	Yes	No

The U.S. supported the language on the WTO but did not sign on to a commitment to support the Paris Climate Accord, which Trump is exiting.	No	No	No	No	No
President Donald Trump is meeting with Turkish President Recep Tayyip Erdogan	No	No	No	No	No
In a proclamation signed Saturday, hours after Bush died, Trump has also designated Wednesday as a national day of mourning.	No	No	No	No	No
President Donald Trump held an "informal conversation" with Russian President Vladimir Putin Friday.	No	No	No	No	Yes
Trump and Xi were set to talk over dinner Saturday at a summit of world leaders in Argentina.	No	No	No	No	No
Both sides have raised tariffs on billions of dollars of each country's goods.	No	No	No	No	No
Trump said they would also discuss the opioid crisis. He said he hopes China will criminalize fentanyl, the synthetic form of opioids largely shipped from China.	No	No	No	No	No
Bush, who was president from 1989 to 1993, died late Friday at age 94.	No	No	No	No	No
Bill Clinton closed his remembrance in the Post with an observation about the way things are now.	No	No	No	No	No
The federal Department of Justice declined to comment on Friday's ruling.	No	No	No	No	No
Mr. Cohen said that, to help Mr. Trump politically, he had falsely testified that the Russia efforts had been dropped by January 2016, when they had in fact continued until June.	No	No	No	No	No
Mr. Giuliani called Mr. Cohen a liar.	No	No	Yes	No	Yes
The president told reporters he did not know exactly when the Russia discussions had ceased, but he said his recollection was early in 2016.	No	No	No	No	No
Soon after the news about Mr. Cohen broke, Mr. Trump canceled his planned meeting with President Vladimir Putin, citing Russia's unresolved naval standoff with Ukraine.	No	No	No	No	No
Good Friday morning. Here are some of the stories making news in Washington and politics today.	No	No	No	No	No
Cohen and Individual 2 discussed efforts to obtain Russian governmental approval for the Moscow Project. COHEN discussed the status and progress of the Moscow Project with Individual 1 on more than the three occasions Cohen claimed to the Committee, and he briefed family members of Individual 1 within the Company about the project.	No	No	No	No	No
Energy infrastructure is an expensive part of the problem, according to the assessment. The report does not provide firm estimates of expected costs.	No	No	No	No	No

As for the weeks of vote-counting, Padilla said it simply takes time to make sure every ballot is tallied in a state with 58 counties and 20 million registered voters.	No	No	No	No	No
Russians have greatly changed their opinion of US President Donald Trump over the past year, with nearly a half of them now thinking negatively of him, Russian sociologists have found.	Yes	Yes	No	Yes	Yes
Within hours it turned out that the email was dated September 14, a day after the WikiLeaks drop, rather than September 4 – as CNN (and later NBC) reported, citing anonymous sources. It was never explained who gave them the glaringly wrong information, or why the reporters never bothered to check the dates on the emails themselves.	No	Yes	No	No	Yes
Vladimir Putin is president of Russia. An editing mistake erroneously identified him as Vladimir Trump in an earlier version of this article	No	No	No	No	No
In a federal court in Manhattan on Thursday, Cohen pleaded guilty to making false statements to Congress in 2017. In statements given to the Senate Select Committee on Intelligence, Cohen said that Trump's efforts to build a Trump Tower in Moscow ended in January 2016 and were not discussed with other members of the Trump company.	No	No	No	No	No
The photo showed Pence being greeted by members of the Broward Sheriff's Office SWAT team at Fort-Lauderdale-Hollywood International Airport on Friday. One of the men pictured was wearing a red and black badge with a Q on it, which appears to be the same design as QAnon merchandise being sold online.	No	No	No	No	No
QAnon is an elaborate conspiracy born on 4Chan last October after an anonymous poster claiming to be a high level US government official called 'Q' began posting cryptic messages.	No	No	No	No	No
It claims President Donald Trump is secretly trying to bring down a child sex trafficking ring and the Mueller investigation is just a cover. The theory has since evolved and spawned numerous alternative narratives.	No	Yes	Yes	No	Yes
Lying to Congress, as Cohen did, is a federal crime. Lying on television, on the other hand, is just a bit shady.	No	No	Yes	Yes	Yes
Votes that arrive first get counted first, while those that get dropped off on Election Day get counted days and weeks later.	No	No	No	No	No
Ohlemacher reported from Washington.	No	No	No	No	No
The hearing will be held by Dec. 21.	No	No	No	No	No
He said Iran could resume enriching uranium to 20 percent purity if it fails to see economic benefits from the 2015 deal that curbed its nuclear program.	No	No	No	No	No

Trump has repeatedly denied any collusion with Russia and criticized Mueller's probe	No	No	No	No	No
Legislators in Canada and Mexico must still approve the pact.	No	No	No	No	No
He threatened to tear it up and withdraw the United States completely at times during the negotiation, which would have left trade between the three neighbors in disarray.	No	No	No	No	No
The three were still bickering over the finer points of the deal just hours before officials were due to sit down and sign it.	No	No	No	No	No
During a Fox News interview in November 2016, Major Golsteyn again said he had killed the Afghan man, who had been captured by American troops, found with bomb-making material and then released.	No	No	No	No	No
The Army has charged a Special Forces soldier with premeditated murder in a shooting that happened nearly nine years ago in southern Afghanistan	No	No	No	No	No
Investigators are examining whether Mr. Dowless or anyone working for him improperly handled, altered or even discarded large numbers of absentee ballots in the 2018 election.	No	No	No	No	No
The ruling was over a lawsuit filed this year by a group of Republican governors and state attorneys general.	No	No	No	No	No
A group of intervening states led by Democrats promised to appeal the decision	No	No	No	No	No
The Weekly Standard, a primary voice of conservative Washington that found itself out of step with the Trumpward turn in the Republican Party, is ceasing publication after 23 years, its owners announced on Friday.	No	No	No	No	No
Trump wants Congress to approve \$5 billion in border wall funding before a Dec. 21 deadline, by which time a new spending or delay measure must be passed by lawmakers on Capitol Hill and signed by the president before current funding levels expire.	No	No	No	No	No
Congressional Democrats, however, are unwilling to agree to more than \$1.6 billion for Trump's proposed impediments along the southern border with Mexico	No	No	No	No	No
He told Reuters his committee may consider future funding restrictions for such missions.	No	No	No	No	No
Michael Cohen will testify next week at a closed hearing of the U.S. House Intelligence Committee after postponing an appearance before another congressional panel	No	No	No	No	No
Mr. Cohen has relayed to the Committee his legitimate concerns for his own safety as well as that of his family, which have been fueled by improper comments made by the President and his lawyer	No	Yes	No	No	Yes

The Left party is the largest opposition group in the German parliament, and has previously called for closer ties with Russia.	Yes	No	No	Yes	Yes
With three hours left to go in the month of February, Sen. Sanders challenged his supporters to bring the month's total individual contributions from its current total of \$41.6 million to \$45 million, reported The Hill.	No	No	No	No	No
Have you seen his hands? They're like this	No	No	No	No	Yes
Speaker John Bercow is to decide which amendments are put forward and voting will take place from 19:00 GMT.	No	No	No	No	No
The votes could guide the next stage of Theresa May's negotiations with the EU	No	No	No	No	No
will find out later if their efforts to change the direction of Brexit will come to pass.	No	No	No	No	No
will then close the debate	No	No	No	No	No
It is indisputable	Yes	No	No	No	No
Mr Greer initially made the comment in response to a Conservative tweet on the anniversary of Churchill's death.	No	No	No	No	No
Green MSP Ross Greer has repeated his controversial claim that Sir Winston Churchill was a "white supremacist" and a "mass murderer."	No	No	Yes	Yes	Yes