



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
ΣΧΟΛΗ ΚΟΙΝΩΝΙΚΩΝ ΚΑΙ ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΟΛΙΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΔΙΕΘΝΩΝ ΣΧΕΣΕΩΝ

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Παγκόσμιες Προκλήσεις και Συστήματα Αναλύσεων»

«Ανάλυση συναισθήματος σε ελληνικές
ειδησεογραφικές ιστοσελίδες σχετικά με το θέμα της

Συμφωνίας των Πρεσπών»

Μεταπτυχιακή Διπλωματική Εργασία

Αλεξάνδρα **Βαγγελάτου**

Τριμελής επιτροπή:

Αναπληρωτής Καθηγητής Ν.Σ. Κουτσούκης

Επίκουρος Καθηγητής Ε. Φακιολάς

Επίκουρος Καθηγητής Ν. Ραχανιώτης

Τελική έκδοση

Κόρινθος, 2018



UNIVERSITY OF THE PELOPONNESE
SCHOOL OF SOCIAL AND POLITICAL SCIENCES
DEPARTMENT OF POLITICAL SCIENCE & INTERNATIONAL RELATIONS

Master of Arts in
“Global Risks and Analytics”

“Sentiment analysis in Greek news sites regarding the
Prespa Agreement”

Master's Dissertation

Alexandra Vangelatou

Supervisors:

Associate Professor N.S. Koutsoukis

Assistant Professor E. Fakiolas

Assistant Professor N. Rachaniotis

Final version

Corinth, 2018

Φύλλο αξιολόγησης

Η διπλωματική εργασία με τίτλο «Ανάλυση συναισθήματος σε ελληνικές ειδησεογραφικές ιστοσελίδες σχετικά με το θέμα της *Συμφωνίας των Πρεσπών*» της Αλεξάνδρας Βαγγελάτου αξιολογήθηκε από την τριμελή επιτροπή, τόσο ως προς την ποιότητα του κειμένου, όσο και ως προς την ποιότητα της προφορικής παρουσίασης και υπεράσπισης της διπλωματικής εργασίας ενώπιον ακροατηρίου.

Η διαδικασία αξιολόγησης της διπλωματικής εργασίας ολοκληρώθηκε την 07/02/2019 με γενική επίδοση:

- Καλώς
- Λίαν Καλώς
- Άριστα

Τα μέλη της τριμελούς επιτροπής:

1. Αναπληρωτής Καθηγητής Ν.Σ. Κουτσούκης
2. Επίκουρος Καθηγητής Ε. Φακιολάς
3. Επίκουρος Καθηγητής Ν. Ραχανιώτης

Abstract

In this thesis, lexicon-based sentiment analysis approach has been applied for the study of data collected from Greek news sites with main subject of interest being the “Prespa Agreement”. The purpose was to identify current trends, over a specific time frame, on this particular issue.

For successful analysis, using the KNIME Analytics Platform, the database to be processed was originally created and included texts centered on the subject under consideration. Then, a specialized sentiment lexicon was created for this analysis and finally, the complete processing model was created and the results have been presented. In addition to a better understanding of the overall process, important elements and definitions of sentiment analysis, as well as the methodology followed at each stage of the procedure, are presented, and each step has been analyzed in depth.

From the final results of the method, it is being concluded that the trend of news story during the selected carrying out period was neutral on the issue of the “Prespa Agreement”.

Περίληψη

Στα πλαίσια της παρούσας εργασίας εφαρμόστηκε η μέθοδος ανάλυσης συναισθήματος με χρήση λεξικού για τη μελέτη δεδομένων που προέκυψαν από ελληνικές ειδησεογραφικές σελίδες και με κύριο θέμα ενδιαφέροντος τη «Συμφωνία των Πρεσπών». Σκοπός ήταν ο εντοπισμός των τάσεων της επικαιρότητας, σε χρονικό διάστημα συγκεκριμένο, ως προς το εν λόγω θέμα.

Για την επιτυχή διεξαγωγή της ανάλυσης και με τη χρήση του συστήματος αναλύσεων KNIME, δημιουργήθηκε αρχικά η προς επεξεργασία βάση δεδομένων, που περιλάμβανε κείμενα επικεντρωμένα στο υπό εξέταση θέμα. Στη συνέχεια, ακολούθησε η δημιουργία εξειδικευμένου λεξικού συναισθήματος για την παρούσα ανάλυση και τέλος, δημιουργήθηκε το ολοκληρωμένο μοντέλο επεξεργασίας των δεδομένων και παρουσιάστηκαν τα αποτελέσματα που προέκυψαν. Συμπληρωματικά και για την καλύτερη κατανόηση της συνολικής διαδικασίας, παρουσιάζονται σημαντικά στοιχεία και ορισμοί της ανάλυσης συναισθήματος, καθώς και η μεθοδολογία που ακολουθήθηκε σε κάθε στάδιο, ενώ κάθε στάδιο αναλύεται εις βάθος.

Από τα τελικά αποτελέσματα της μεθόδου, προκύπτει το συμπέρασμα ότι η τάση της ειδησεογραφικής επικαιρότητας, κατά το χρονικό διάστημα που επιλέχθηκε για τη διεξαγωγή της ανάλυσης, ήταν ουδέτερη ως προς το ζήτημα της «Συμφωνίας των Πρεσπών».

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Με την παρούσα δήλωση:

1. Δηλώνω ρητά και ανεπιφύλακτα ότι η διπλωματική εργασία που σας καταθέτω αποτελεί προϊόν δικής μου πνευματικής προσπάθειας, δεν παραβιάζει τα δικαιώματα τρίτων μερών και ακολουθεί τα διεθνώς αναγνωρισμένα πρότυπα επιστημονικής συγγραφής, τηρώντας πιστά την ακαδημαϊκή δεοντολογία.
2. Επίσης, οι απόψεις που εκφράζονται αποτελούν αποκλειστικά ευθύνης της συγγραφέως και ο επιβλέπων, οι εξεταστές, το Τμήμα και το Πανεπιστήμιο Πελοποννήσου δεν υιοθετούν κατ' ανάγκη τις εκφραζόμενες απόψεις ούτε φέρουν οποιαδήποτε ευθύνη για τυχόν λάθη και παραλείψεις.

Η δηλούσα

Βαγγελιάτου Αλεξάνδρα

Περιεχόμενα

Abstract	I
Περίληψη.....	II
Περιεχόμενα	IV
Περιεχόμενα πινάκων.....	V
1 Εισαγωγή.....	1
1.1 Διαδίκτυο.....	1
1.2 Ανάλυση συναισθήματος – Εξόρυξη γνώμης	2
1.3 Στόχος και δομή της εργασίας.....	3
2 Συμφωνία των Πρεσπών.....	5
2.1 Ιστορική αναδρομή.....	5
2.1.1 Περίοδος αναφοράς 1991-1999.....	5
2.1.2 Περίοδος αναφοράς 2000-2009.....	6
2.1.3 Περίοδος αναφοράς 2010-2018.....	7
2.2 Η Συμφωνία των Πρεσπών.....	8
3 Ανάλυση συναισθήματος	10
3.1 Μέθοδοι μηχανικής μάθησης	10
3.1.1 Μέθοδοι επιβλεπόμενης μάθησης	10
3.1.2 Μέθοδοι μη επιβλεπόμενης μάθησης.....	11
3.1.3 Μέθοδοι ημι-επιβλεπόμενης μάθησης.....	12
3.2 Ανάλυση συναισθήματος με τη χρήση λεξικού	12
3.3 Προκλήσεις στην ανάλυση συναισθήματος	13
4 Μεθοδολογία	15
4.1 Σύστημα αναλύσεων KNIME	15
4.2 Συλλογή δεδομένων.....	16
4.3 Προεπεξεργασία δεδομένων.....	18
4.4 Δημιουργία λεξικού συναισθήματος	19
4.5 Τελική επεξεργασία δεδομένων	20
5 Εμπειρικά αποτελέσματα	21

5.1 Επιλογή δεδομένων και περίοδος αναφοράς	21
5.2 Ανάλυση συναισθήματος με τη χρήση λεξικού	22
5.2.1 Συλλογή δεδομένων.....	22
5.2.2 Δημιουργία εξειδικευμένου λεξικού.....	25
5.3 Επεξεργασία δεδομένων και αποτελέσματα.....	26
6 Συμπεράσματα.....	29
Κατάλογος πηγών.....	31
Ξενόγλωσσες πηγές.....	31
Διαδικτυακοί τόποι/ Ιστοσελίδες.....	33
Παράρτημα Α	34
Παράρτημα Β	42

Περιεχόμενα πινάκων

Πίνακας 5.1.....	23
Πίνακας 5.2.....	24
Πίνακας 5.3.....	27

1 Εισαγωγή

Τα συναισθήματα αποτελούν σημαντικό παράγοντα στην εξέλιξη των ανθρώπινων σχέσεων και η κατανόησή τους παίζει ιδιαίτερο ρόλο, κυρίως τα τελευταία χρόνια, στην προσωπική ανάπτυξη των ανθρώπων. Πολύ συχνά οι άνθρωποι εξαρτώνται από τη γνώμη και τα συναισθήματα των άλλων για τη λήψη σημαντικών αποφάσεων για κρίσιμα θέματα που τους αφορούν σε διάφορες στιγμές της ζωής τους. Σήμερα το ρόλο του κοινωνικού περίγυρου των ανθρώπων παίζει πολύ συχνά το διαδίκτυο, καθώς και οι ταχύτατα αναπτυσσόμενες εφαρμογές μέσω κοινωνικής δικτύωσης.

Την ευκαιρία αυτή, για άμεση πρόσβαση σε υπεράριθμες διαφορετικές απόψεις των χρηστών του διαδικτύου και των μέσων κοινωνικής δικτύωσης, εντόπισαν πολύ γρήγορα, τόσο η ακαδημαϊκή κοινότητα όσο και οι επιχειρήσεις. Παράλληλα, πολιτικοί παράγοντες χρησιμοποιούν δεδομένα που αντλούνται από το διαδίκτυο και τα μέσα κοινωνικής δικτύωσης με σκοπό τον εντοπισμό των τάσεων της κοινής γνώμης είτε για θέματα πολιτικού ενδιαφέροντος και πολιτικά γεγονότα είτε για να προσδιορίσουν το ενδιαφέρον, ενίοτε και την αδιαφορία, των πολιτών για συγκεκριμένα πολιτικά πρόσωπα.

Υπάρχει πληθώρα μη επεξεργασμένων δεδομένων, γνωστά και ως μη-δομημένα (unstructured data), που μπορούν να αντληθούν από διάφορες πηγές του διαδικτύου με διάφορους τρόπους και που με την κατάλληλη επεξεργασία, μπορούν να διευκολύνουν σημαντικά την επίτευξη εμπορικών, πολιτικών ή ερευνητικών στόχων. Για την επίτευξη των στόχων αυτών αναπτύχθηκε η τεχνητή νοημοσύνη και η ανάλυση συναισθήματος ή εξόρυξη γνώμης (sentiment analysis – opinion mining). Υπάρχουν πολλά εργαλεία εμπορικής ή ακαδημαϊκής χρήσης που έχουν σχεδιαστεί με στόχο την ανάλυση δεδομένων και τον εντοπισμό του συναισθήματος διαφόρων κοινωνικών ομάδων, ειδικά μέσω διαδικτύου.

1.1 Διαδίκτυο

Η πορεία εξέλιξης του διαδικτύου και οι δυνατότητες που θα προσφέρει στο άμεσο μέλλον είναι ασύλληπτες σήμερα. Ήδη, βρίσκεται σε εξέλιξη η εποχή όπου οι διαδικτυακές κοινότητες επηρεάζουν και παίζουν ενίοτε καθοριστικό ρόλο, όχι μόνο στην εξέλιξη προϊόντων και υπηρεσιών, αλλά και στην εξέλιξη πολιτικών γεγονότων. Σημαντικές εφαρμογές στην εξέλιξη του διαδικτύου έχουν αποτελέσει τόσο οι ιστοσελίδες όσο και τα μέσα κοινωνικής δικτύωσης, για τη χρήση των οποίων δεν απαιτούνται ιδιαίτερες τεχνικές γνώσεις από τους χρήστες, γεγονός που αποτελεί κρίσιμο παράγοντα στην αυξανόμενη χρήση του διαδικτύου.

Το διαδίκτυο σήμερα δεν αποτελείται πλέον από στατικές εικόνες. Οι δυνατότητες της αναζήτησης, του σχολιασμού των χρηστών, της παράθεσης σχετικών συνδέσμων και επισημάνσεων, καθώς και δυνατότητα συγγραφής και ανάρτησης κειμένων, προσφέρουν μία πληθώρα δεδομένων, που με την κατάλληλη επεξεργασία μπορούν να αποτελέσουν τη βάση εξόρυξης σημαντικών αποτελεσμάτων για τις τάσεις (Benedetto & Tedeschi, 2016). Επιπλέον, η επεξεργασία των δεδομένων αυτών επιτρέπει τη χαρτογράφηση των ενδιαφερόντων του χρήστη και έμμεσα του ασκείται σημαντική επιρροή είτε μέσω αναδυόμενων παραθύρων είτε μέσω διαφημίσεων σχετικών με τα θέματα που ενδιαφέρουν το χρήστη.

Συγκεκριμένα, στον τομέα της επικαιρότητας, παίζουν σημαντικό ρόλο οι δημοσιογραφικές ιστοσελίδες και οι ιστοσελίδες εφημερίδων, δεδομένου ότι ο χρήστης μπορεί με απλή εγγραφή να λαμβάνει ενημερώσεις και ειδοποιήσεις αναρτήσεων νέων άρθρων και σχολίων, σχετικών με τα θέματα που τον ενδιαφέρουν σε πραγματικό χρόνο, δηλαδή αμέσως μετά την ανάρτησή τους. Η δυνατότητα αυτή ενισχύθηκε μέσω της ανάπτυξης της μεθόδου ανταλλαγής ψηφιακού πληροφοριακού περιεχομένου μέσω του διαδικτύου, επωνομαζόμενο ως RSS (Rich Site Summary, συχνά παραφρασμένου ως Real Simple Syndication).

Η μέθοδος του RSS υποστηρίζεται από τη γλώσσα σήμανσης XML, περιέχει όλη την απαραίτητη πληροφόρηση που μπορεί να ενδιαφέρει το χρήστη, όπως πλήρη δεδομένα, ημερομηνία έκδοσης και μεταδεδομένα (metadata, τα οποία αναφέρονται σε πληροφορίες για τα ίδια τα δεδομένα και περιγράφουν χαρακτηριστικά των πληροφοριακών οντοτήτων), ενώ παράλληλα διευκολύνει σημαντικά την αυτόματη ενημέρωση των χρηστών για νέα και γεγονότα, (Zadrozny & Kodali, 2013).

1.2 Ανάλυση συναισθήματος – Εξόρυξη γνώμης

Ο όρος ανάλυση συναισθήματος, εναλλακτικά εξόρυξη γνώμης, αναφέρεται στη δυνατότητα επεξεργασίας και ανάλυσης κειμένων μέσω διαφόρων τεχνικών προγραμματισμού, με σκοπό τον προσδιορισμό της άποψης του συγγραφέα ως προς ένα συγκεκριμένο θέμα. Η άποψη αυτή μπορεί να αφορά είτε την κριτική του συγγραφέα ως προς το υπό διερεύνηση θέμα είτε τη σκόπιμη επιθυμία του συγγραφέα να επηρεάσει προς συγκεκριμένη κατεύθυνση τον αναγνώστη.

Παράλληλα, οι εκφράσεις συνήθως χωρίζονται σε δύο κατηγορίες: τα γεγονότα, που είναι πιο αντικειμενικά και την άποψη, που είναι πιο υποκειμενική. Είναι σημαντικό κατά την ανάλυση συναισθήματος να αναγνωριστούν και να διαχωριστούν τα γεγονότα από τη γνώμη (Luo, Chen, Xu, & Zhou, 2013). Τα χαρακτηριστικά που προσδιορίζουν την ανάλυση συναισθήματος χωρίζονται σε τέσσερις κατηγορίες και είναι συντακτικά, σημασιολογικά, μορφολογικά και βασισμένα σε συσχετισμούς (link-based) (Chen, 2012).

Ως βασικές αρχές της ανάλυσης συναισθήματος έχουν χαρακτηριστεί η αναγνώριση συναισθήματος, που εστιάζει κυρίως στην εξαγωγή συναισθημάτων που προκύπτουν μέσα από ένα κείμενο και αποτελεί μία δραστηριότητα που περικλείεται στη διαδικασία ανίχνευσης τάσεων, και η ανίχνευση τάσεων, που κατηγοριοποιεί τα κείμενα ως «θετικά» ή «αρνητικά» προσκείμενα σε σχέση με το υπό διερεύνηση θέμα και στην ουσία μπορεί να αποτελέσει

δραστηριότητα που περικλείεται σε διαδικασίες πιο προχωρημένων αναλύσεων (Cambria, Das, Bandyopadhyay, & Feraco, 2017).

Στην παρούσα εργασία, η ανάλυση συναισθήματος εφαρμόζεται κυρίως σε δεδομένα που έχουν συλλεχθεί μέσω ενημερωτικών ιστοσελίδων με τη χρήση του RSS κι έχουν επεξεργαστεί με το σύστημα αναλύσεων KNIME. Ωστόσο, πολύ συχνή είναι και η εφαρμογή της σε δεδομένα που συλλέγονται από μέσα κοινωνικής δικτύωσης και ιστοσελίδες εμπορικού ή άλλου περιεχομένου.

Επιπλέον, για τη σωστή εφαρμογή της ανάλυσης συναισθήματος πρέπει να προσδιοριστούν εξαρχής οι τάξεις πολικότητας κατά την ανίχνευση τάσεων, για παράδειγμα ανάλυση δύο τάξεων αναφέρεται στον προσδιορισμό της πολικότητας του κειμένου ως θετικό ή αρνητικό, καθώς και ο προσδιορισμός του επιπέδου στο οποίο θα εφαρμοστεί η ανάλυση, δηλαδή την επιλογή μεταξύ τριών επιπέδων ανάλυσης: σε επίπεδο κειμένου (document level), σε επίπεδο πρότασης (sentence level) και σε επίπεδο φράσης (phrase level). Μία ακόμη σημαντική εργασία κατά την εφαρμογή της ανάλυσης συναισθήματος είναι οι υποθέσεις που τίθενται για το υπό διερεύνηση θέμα, που αποτελεί την πηγή του συναισθήματος (Chen, 2012). Στην παρούσα εργασία, για τον καλύτερο προσδιορισμό των τάσεων της επικαιρότητας, η ανάλυση εφαρμόζεται σε επίπεδο κειμένου, λαμβάνοντας ως κείμενα τα άρθρα που συλλέχθηκαν από ειδησεογραφικές ιστοσελίδες μέσω της χρήσης του RSS.

Οι τεχνικές ανάλυσης συναισθήματος, που αναφέρονται συνηθέστερα στη βιβλιογραφία, χωρίζονται σε δύο βασικές κατηγορίες: α) στην ανάλυση με αλγόριθμους μηχανικής μάθησης, που περιλαμβάνουν τις μεθόδους επιβλεπόμενης και μη επιβλεπόμενης μάθησης, και β) στην ανάλυση με τη χρήση λεξικού (Benedetto & Tedeschi, 2016).

Η ανάλυση συναισθήματος, μέχρι σήμερα, έχει εφαρμοστεί εκτενώς κι έχει εξελιχθεί σε κείμενα αγγλικής γλώσσας, ενώ η ανάλυση κειμένων στην ελληνική γλώσσα είναι σαφώς πιο περιορισμένη, παρά την ολοένα και αυξανόμενη χρήση του διαδικτύου από τους Έλληνες χρήστες. Επιπλέον, η ανάλυση συναισθήματος σε κείμενα με ελληνικό περιεχόμενο αποτελεί μία ιδιαίτερη πρόκληση, λαμβάνοντας υπόψη το λεκτικό πλούτο και τις γραμματικές και συντακτικές ιδιαιτερότητες της ελληνικής γλώσσας. Τα δύο αυτά χαρακτηριστικά, σε συνδυασμό με τη μερική ανάλυση και δημιουργία ελληνικών λεξικών συναισθήματος σε σχέση με τα λεξικά για την αγγλική ή άλλες γλώσσες όπως η ισπανική, δημιουργούν την ανάγκη περαιτέρω μελέτης και ανάπτυξης της ανάλυσης συναισθήματος σε κείμενα με ελληνικό περιεχόμενο.

1.3 Στόχος και δομή της εργασίας

Δεδομένης της τάσης των χρηστών κατά τη διάρκεια των τελευταίων ετών να ενημερώνονται από ιστοσελίδες ειδησεογραφικού περιεχομένου σχετικά με οικονομικά, κοινωνικά και κυρίως πολιτικά ζητήματα, προκύπτει η ανάγκη διερεύνησης των τάσεων της αρθρογραφίας που σχετίζεται με θέματα της επικαιρότητας. Πρόκειται για μία πηγή πληροφόρησης από την οποία μπορούν να προκύψουν αποτελέσματα που πιθανώς να σχετίζονται και με τις τάσεις της κοινής γνώμης.

Σκοπός της παρούσας εργασίας είναι η δημιουργία ενός συστήματος εξόρυξης γνώμης με βάση την επίκαιρη αρθρογραφία και μέσω της χρήσης του RSS ενημερωτικών ιστοσελίδων και από το οποίο θα προκύπτει η τάση (θετική, αρνητική ή ουδέτερη πολικότητα) των κειμένων σχετικά με τη «Συμφωνία των Πρεσπών» μεταξύ της Ελλάδος και της Πρώην Γιουγκοσλαβικής Δημοκρατίας της Μακεδονίας, κατά την περίοδο 27 Σεπτεμβρίου 2018 έως 31 Οκτωβρίου 2018. Το κύριο ερώτημα που τίθεται υπό διερεύνηση είναι, για το συγκεκριμένο διάστημα, ποια είναι η τάση που προκύπτει από ειδησεογραφικές ιστοσελίδες σχετικά με το θέμα της «Συμφωνίας των Πρεσπών».

Προς την κατεύθυνσή αυτή, στη συνέχεια της παρούσας εργασίας, θα παρουσιαστούν αναλυτικά τα εξής: Στο δεύτερο κεφάλαιο θα γίνει αναφορά στο ιστορικό και θεωρητικό πλαίσιο της «Συμφωνίας των Πρεσπών». Στο τρίτο κεφάλαιο παρουσιάζεται η μεθοδολογία που πραγματεύεται περισσότερο η τρέχουσα βιβλιογραφία και οι δυνατές μέθοδοι ανάλυσης συναισθήματος, ενώ παρουσιάζονται και οι προκλήσεις που προκύπτουν κατά την εφαρμογή της. Στο τέταρτο κεφάλαιο αναπτύσσεται η θεωρία της μεθόδου ανάλυσης συναισθήματος που ακολουθήθηκε για τη δημιουργία του συστήματος εξόρυξης γνώμης και την ανίχνευση των τάσεων της επικαιρότητας σχετικά με το ζήτημα της «Συμφωνίας των Πρεσπών», ενώ στο πέμπτο κεφάλαιο παρουσιάζονται τα εμπειρικά αποτελέσματα που προέκυψαν κατά την εφαρμογή του εν λόγω συστήματος και οι προκλήσεις που έπρεπε να αντιμετωπιστούν για την ολοκλήρωση της ανάλυσης. Τέλος, στο έκτο κεφάλαιο συνοψίζονται τα συμπεράσματα και γίνεται αναφορά στις δυνατότητες επέκτασης και βελτίωσης του συγκεκριμένου συστήματος.

2 Συμφωνία των Πρεσπών

Κατά τη διάρκεια των τελευταίων ετών και για διάστημα πάνω από 20 χρόνια επικράτησε η διπλωματική διαμάχη μεταξύ των χωρών της Ελλάδας και της πρώην Γιουγκοσλαβικής Δημοκρατίας της Μακεδονίας (π.Γ.Δ.Μ.). Οι προσπάθειες επίλυσης του προβλήματος ήταν συνεχείς με χρονικά διαστήματα έντονης κορύφωσης διαπραγματεύσεων, ενώ έλαβαν χώρα και γεγονότα όξυνσης του προβλήματος που επηρεάστηκαν από πολιτικές αποφάσεις που έλαβαν οι δύο χώρες.

Η «Συμφωνία των Πρεσπών» ήταν η κατάληξη του τελευταίου κύκλου διαπραγματεύσεων, που ξεκίνησε κατά το έτος 2017. Ωστόσο, δεν μπορεί να θεωρηθεί ιστορικά ότι το ζήτημα έχει επιλυθεί πλήρως, δεδομένου ότι ακόμη και τώρα λαμβάνουν χώρα γεγονότα στο εσωτερικό των δύο χωρών, τα οποία θα κρίνουν και την τελική έκβαση.

Στη συνέχεια και για την καλύτερη κατανόηση του πλαισίου της Συμφωνίας, παρουσιάζονται τα κύρια ιστορικά γεγονότα που οδήγησαν στην τελική υπογραφή της «Συμφωνίας των Πρεσπών», καθώς και τα ζητήματα που αποτελούν την ουσία των διαπραγματεύσεων μέχρι σήμερα.

2.1 Ιστορική αναδρομή

2.1.1 Περίοδος αναφοράς 1991-1999

Το 1991, μετά τη διάλυση του κράτους της Γιουγκοσλαβίας, σχηματίστηκε το κράτος της Γιουγκοσλαβικής Δημοκρατίας της Μακεδονίας. Αμέσως προέκυψαν αντιδράσεις από την πλευρά της Ελλάδας, που πυροδοτήθηκαν κυρίως με αφορμή την αναφορά του όρου «Δημοκρατία της Μακεδονίας» στο σύνταγμα του νεοσύστατου κράτους, τον καθορισμό της εθνικότητας, της γλώσσας, της ιστορίας, καθώς και τη χρήση ιστορικών συμβόλων από το γειτονικό κράτος (Ceka, 2018). Επιπλέον, εκφράστηκαν φόβοι από την πλευρά του Ελληνικού κράτους για πιθανές εδαφικές διεκδικήσεις από το γειτονικό κράτος προς την Ελλάδα. Οι φόβοι αυτοί θα μπορούσε να θεωρηθεί ότι προέρχονται από τις εκφρασμένες πολιτικές διεκδικήσεις πρώην πολιτικού προσώπου του κράτους της πρώην Γιουγκοσλαβίας περί επέκτασης των συνόρων του τότε κράτους μέχρι την πόλη της Θεσσαλονίκης (Mavromatidis, 2010).

Σε μία πρώτη προσπάθεια μη δημιουργίας σημαντικών προβλημάτων, η Ελλάδα και η Ε.Ο.Κ. (τότε) το 1992 αναγνώρισαν το νέο κράτος, θέτοντας όμως σαφείς όρους περί αλλαγής άρθρων και όρων του Συντάγματος της γειτονικής χώρας και του συμβόλου της σημαίας. Η κυβέρνηση του γειτονικού κράτους, σε μία ένδειξη προσπάθειας συνεργασίας, προχώρησε στην τροποποίηση κάποιων άρθρων του συντάγματος, χωρίς όμως να υποχωρήσει στο θέμα της ονομασίας (Tziampiris, 2005). Συγκεκριμένα, έγιναν δύο πολύ σημαντικές αλλαγές, η μία σχετικά με τη μη υποστήριξη εδαφικών διεκδικήσεων προς τα γειτονικά κράτη της χώρας και η άλλη σχετικά τη φροντίδα μειονοτήτων που ζουν σε γειτονικά κράτη χωρίς καμία παρέμβαση στις εσωτερικές υποθέσεις των γειτονικών αυτών κρατών (Tanev, 1999). Ωστόσο, από την πλευρά της Ελλάδας οι πιέσεις για εξαίρεση του όρου «Μακεδονία» από την ονομασία του νέου κράτους συνεχίστηκαν, ενώ υπήρξαν μαζικές αντιδράσεις από πολίτες του γεωγραφικού διαμερίσματος της Μακεδονίας (Fidanovski, 2018).

Το 1993, ο Οργανισμός Ηνωμένων Εθνών (Ο.Η.Ε.) αναγνώρισε το νέο κράτος με την ονομασία «πρώην Γιουγκοσλαβική Δημοκρατία της Μακεδονίας» (π.Γ.Δ.Μ.), παράλληλα με την αποδοχή ότι η ονομασία τίθεται επίσημα ως θέμα διαπραγμάτευσης, γεγονός που εκκίνησε τη διαδικασία διαπραγματεύσεων μεταξύ των δύο χωρών για τα επόμενα χρόνια (Ceka, 2018). Για το κράτος της π.Γ.Δ.Μ., νομικά, ήταν ιδιαίτερα σημαντική η αναγνώριση του ονόματος ακόμη και με τη συγκεκριμένη τροποποίηση, δεδομένου ότι η ύπαρξη ονόματος προσφέρει στο κράτος νομική οντότητα, που κατά συνέπεια επιτρέπει τη σύναψη συμβάσεων και την δυνατότητα διαπραγμάτευσης με άλλα νομικά πρόσωπα (Tanev, 1999).

Στα πλαίσια των διαπραγματεύσεων και ως μέσο πίεσης για την επίλυση του θέματος, η Ελλάδα το 1994 αποφάσισε την εφαρμογή της πολιτικής οικονομικού αποκλεισμού προς τη γειτονική χώρα, η οποία οδήγησε τελικά στην απόφαση από μέρους της π.Γ.Δ.Μ. να αλλάξει το σύμβολο της σημαίας του κράτους (Ceka, 2018). Επιπλέον, η απόφαση περί οικονομικού αποκλεισμού είχε σημαντικές οικονομικές και πολιτικές συνέπειες στη γειτονική χώρα και αποτέλεσε την αφορμή μεσολάβησης τρίτων χωρών με σκοπό την επίλυση του θέματος, γεγονός που είχε ως αποτέλεσμα την επίτευξη «Ενδιάμεσης Συμφωνίας» το 1995 (Fidanovski, 2018).

Η «Ενδιάμεση Συμφωνία» δημιούργησε έντονες τριβές στο εσωτερικό της π.Γ.Δ.Μ., πυροδοτώντας εξελίξεις σε άλλα προβλήματα όπως την κρίση στο Κόσοβο και την αυξανόμενη ένταση στο θέμα της Αλβανικής μειονότητας στη γειτονική χώρα, ενώ παρά την υπογραφή της συμφωνίας οι προσπάθειες συμβιβασμού και διαπραγμάτευσης μεταξύ των δύο χωρών δεν βελτιώθηκαν σχεδόν καθόλου. Αντίθετα, σε μία προσπάθεια για τη διεθνή αναγνώριση του κράτους της π.Γ.Δ.Μ., πολιτικά μέλη της τότε κυβέρνησης του γειτονικού κράτους υιοθέτησαν μια μάλλον επιθετική πολιτική (Tziampiris, 2005).

Παράλληλα με το ζήτημα της ονομασίας της π.Γ.Δ.Μ. που είχε προκύψει με την Ελλάδα, προβλήματα δημιουργήθηκαν και με άλλα γειτονικά κράτη της π.Γ.Δ.Μ., όπως η Βουλγαρία, η Σερβία και η Αλβανία, ευνοώντας την ανάπτυξη αστάθμητων παραγόντων στην περιοχή των Βαλκανίων (Nikodinovska-Stefanovska, 2014).

2.1.2 Περίοδος αναφοράς 2000-2009

Δεδομένων όλων των γεγονότων που είχαν λάβει χώρα έως το 2000, είχε γίνει κατανοητό ότι η καλύτερη προοπτική για το κράτος της π.Γ.Δ.Μ. ήταν η ένταξη του κράτους στην Ε.Ε. στο

απώτερο μέλλον. Έτσι, το 2000 έγινε η έναρξη της Διαδικασίας Σταθεροποίησης και Συνεργασίας (Stabilisation and Association Process – SAP) για το γειτονικό κράτος υποστηριζόμενη από την Ε.Ε., η οποία κατέληξε στην υπογραφή της Συμφωνίας Σταθερότητας και Συνεργασίας (Stabilisation and Association Agreement – SAA) το 2001 και που τελικά τέθηκε σε οριστική εφαρμογή το 2004 (Mavromatidis, 2010). Κατά την αμέσως επόμενη χρονιά, 2005, η π.Γ.Δ.Μ. κατέθεσε αίτηση για ένταξη του κράτους στην Ε.Ε., η οποία έγινε δεκτή χωρίς καθορισμό συγκεκριμένης ημερομηνίας ένταξης, ενώ δεν σημειώθηκε πρόοδος στη διαδικασία για αρκετά χρόνια μετέπειτα, δεδομένου ότι η Ελλάδα ήταν η χώρα που είχε σοβαρές αντιρρήσεις για το ζήτημα της ονομασίας (Ceka, 2018).

Το 2007 ξεκίνησε ένας νέος κύκλος διαπραγματεύσεων μεταξύ των δύο χωρών, με κύρια θέματα συζήτησης την αλλαγή του ονόματος της π.Γ.Δ.Μ. με τη χρήση διευκρινιστικού όρου είτε γεωγραφικού προσδιορισμού όπως «Ανω» ή «Νέα» και το σαφή διαχωρισμό του όρου «Μακεδονία» σε σχέση με την αναφορά σε ιστορικούς αντίστοιχους όρους. Η αποδοχή πλέον, από την ελληνική πλευρά, του όρου «Μακεδονία» στην ονομασία του γειτονικού κράτους αποτέλεσε ένα θετικό βήμα στην εξέλιξη των διαπραγματεύσεων (Fidanovski, 2018).

Τελικά δεν κατέληξαν σε ουσιαστικό αποτέλεσμα οι διαπραγματεύσεις της εν λόγω περιόδου και το 2008 η Ελλάδα άσκησε το δικαίωμα της αρνησικυρίας στη διαδικασία ένταξης της γειτονικής χώρας στο NATO, εμποδίζοντας ταυτόχρονα και τη διαδικασία ένταξης της π.Γ.Δ.Μ. στην Ε.Ε. (Bogdanovski, 2015). Παράλληλα, την ίδια χρονιά, η αναγνώριση από το κράτος της π.Γ.Δ.Μ. του Κοσόβου ως ανεξάρτητο κράτος, προκάλεσε νέες τριβές στις σχέσεις του κράτους με τη Σερβία, προκαλώντας ανησυχία για τη σταθερότητα της περιοχής των Βαλκανίων (Janev, Dimeski, & Ristov, 2014).

Στο ενδιάμεσο, κατά τη δεκαετία του 2000 οι οικονομικές και εμπορικές σχέσεις της π.Γ.Δ.Μ. με την Ε.Ε. και την Ελλάδα είχαν ενισχυθεί σημαντικά, με τη δεύτερη να έχει καταφέρει να εξασφαλίσει ιδιαίτερα σημαντική οικονομική και επενδυτική παρουσία στη γειτονική χώρα (Mavromatidis, 2010).

Ωστόσο, οι καλές οικονομικές και εμπορικές σχέσεις μεταξύ των χωρών της Ελλάδας και της π.Γ.Δ.Μ. δεν βοήθησαν όταν, κατά το έτος 2009 που τέθηκαν εκ νέου υπό συζήτηση τα θέματα της ένταξης της δεύτερης στο NATO και του ορισμού ημερομηνίας για ένταξη στην Ε.Ε.. Το αποτέλεσμα ήταν να παρεμποδιστούν οι διαδικασίες από την πλευρά της Ελλάδας και η π.Γ.Δ.Μ. να προσφύγει κατά της Ελλάδας στο Διεθνές Δικαστήριο, με αιτιολογία την παραβίαση της «Ενδιάμεσης Συμφωνίας» του 1995 (Fidanovski, 2018).

2.1.3 Περίοδος αναφοράς 2010-2018

Το 2011, το Διεθνές Δικαστήριο απεφάνθη υπέρ της π.Γ.Δ.Μ. αναγνωρίζοντας την παραβίαση από την πλευρά της Ελλάδας, ωστόσο, απέρριψε την πρόταση της πρώτης για δέσμευση της Ελλάδος περί μη υποβολής ένστασης στο μέλλον στις διαδικασίες ένταξης της π.Γ.Δ.Μ. στο NATO και την Ε.Ε. (Janev, Dimeski, & Ristov, 2014).

Κατά την περίοδο αυτή, η π.Γ.Δ.Μ. έπρεπε να αντιμετωπίσει σωρεία προβλημάτων στο εσωτερικό του κράτους. Η συνεχής πολιτική αστάθεια της χώρας, με αποκορύφωμα το πολιτικό σκάνδαλο περί εμπλοκής πολιτικών προσώπων σε εγκληματικές δραστηριότητες, δημιούργησαν ένα αρνητικό κλίμα κατά της π.Γ.Δ.Μ. (Bogdanovski, 2015). Αντίθετα, το κλίμα υπέρ της Ελλάδας ενισχύθηκε και ως μέλος του NATO και της Ε.Ε. θεωρήθηκε σημαντικός

παράγοντας σταθερότητας και ανάπτυξης της περιοχής, γεγονός ιδιαίτερα θετικό για την εξέλιξη του ζητήματος (Fidanovski, 2018).

Παράλληλα, οι ανησυχίες της Ε.Ε. για τις τεταμένες σχέσεις μεταξύ των Βαλκανικών κρατών και την πιθανή αύξηση της επιρροής της Ρωσίας στην ευρύτερη περιοχή των Βαλκανίων, κυρίως το διάστημα από το 2014 κι έπειτα, προκάλεσαν την ανάγκη περαιτέρω ενίσχυση της συνεργασίας μεταξύ των Βαλκανικών κρατών και τη σταθεροποίηση στην περιοχή, με σκοπό τη δημιουργία μίας περιοχής ειρήνης, ασφάλειας, σταθερότητας και συνεργασίας (Ivanovski, 2013).

Η σταθερότητα της περιοχής, γενικότερα, κρίθηκε από την Ε.Ε. ως βαρύνουσα σημασία στο πλαίσιο της μεγέθυνσης της, προσφέροντας έτσι την ευκαιρία στην Ελλάδα να εκμεταλλευτεί τη θέση της και να επιτύχει μία συμφωνία με την π.Γ.Δ.Μ. με τους ευνοϊκότερους δυνατούς όρους (Nikodinovska-Stefanovska, 2014).

Σημαντικό σημείο της δεκαετίας, φαίνεται να ήταν το έτος 2017, με την αλλαγή κυβέρνησης στο κράτος της π.Γ.Δ.Μ.. Το γεγονός αυτό, σε συνδυασμό με τα γεγονότα που είχαν προηγηθεί, αποτέλεσε την αφορμή για την έναρξη εκ νέου των διαδικασιών ένταξης της χώρας στην Ε.Ε με βασική προϋπόθεση την οριστική επίλυση του ζητήματος της ονομασίας. Η απόφαση της Ε.Ε. για επίλυση τυχόν θεμάτων με γειτονικές χώρες δεν αφορά μόνο την π.Γ.Δ.Μ., αλλά τέθηκε το Φεβρουάριο του 2018 ως απαραίτητη προϋπόθεση για την έναρξη προ-ενταξιακών διαπραγματεύσεων με την Ε.Ε. για τα κράτη των π.Γ.Δ.Μ., Αλβανίας, Βοσνίας-Ερζεγοβίνης και Κοσόβου (Dabrowski & Myachenkova, 2018). Παράλληλα, η κίνηση αυτή διευκολύνει και την εξέλιξη του ζητήματος της ένταξης της χώρας στο ΝΑΤΟ, γεγονός που αποτελεί επιθυμητό στόχο του ίδιου του κράτους (Ceka, 2018).

Οι διαπραγματεύσεις μεταξύ των δύο χωρών ξεκίνησαν άμεσα με τη θερμή υποστήριξη των πρωθυπουργών των δύο κρατών. Παρόλα αυτά, οι αντιδράσεις και στις δύο χώρες υπήρξαν έντονες κατά τα έτη 2017-2018, τόσο σε πολιτικό επίπεδο όσο και σε κοινωνικό.

2.2 Η Συμφωνία των Πρεσπών

Ως αποτέλεσμα των διαπραγματεύσεων και των γεγονότων που έλαβαν χώρα για διάστημα μεγαλύτερο των 25 ετών, προέκυψε η «Συμφωνία των Πρεσπών», που υπεγράφη τον Ιούνιο του 2018 μεταξύ των χωρών της Ελλάδος και της π.Γ.Δ.Μ..

Η συμφωνία αυτή διευθέτησε ένα χρόνιο πρόβλημα μεταξύ των δύο χωρών. Η ελληνική πλευρά επικεντρώθηκε κυρίως στην αποσαφήνιση των όρων «Μακεδονία» και «Μακεδόνας» που αναφέρεται στο σύνταγμα της π.Γ.Δ.Μ. σε σχέση με την ονομασία του όμορου γεωγραφικού διαμερίσματος και των ιστορικών όρων και αναφορών. Επιπλέον, όπως προκύπτει από τα θέματα που τέθηκαν υπό διαπραγμάτευση όλο αυτό το διάστημα, θεωρείται ιδιαίτερα σημαντική η διασφάλιση μη εδαφικών διεκδικήσεων από την πλευρά του γειτονικού κράτους και το σεβασμό ως προς τα υπάρχοντα σύνορα μεταξύ των δύο χωρών, σύμφωνα με την ερμηνεία από το Ελληνικό κράτος συγκεκριμένων συνταγματικών άρθρων της π.Γ.Δ.Μ.. Παράλληλα, στα πλαίσια της συμφωνίας διασφαλίστηκε η χρήση διευκρινιστικού όρου ή γεωγραφικού προσδιορισμού στο όνομα της γειτονικής χώρας και ο καθορισμός της γλώσσας και της εθνικότητας των πολιτών της.

Από μία άλλη οπτική, το ζήτημα ήταν ιδιάζουσας πολιτικής σημασίας για την Ελλάδα και δεν μπορεί να εξετάζεται μόνο στο πλαίσιο των ιστορικών γεγονότων και συμβόλων, σαν μία πολιτική κρίση στο εσωτερικό της Ελλάδος. Η ηθική διάσταση του ζητήματος θα μπορούσε να εντοπιστεί κυρίως στη διατάραξη των σχέσεων της καλής γειτονίας από τη γειτονική χώρα, μέσω της ιδεολογικής χρήσης των ιστορικών γεγονότων, που αλλάζουν την οπτική απέναντι στην κοινωνική πραγματικότητα και την ιστορική κληρονομιά (Sjoeberg, 2011).

Για το λόγο αυτό και με αφορμή το γεγονός ότι η Ε.Ε. εμφανίζεται πιο σταθερή ως προς τη σημασία της περιοχής των Βαλκανίων και διατηρώντας πλέον μεγαλύτερη συνοχή ως προς την απόφαση της διεύρυνσής της στην ευρύτερη περιοχή (Illik & Gjurovski, 2014), η σημασία της «Συμφωνίας των Πρεσπών» είναι μεγάλη και ιδιαίτερα καθοριστική για την ασφάλεια και σταθερότητα της περιοχής.

Ωστόσο, η «Συμφωνία των Πρεσπών» αποτέλεσε σημείο έντονης πολιτικής τριβής στο εσωτερικό και την πολιτική ζωή τόσο της Ελλάδος όσο και της π.Γ.Δ.Μ., ενώ προκάλεσε ποικίλα σχόλια και αντιδράσεις. Στην Ελλάδα, έντονες αντιδράσεις προήλθαν και από μεγάλη μερίδα της κοινής γνώμης, η οποία αντέδρασε με κινητοποιήσεις και συλλαλητήρια. Ιδιαίτερη ανησυχία θα μπορούσε να προκαλέσει το γεγονός ότι μερίδα της κυβερνούσας παράταξης δηλώνει ξεκάθαρα τη μη υποστήριξη της συμφωνίας, όταν κατατεθεί για ψήφιση στη Βουλή.

Στη γειτονική χώρα και προς αντιμετώπιση των αντιδράσεων, ανακοινώθηκε τον Ιούλιο του 2018 η διεξαγωγή δημοψηφίσματος, το Σεπτέμβριο του 2018, με θέμα την υιοθέτηση κι εφαρμογή ή μη της συγκεκριμένης συμφωνίας, ενώ τον Οκτώβριο του 2018 οι προσπάθειες της κυβέρνησης του γειτονικού κράτους ήταν έντονες για τη διασφάλιση της πλειοψηφίας των βουλευτών κατά την κατάθεση του ζητήματος αλλαγής του Συντάγματος στη Βουλή.

Συνοψίζοντας, η συμφωνία μεμονωμένα δεν μπορεί να θεωρηθεί ακόμη ως η οριστική επίλυση του προβλήματος, ενώ πρέπει να ληφθεί σοβαρά υπόψη ότι τα γεγονότα εκτυλίσσονται κατά το τρέχον διάστημα και δεν έχει προκύψει τελικό αποτέλεσμα. Υπάρχουν ζητήματα προς διευθέτηση στο εσωτερικό και των δύο εμπλεκόμενων κρατών.

3 Ανάλυση συναισθήματος

Οι τρόποι με τους οποίους μπορεί να διεξαχθεί η ανάλυση συναισθήματος παρουσιάζουν τόσο κοινά σημεία όσο και σημαντικές διαφορές. Το βασικό κοινό σημείο όλων των μεθόδων με τις οποίες μπορεί να πραγματοποιηθεί ανάλυση συναισθήματος σε δεδομένα που έχουν αντληθεί από το διαδίκτυο είναι η χρήση υπολογιστικών εργαλείων, με σκοπό την ανάπτυξη μοντέλων ή συστημάτων τεχνητής νοημοσύνης. Επιπλέον, κοινό σημείο είναι η ανάγκη συμμετοχής του αναλυτή κατά την ανάπτυξη του αλγορίθμου ή των εργαλείων που χρησιμοποιούνται.

Ωστόσο, υπάρχουν διαφορές ανάλογα με τη μέθοδο που επιλέγεται για την ανάλυση και αυτές εντοπίζονται κυρίως στη χρήση στατιστικών μεθόδων ή άλλων μεθόδων, ενώ μπορεί να υπάρχουν και διαφορές στα δεδομένα που χρησιμοποιούνται για την εξαγωγή συμπερασμάτων.

Στο κεφάλαιο αυτό παρουσιάζονται συνοπτικά οι υπάρχουσες μέθοδοι ανάλυσης συναισθήματος ανά κατηγορία. Δηλαδή, παρουσιάζονται οι τρεις μέθοδοι μηχανικής μάθησης και η μέθοδος ανάλυσης συναισθήματος με τη χρήση λεξικού.

3.1 Μέθοδοι μηχανικής μάθησης

3.1.1 Μέθοδοι επιβλεπόμενης μάθησης

Η χρήση της μηχανικής μάθησης με επίβλεψη στην επίλυση προβλημάτων ταξινόμησης είναι ευρεία και στόχος της είναι η ορθή και αυτόματη λειτουργία ενός συστήματος ταξινόμησης που έχει δημιουργηθεί, γεγονός που επιτυγχάνεται με την εκπαίδευση ενός συστήματος για την αυτόματα επίλυση προβλημάτων.

Αρχικά, είναι ιδιαίτερα σημαντική για την εκπαίδευση του συστήματος η δημιουργία ενός συνόλου δεδομένων, το οποίο ονομάζεται σύνολο εκπαίδευσης (training set) και το οποίο δημιουργείται από τον αναλυτή με σαφή προσδιορισμό των εισροών και εκροών του συστήματος, καθώς και της συνάρτησης f που τα συσχετίζει. Στην ουσία, με βάση το σύνολο εκπαίδευσης, δημιουργείται το μοντέλο με τη χρήση του οποίου πρόκειται να κατηγοριοποιηθούν τα υπό διερεύνηση δεδομένα. Το σύστημα αυτόματα ορίζει μία συνάρτηση συναφή με τη συνάρτηση του μοντέλου που έχει δημιουργηθεί και μπορούν να προκύψουν

αποτελέσματα, κάθε φορά που η ορισμένη από το σύστημα συνάρτηση πλησιάζει τη συνάρτηση του μοντέλου.

Ωστόσο, υπάρχουν δύο περιπτώσεις που χρήσουν προσοχής για την επίτευξη της βέλτιστης αποτελεσματικότητας του μοντέλου. Στην πρώτη περίπτωση, για να μην αποτύχει το σύστημα και για να έχουν ουσία τα αποτελέσματα, το σύνολο εκπαίδευσης δεν πρέπει να είναι πολύ μικρό και στη δεύτερη περίπτωση, για να μην είναι πολύ αργό στην παραγωγή αποτελεσμάτων, το σύνολο εκπαίδευσης δεν πρέπει να είναι πολύ μεγάλο.

Τα κύρια βήματα της διαδικασίας για την επίλυση προβλημάτων συμπεριλαμβάνουν: α) το σαφή καθορισμό και τη δημιουργία από τον αναλυτή των παραδειγμάτων που θα χρησιμοποιηθούν κατά την εκπαίδευση του συστήματος, β) τη μορφή που θα έχει η συνάρτηση εισροών, δεδομένου ότι εξαρτάται άμεσα από τη μορφή των εισροών του συστήματος, γ) το σχεδιασμό της δομής των συναρτήσεων και αλγορίθμων και την εκπαίδευσή τους καθώς και δ) την αξιολόγηση της αξιοπιστίας του μοντέλου. Οι συνηθέστερες προσεγγίσεις της επιβλεπόμενης μάθησης είναι οι μέθοδοι με τη χρήση του κατηγοριοποιητή Naïve Bayes και της μεθόδου με τη χρήση Διανυσμάτων Στήριξης (Support Vector Machine - SVM) (Benedetto & Tedeschi, 2016).

3.1.2 Μέθοδοι μη επιβλεπόμενης μάθησης

Σε περιπτώσεις μεγάλου όγκου δεδομένων (big data) παρατηρείται μία τάση υπέρ των μεθόδων μη επιβλεπόμενης μάθησης έναντι των μεθόδων επιβλεπόμενης μάθησης (Benedetto & Tedeschi, 2016). Ιδιαίτερα σε αυτές τις περιπτώσεις μεγάλου όγκου δεδομένων και όταν γίνεται χρήση μεθόδων επιβλεπόμενης μάθησης, ο χρόνος εκπαίδευσης του μοντέλου, η ποσότητα και ποιότητα των υπό διερεύνηση δεδομένων μπορούν να αποτελέσουν προβλήματα που επηρεάζουν σημαντικά την ανάλυση. Για την αντιμετώπιση των προβλημάτων αυτών αναπτύχθηκαν οι μέθοδοι μη επιβλεπόμενης μάθησης.

Στις μεθόδους μη επιβλεπόμενης μάθησης που έχουν αναπτυχθεί, στόχος είναι ο εντοπισμός συσχετίσεων μεταξύ φαινομενικά άσχετων δεδομένων και κυρίως λέξεων. Οι συσχετισμοί αυτοί μπορούν να συμβάλλουν τόσο στη λήψη αποφάσεων όσο και στην κατηγοριοποίηση δεδομένων. Επιπλέον, η χρήση μεθόδων μη επιβλεπόμενης μάθησης μπορεί να δώσει αποτελέσματα χωρίς να είναι απαραίτητη η εκπαίδευση του συστήματος. Ωστόσο, ιδιαίτερη προσοχή χρειάζεται το γεγονός ότι τα αποτελέσματα που θα προκύψουν μπορεί να μην έχουν συνοχή, λαμβάνοντας υπόψη το γεγονός ότι τα δεδομένα που χρησιμοποιήθηκαν δεν είναι κατηγοριοποιημένα (labeled data).

Οι μέθοδοι μη επιβλεπόμενης μάθησης μπορούν είτε να βασίζονται σε στατιστικά μοντέλα που εντοπίζουν την κατανομή πιθανοτήτων λέξεων ή φράσεων μέσα στο κείμενο, είτε σε μοντέλα πιθανοτήτων που εντοπίζουν εσωτερικές στατιστικές δομές μέσα στο κείμενο. Δύο από τις πιο διαδεδομένες μεθόδους μη επιβλεπόμενης μάθησης είναι οι: α) pLSI (probabilistic Latent Semantic Indexing) σύμφωνα με την οποία κάθε κείμενο χαρακτηρίζεται από τα θέματα που εμπεριέχει και την κατανομή πιθανοτήτων κάθε θέματος βάσει συγκεκριμένου λεξικού, και β) LDA (Latent Dirichlet allocation model) που σύμφωνα με την κατανομή των θεμάτων μέσα στο κείμενο ελέγχει τις παραμέτρους κάθε θέματος (Wang & Zhai, 2017).

3.1.3 Μέθοδοι ημι-επιβλεπόμενης μάθησης

Μία εναλλακτική μέθοδος που, σε σχέση με τις μεθόδους επιβλεπόμενης και μη επιβλεπόμενης μάθησης, μπορεί να θεωρηθεί και νέα προσέγγιση στη διαδικασία ανάλυσης συναισθήματος είναι η μέθοδος ημι-επιβλεπόμενης μάθησης. Η ανάπτυξη της μεθόδου επιτυγχάνεται με την εκπαίδευση του συστήματος μέσω χρήσης τόσο κατηγοριοποιημένων (labeled data) όσο και μη κατηγοριοποιημένων δεδομένων (unlabeled data).

Συγκεκριμένα, κατά την εκπαίδευση του μοντέλου δεν δίνονται στο σύστημα όλα τα αποτελέσματα, ενώ τα δεδομένα εισόδου που χρησιμοποιούνται δεν κατηγοριοποιούνται πλήρως. Συνήθως πρόκειται για μεγάλο ποσοστό κατηγοριοποιημένων δεδομένων, αλλά εισάγονται και μη κατηγοριοποιημένα δεδομένα. Η χρήση του μοντέλου που δεν έχει εκπαιδευτεί πλήρως σε συνδυασμό με τα διαφορετικά δεδομένα εισόδου, βελτιώνει σημαντικά την ακρίβεια του μοντέλου.

Οι μέθοδοι που ανήκουν στην κατηγορία της ημι-επιβλεπόμενης μάθησης μπορούν να χωριστούν σε δύο κατηγορίες. Στην πρώτη κατηγορία το μοντέλο εκπαιδεύεται μόνο του κατά την επεξεργασία των δεδομένων, ενώ στη δεύτερη κατηγορία παίζει σημαντικό ρόλο η καλή δομή του μοντέλου παράλληλα με τη χρήση μη κατηγοριοποιημένων δεδομένων (Benedetto & Tedeschi, 2016).

3.2 Ανάλυση συναισθήματος με τη χρήση λεξικού

Μία επιπλέον διαδεδομένη μέθοδος που χρησιμοποιείται στην ανάλυση συναισθήματος είναι η μέθοδος της ανάλυσης με χρήση ενός λεξικού συναισθήματος (sentiment lexicon), το οποίο διαφέρει σημαντικά από τα συνηθισμένα. Η διαδικασία δημιουργίας του λεξικού συναισθήματος μπορεί γίνει με την εφαρμογή τριών διαφορετικών προσεγγίσεων: α) προσέγγιση βασισμένη στη σημασιολογία των λέξεων, β) προσέγγιση βασισμένη στη βάση δεδομένων που έχει δημιουργηθεί για την ανάλυση και γ) προσέγγιση βασισμένη σε συνδυαστικές εκφράσεις (Agathangelou, Katakis, Rori, Gunopulos, & Richards, 2017).

Στη συγκεκριμένη περίπτωση, το λεξικό συναισθήματος, δηλαδή η λίστα λέξεων με βάση την οποία διεξάγεται η ανάλυση, βασίζεται στο συναίσθημα και την πολικότητα που εμπερικλείει κάθε λέξη ως οντότητα και στη συνέχεια καθορίζει το συναίσθημα και την πολικότητα που προκύπτει από ένα ολόκληρο κείμενο, αθροίζοντας το συναίσθημα όλων των λέξεων του κειμένου που αναλύεται, συνυπολογίζοντας την αξία λέξεων που είτε υποδηλώνουν άρνηση είτε ενισχύουν το συναίσθημα των λέξεων. Το συνολικό τελικό άθροισμα που προκύπτει από τις λέξεις ενός κειμένου καθορίζει και την τελική πολικότητα, θετική ή αρνητική, του κειμένου συνολικά (Benedetto & Tedeschi, 2016).

Δύο λεξικά συναισθήματος που χρησιμοποιούνται κυρίως και αναφέρονται στην Αγγλική γλώσσα είναι τα SentiWordNet και WordNet-Affect, που επικεντρώνονται και τα δύο κυρίως

λέξεις. Ωστόσο, είναι σημαντικό να αναφερθεί ότι η χρήση των υφιστάμενων λεξικών συναισθήματος υπόκειται σε κάποιους περιορισμούς, όπως το γεγονός ότι είναι ανεξάρτητα από το υπο-διερεύνηση θέμα κι αυτό μπορεί να επηρεάζει την αποτελεσματικότητά τους σε συγκεκριμένα θέματα. Επιπλέον, οι ίδιες οι λέξεις μπορούν να επηρεάσουν το αποτέλεσμα, καθώς μπορεί να περιλαμβάνονται σε ένα λεξικό χωρίς στην πραγματικότητα να εκφράζουν συναίσθημα ή ακόμη και να εκφράζουν αντίθετα συναισθήματα σε διαφορετικά θέματα (Biagioni, 2016).

Κατά τη διαδικασία ανάλυσης με τη συγκεκριμένη μέθοδο, αποτελεί πρόκληση η αντιμετώπιση προβλημάτων που σχετίζονται με λέξεις της καθομιλουμένης (αργκό) ή τύπους λέξεων που δεν υφίστανται και προκαλούν «θόρυβο», δεδομένου ότι δεν εντοπίζονται αντιστοιχίες με λέξεις που έχουν καθοριστεί στα λεξικά συναισθήματος. Για την αντιμετώπιση των εν λόγω προβλημάτων είναι σημαντική η παράλληλη χρήση υφιστάμενων λεξικών συναισθήματος και εξειδικευμένου λεξικού που δημιουργείται από τον αναλυτή για το εκάστοτε υπό διερεύνηση θέμα. Ιδιαίτερο ρόλο για τη δημιουργία του εξειδικευμένου λεξικού παίζει το νόημα της εκάστοτε λέξης με βάση τα συμφραζόμενα για το θέμα, καθώς και οι φορές που η λέξη έχει εμφανιστεί σε φράσεις θετικής, αρνητικής ή ουδέτερης πολικότητας. Τα κύρια γραμματικά στοιχεία που χρησιμοποιούνται για τη δημιουργία του εξειδικευμένου λεξικού είναι τα ρήματα, τα επίθετα, τα επιρρήματα και τα ουσιαστικά .

Η διπλή χρήση υφιστάμενων λεξικών συναισθήματος και εξειδικευμένου λεξικού, ειδικά για την ελληνική γλώσσα, πιθανώς να μπορεί να βοηθήσει την αντιμετώπιση προβλημάτων που προκύπτουν από τις ιδιαιτερότητες και το λεκτικό πλούτο της ελληνικής γλώσσας, ενώ μπορεί να επιλύσει προβλήματα φράσεων που εμπεριέχουν σαρκασμό ή ειρωνεία.

3.3 Προκλήσεις στην ανάλυση συναισθήματος

Δεδομένου του γεγονότος ότι στην ανάλυση συναισθήματος κύριο στοιχείο έρευνας είναι η επεξεργασία της φυσικής γλώσσας (Natural Language Processing), είναι συχνό φαινόμενο να πρέπει να αντιμετωπιστούν πολλές ιδιομορφίες και προκλήσεις που σχετίζονται κυρίως με το περιεχόμενο των υπό ανάλυση δεδομένων.

Σημαντική πρόκληση αποτελεί το ίδιο το κείμενο, αφού του συναίσθημα και οι πολικότητες εξαρτώνται συχνά από το περιεχόμενο και τις εκφράσεις που χρησιμοποιούνται. Οι λέξεις που εκφράζουν άρνηση χρήζουν ιδιαίτερης προσοχής, καθώς η χρήση τους συνήθως αντιστρέφει το συναίσθημα λέξεων ή φράσεων, ενώ υπάρχει το ενδεχόμενο της χρήσης άρνησης είτε σε ιδιωτισμούς είτε σε εκφράσεις σαρκασμού και ειρωνείας, γεγονός που δύσκολα εντοπίζεται κατά την ανάλυση συναισθήματος.

Επιπροσθέτως, υπάρχουν κάποιες ιδιαιτερότητες στο περιεχόμενο κειμένων ή φράσεων, όπως λέξεις που μπορεί να τονίζουν ή να μειώνουν το συναίσθημα ή ιδιωτισμοί εκφράσεων που χρησιμοποιούν ουδέτερες λέξεις χωρίς συναίσθημα και από τις οποίες τελικά προκύπτει συναίσθημα. Ακόμη, δεν είναι σπάνιο το ενδεχόμενο να εντοπίζονται μέσα σε κείμενα φράσεις από τις οποίες προκύπτουν διαφορετικά συναισθήματα, καθώς και να προκύπτει αμφιβολία ως προς το τελικό συναίσθημα που απορρέει κατά την ανάγνωση μίας φράσης ή ενός κειμένου.

Οι ανωτέρω προκλήσεις μπορούν να αντιμετωπιστούν εύκολα από έναν άνθρωπο, αλλά μπορεί να είναι ιδιαίτερα σοβαρές και να μην μπορούν να αντιμετωπιστούν το ίδιο εύκολα από τις μεθόδους μηχανικής μάθησης. Συγκεκριμένα, οι μέθοδοι μηχανικής μάθησης με επίβλεψη έχουν το πλεονέκτημα μεν της επίτευξης αποτελεσμάτων μεγαλύτερης ακρίβειας στην ανάλυση συναισθήματος, ωστόσο καταναλώνουν σημαντικά πολύ χρόνο.

Αντίθετα, οι μέθοδοι μηχανικής μάθησης χωρίς επίβλεψη, που επικεντρώνονται στην ανάλυση συναισθήματος σε επίπεδο λέξεων ή φράσεων, μπορούν να υλοποιήσουν την ανάλυση πιο γρήγορα (Biagioni, 2016), ενώ πρέπει να ληφθεί υπόψη και το γεγονός ότι οι μέθοδοι μηχανικής μάθησης σχεδιάζονται κι εκπαιδεύονται για συγκεκριμένους τομείς και συχνά παρουσιάζονται αστοχίες και δυσκολίες κατά την προσπάθεια εφαρμογής τους σε διαφορετικούς τομείς (Fernandez-Gavilanez, Juncal-Martinez, Garcia-Mendez, Costa-Montenegro, & Gonzalez-Castano, 2018).

4 Μεθοδολογία

Λαμβάνοντας υπόψη την αναφορά στις διάφορες μεθόδους ανάλυσης συναισθήματος, η μέθοδος ανάλυσης με τη χρήση λεξικού αποτελεί μια ενδιαφέρουσα πρόκληση και για το λόγο αυτό έχει επιλεγεί για την ολοκλήρωση της ανάλυσης στην παρούσα εργασία. Η ανάλυση πρόκειται να ολοκληρωθεί εξολοκλήρου με τη χρήση του συστήματος KNIME σε όλα τα στάδια: συλλογή δεδομένων, δημιουργία του λεξικού, προεπεξεργασία, επεξεργασία δεδομένων, ανάλυση και κατηγοριοποίηση δεδομένων.

Στο σημείο αυτό, κρίνεται σκόπιμο να σημειωθεί ότι πληροφορίες που περιέχουν κυρίως πολιτικό περιεχόμενο, μπορούν να γίνουν κατανοητές μόνο σαν μέρος ενός ευρύτερου συνόλου σχέσεων, αλληλεπιδράσεων, ιστορικότητας και γεγονότων που έχουν λάβει χώρα σε ένα πολιτικό σύστημα, γεγονός που δεν επιτρέπει τον προσδιορισμό τους απλά ως μία σχέση μεταξύ του κατόχου της γνώμης και του παραλήπτη (Guidi, Ruiz-Agundez, & Canga-Sanchez, 2012). Αυτή η ιδιαιτερότητα λαμβάνεται υπόψη στη μεθοδολογία της ανάλυσης στην παρούσα εργασία.

Στη συνέχεια του κεφαλαίου αυτού παρουσιάζονται το σύστημα αναλύσεων KNIME και τα αναλυτικά τα βήματα που χρειάστηκαν κατά τη διαδικασία ανάπτυξης του μοντέλου της παρούσας έρευνας. Ιδιαίτερα αναλύονται τα στάδια της συλλογής δεδομένων, της επεξεργασίας τους, καθώς και η μεθοδολογία ανάπτυξης και υλοποίησης του λεξικού συναισθήματος που δημιουργήθηκε και χρησιμοποιήθηκε για τους σκοπούς της παρούσας ανάλυσης. Τέλος γίνεται αναφορά στην αυτοματοποιημένη μέθοδο κατηγοριοποίησης των κειμένων.

4.1 Σύστημα αναλύσεων KNIME

Το σύστημα αναλύσεων KNIME είναι κατά βάση μία πλατφόρμα διαθέσιμη στο διαδίκτυο που έχει ως κύριο σκοπό λειτουργίας να διευκολύνει το χρήστη στη διεξαγωγή ανάλυσης δεδομένων κάθε τύπου. Προσφέρεται για πολλές λειτουργίες ανάλυσης, δεδομένου ότι μπορεί να χρησιμοποιηθεί με πολλαπλούς τρόπους. Όλες οι λειτουργίες παρέχονται με τη μορφή κόμβων, που εκτελούν συγκεκριμένη προγραμματισμένη λειτουργία ο καθένας.

Ο χρήστης μπορεί, κατά πρώτον, να συνθέσει ένα δικό του σύστημα ανάλυσης (workflow) με τη χρήση επί μέρους κόμβων, ή, κατά δεύτερον, να κάνει χρήση έτοιμων workflow που παρέχονται και υποστηρίζονται από την ιστοσελίδα του συστήματος KNIME.

Οι λειτουργίες του συστήματος περιλαμβάνουν μεταξύ άλλων την ανάγνωση κι εξόρυξη δεδομένων από διαφορετικές πηγές, την επεξεργασία τους με διάφορους τρόπους, όπως π.χ. με μεθόδους στατιστικής ανάλυσης, τη διαχείριση βάσεων δεδομένων, την επεξεργασία κειμένων και δομημένων δεδομένων (structured data), εξόρυξη δεδομένων από μέσα κοινωνικής δικτύωσης, τη σύνταξη αναφορών, καθώς προσφέρει και δυνατότητες απεικόνισής τους.

Η ευκολία της χρήσης του συστήματος, πέραν του εύρους των επιλογών που προαναφέρθηκαν, έγκειται στο γεγονός ότι δεν απαιτεί πολύπλοκες γνώσεις προγραμματισμού του χρήστη και το workflow μπορεί να σχεδιαστεί εύκολα, απλά με τη χρήση κόμβων που συνδέονται μεταξύ τους και στους οποίους καθορίζονται από το χρήστη οι όροι με τους οποίους επιθυμεί να εκτελεστεί η επεξεργασία των δεδομένων.

Ιδιαίτερα σε περιπτώσεις όπως η ανάλυση συναισθήματος, στις οποίες ο χρήστης θέτει συνεχώς όρους προσαρμοσμένους στο υπό διερεύνηση θέμα με σκοπό τη μεγιστοποίηση της αποτελεσματικότητας του συστήματος, το KNIME μπορεί να αποτελέσει πολύτιμο εργαλείο. Ωστόσο, έχουν αναπτυχθεί πολλά συστήματα αναλύσεων με παρόμοια φιλοσοφία, τα οποία μπορούν να χρησιμοποιηθούν για τη διεξαγωγή ανάλυσης δεδομένων. Μερικά ευρέως γνωστά συστήματα αναλύσεων είναι τα εξής: Python, R, Apache Spark, Splunk, RapidMiner, Pentaho, Sisense, Stata και πολλά άλλα δεδομένης της αυξανόμενης ανάγκης των χρηστών για την ανάλυση δεδομένων μεγάλου όγκου.

4.2 Συλλογή δεδομένων

Οι ειδησεογραφικές ιστοσελίδες μπορούν να αποτελέσουν εξαιρετική πηγή ανάλυσης συναισθήματος δεδομένων για θέματα σχετικά με την πολιτική, οικονομική και κοινωνική επικαιρότητα. Το περιεχόμενό τους απεικονίζεται συνήθως σε καθαρή γλώσσα τόσο γραμματικά όσο και συντακτικά, με ενσωματωμένη την προσωπική άποψη των συντακτών των κειμένων, τα οποία είναι διαθέσιμα σε όλους.

Η χρήση της μεθόδου RSS προσφέρει τη δυνατότητα της συνεχούς ανανέωσης του περιεχομένου και της άμεσης δημοσιοποίησής του, ενώ μέσω αυτού παρέχεται η δυνατότητα αναζήτησης κειμένων συγκεκριμένης κατηγορίας ανάλογα με το ενδιαφέρον του αναλυτή, αν η επιλογή αυτή είναι δυνατή μέσω της ιστοσελίδας. Η μέθοδος αυτή υποστηρίζεται από τη γλώσσα σήμανσης XML, που επιτρέπει σχετικά εύκολα την εξόρυξη των δεδομένων για το υπό διερεύνηση θέμα.

Ιδιαίτερα σημαντικό γεγονός για τους σκοπούς της έρευνας ήταν η σωστή δημιουργία της βάσης δεδομένων που χρησιμοποιήθηκε για την ανάλυση συναισθήματος. Για τη δημιουργία της βάσης δεδομένων, πληκτρολογήθηκαν στο σύστημα οι διευθύνσεις URL των πηγών RSS και επιλεγμένων κατηγοριών, με σκοπό τη συλλογή κειμένων. Εντοπίστηκαν στη γλώσσα XML τα εξής στοιχεία: Τίτλος, Ημερομηνία, Κείμενο, Συντάκτης και Κατηγορία. Τα κείμενα, κυρίως άρθρα, συλλέχθηκαν σε καθημερινή βάση για διάστημα περίπου ενός μηνός από συγκεκριμένες ιστοσελίδες, οι οποίες επιλέχθηκαν κατόπιν έρευνας σχετικής με τη δυνατότητα

παροχής της υπηρεσίας RSS. Σε πρώτο χρόνο εντοπίστηκαν οι ειδησεογραφικές ιστοσελίδες που παρείχαν τη δυνατότητα RSS και σε δεύτερο χρόνο ξεχώρισαν αυτές, των οποίων το RSS λειτουργούσε σωστά και επέτρεπε τη συλλογή δεδομένων.

Η χρήση της πλατφόρμας KNIME σε αυτό το στάδιο ήταν καθοριστική, δεδομένης της ευελιξίας που παρέχει για την επιτυχή ολοκλήρωση τριών σημαντικών βημάτων: α) τη συνεχή πρόσβαση στις πηγές δεδομένων, β) περιοδική ανανέωση του μοντέλου και αντιμετώπιση αστοχιών άμεσα και γ) τη δυνατότητα άμεσης επεξεργασίας των δεδομένων (Kranjc, Smailovic, Podpecan, Grear, Znidarsic, & Lavrac, 2015).

Μία από τις προκλήσεις της παρούσας έρευνας ήταν η συλλογή της πεντάδας των βασικών στοιχείων του εκάστοτε κειμένου που προαναφέρθηκε, δεδομένου ότι δεν ήταν εύκολος ο εντοπισμός των δεδομένων στη γλώσσα XML σε όλες τις πηγές, γεγονός που οδήγησε στο συμπέρασμα ότι δεν είναι σωστά προγραμματισμένα τα πεδία σε όλα τα RSS. Σε πολλές περιπτώσεις χρειάστηκε αρκετός χρόνος για τον εντοπισμό των στοιχείων αυτών.

Για την ορθή διεξαγωγή ανάλυσης συναισθήματος, αποτελεί ιδιαίτερης σημασίας παράγοντα η συλλογή των βασικών στοιχείων κάθε κειμένου, τα οποία είναι:

- e (entity), που αναφέρεται στον τίτλο και στο σώμα του εκάστοτε κειμένου,
- a (aspect), που αναφέρεται στην κατηγορία στην οποία ανήκει το εκάστοτε κείμενο,
- h (holder), που αναφέρεται στο συντάκτη του εκάστοτε κειμένου και κατά συνέπεια στον κάτοχο της γνώμης που τελικά προκύπτει από την επεξεργασία των δεδομένων,
- s (sentiment), που αναφέρεται στο συναίσθημα του εκάστοτε κειμένου κατά την ανάλυση,
- t (time), που αναφέρεται στο χρόνο δημοσίευσης του εκάστοτε κειμένου.

Κάθε στοιχείο μόνο του δεν παρέχει επαρκή πληροφόρηση για την ανάλυση συναισθήματος. Κατ' ελάχιστο, η πεντάδα των ανωτέρω βασικών στοιχείων παρέχει επαρκή πληροφόρηση για το κάθε κείμενο ως οντότητα και θεωρείται σημαντική στην ανάλυση συναισθήματος (Liu, 2017). Επιπλέον, η πεντάδα των βασικών στοιχείων αποτελεί μία πολύ καλή βάση χρήσιμων πληροφοριών στη διαδικασία της ανάλυσης συναισθήματος με οποιαδήποτε μέθοδο κι αν διεξαχθεί (Taylor, Rodriguez, Velasquez, Ghosh, & Banerjee, 2013).

Στη συνέχεια, μετά τη συλλογή τους, τα κείμενα αποθηκεύονται ως πίνακες, παρέχοντας όλη την πληροφόρηση που χρειάζεται για την ανάλυση, οι οποίοι, μετά το πέρας της περιόδου συλλογής, συμπύκνθηκαν σε ένα τελικό πίνακα για κάθε ιστοσελίδα, ενώ το σύνολο των κειμένων που συλλέχθηκαν αποτελεί και τη βάση δεδομένων της έρευνας.

Σημεία ιδιαίτερης προσοχής και για λόγους ομοιομορφίας και αποτελεσματικότητας κατά τη συλλογή δεδομένων ήταν η επεξεργασία της ημερομηνίας δημοσίευσης των κειμένων, καθώς προέκυψε ότι κάθε ιστοσελίδα την απεικονίζει με διαφορετική μορφοποίηση, ενώ σημαντικά ήταν το φιλτράρισμα των κειμένων με βάση τον τίτλο τους και η αφαίρεση των πολλαπλών εγγραφών των κειμένων, ώστε να μην περιέχονται στους πίνακες περισσότερες από μία φορές.

Μετά τη δημιουργία του σώματος των προς επεξεργασία δεδομένων, έγινε επιλογή των κειμένων που περιλάμβαναν τον όρο «Πρέσπες», ο οποίος ήταν και η ορθότερη επιλογή από ένα σύνολο όρων όπως: «Μακεδονικό», «ΠΓΔΜ», «Σκόπια», «Συμφωνία», «Σκοπιανό» κτλ. Όλοι οι προαναφερθέντες όροι ελέγχθηκαν ως προς το κατά πόσο έδιναν αποτελέσματα κειμένων και άρθρων σχετικών με το θέμα της «Συμφωνίας των Πρεσπών», που είναι και το κύριο θέμα ενδιαφέροντος για το οποίο διεξάγεται η ανάλυση. Με τη χρήση των υπόλοιπων

όρων, προέκυπταν αποτελέσματα κειμένων και άρθρων τόσο σχετικών με τη «Συμφωνία των Πρεσπών», όσο και γενικού πολιτικού ενδιαφέροντος. Αντίθετα, ο τελικά επιλεγμένος όρος απέδωσε τα μέγιστα αποτελέσματα κειμένων και άρθρων συναφών με το θέμα. Στο τέλος, εξετάστηκε δείγμα των κειμένων ως προς της συνάφειά τους με τον όρο, με σκοπό τον έλεγχο της ορθής λειτουργίας του συστήματος.

4.3 Προεπεξεργασία δεδομένων

Κατά τη διαδικασία του σταδίου συλλογής δεδομένων, η βάση δεδομένων που δημιουργήθηκε περιλαμβάνει τα κείμενα αυτούσια, χωρίς καμία επεξεργασία και παρεμβολή στο κυρίως κείμενο. Η χρήση των κειμένων στην ανάλυση χωρίς καμία επεξεργασία μπορεί να επηρεάσει την ορθή διεξαγωγή αποτελεσμάτων.

Για τη βελτιστοποίηση της αποδοτικότητας της ανάλυσης συναισθήματος και την αποφυγή αστοχιών στη λειτουργία της διαδικασίας ανάλυσης που δημιουργείται, κρίνεται απαραίτητη η προεπεξεργασία (preprocessing) και ο καθαρισμός των δεδομένων από τυχόν παρεμβολές και «θορύβους» που μπορεί να υπάρχουν. Ως στοιχεία που μπορεί να αποτελέσουν πρόβλημα και να οδηγήσουν σε λανθασμένα συμπεράσματα κατά την ανάλυση με , ενδεικτικά αναφέρονται οι συντομογραφίες και οι ειδικοί χαρακτήρες, οι αριθμοί και η διαφορά μεταξύ πεζών και κεφαλαίων χαρακτήρων. Το στάδιο της προεπεξεργασίας συμβάλλει σημαντικά, πέρα από τη βελτιστοποίηση της απόδοσης του συστήματος, στην ταχύτερη κατηγοριοποίηση των κειμένων (McKenna, Myers, & Newman, 2017).

Τα βήματα που ακολουθήθηκαν στο στάδιο της προεπεξεργασίας για την παρούσα έρευνα, είναι τα εξής:

1. Αφαίρεση αριθμών, δεδομένου ότι η χρήση τους στο κείμενο δεν συνεισφέρει στην εξόρυξη γνώμης και ανάλυσης του συναισθήματος του κειμένου.
2. Αφαίρεση των σημείων στίξης του κειμένου, δεδομένου ότι η χρήση τους δεν συνεισφέρει στην εξόρυξη γνώμης. Ωστόσο, σε άλλες περιπτώσεις όπως η εξόρυξη γνώμης από μηνύματα χρηστών κοινωνικών δικτύων, θα μπορούσαν να αποτελέσουν δείκτες της έντασης του συναισθήματος του συγγραφέα του εκάστοτε μηνύματος.
3. Αφαίρεση λέξεων με συγκεκριμένο αριθμό χαρακτήρων (N-chars), οι οποίες συνήθως είναι πολύ μικρές λέξεις, όπως για παράδειγμα άρθρα και προθέσεις, που δεν εκφράζουν συναίσθημα και δεν συνεισφέρουν στην ανάλυση. Συγκεκριμένα, για την περίπτωση της παρούσας ανάλυσης και δεδομένων των ιδιαιτεροτήτων της ελληνικής γλώσσας, επιλέχθηκε η αφαίρεση λέξεων που περιλαμβάνουν έως και τέσσερις χαρακτήρες.
4. Μετατροπή κεφαλαίων χαρακτήρων σε πεζούς, γιατί είναι σημαντική η ομοιομορφία όλων των χαρακτήρων πριν την τελική τους επεξεργασία.

Επιπλέον, θα ήταν χρήσιμο για την ανάλυση να αφαιρεθούν όλες οι καταλήξεις των λέξεων και να διατηρηθεί μόνο το θέμα κάθε λήμματος (stemming). Στην ανάλυση συναισθήματος κειμένων στην Αγγλική ή την Ισπανική γλώσσα έχουν αναπτυχθεί και προσφέρεται η δυνατότητα χρήσης συγκεκριμένων έτοιμων εφαρμογών, π.χ. Porter stemmer (Ryan, 2013). Ωστόσο, στην περίπτωση της ελληνικής γλώσσας δεν υπάρχει ακόμη αντίστοιχο έτοιμο

εργαλείο, γεγονός που αποτελεί μία πρόκληση, και θα έπρεπε να δημιουργηθεί με χρήση κάποιας γλώσσας προγραμματισμού μία σχετική εφαρμογή αποκλειστικά για την παρούσα ανάλυση.

4.4 Δημιουργία λεξικού συναισθήματος

Τα λεξικά συναισθήματος διαφέρουν από τα τυπικά, γιατί εκφράζουν το συναίσθημα που προκύπτει και εξάγεται από την ίδια τη λέξη. Το συναίσθημα αυτό, που μπορεί να εμπεριέχεται σε μία λέξη, μπορεί να είναι δύο διαστάσεων: απόλυτο, εκφράζοντας ένα συναίσθημα που παραμένει ίδιο κάθε φορά που χρησιμοποιείται η λέξη, και σχετικό, εκφράζοντας διαφορετικό συναίσθημα ανάλογα με τα συμφραζόμενα και τις εκφράσεις που χρησιμοποιούνται στο κείμενο.

Ιδανική περίπτωση για τη βελτιστοποίηση των αποτελεσμάτων της ανάλυσης είναι η από κοινού χρήση τόσο ενός υφιστάμενου λεξικού συναισθήματος όσο και η δημιουργία ενός λεξικού που θα επικεντρώνεται και θα αναφέρεται συγκεκριμένα στο υπό διερεύνηση θέμα (Mohammad, 2017).

Δεδομένης της ευρείας χρήσης της ανάλυσης συναισθήματος έχουν αναπτυχθεί ήδη αρκετά λεξικά συναισθήματος, τα οποία όμως έχουν επικεντρωθεί στην Αγγλική γλώσσα. Στην περίπτωση της ελληνικής γλώσσας, η δημιουργία λεξικών συναισθήματος δεν έχει αναπτυχθεί αρκετά και κατά συνέπεια κρίθηκε σκόπιμη η δημιουργία λεξικού συναισθήματος, σχετικού με πολιτικού θέματος λέξεις. Η πολικότητα των λέξεων που περιλαμβάνονται στο λεξικό διαχωρίζεται σε θετική και αρνητική.

Αρχικά, επιλέχθηκαν ως δείγμα κάποια από τα κείμενα που είχαν συλλεχθεί. Ακολούθησε το στάδιο της προεπεξεργασίας τους κι έπειτα διαχωρίστηκαν οι λέξεις των κειμένων του δείγματος, ώστε να δημιουργηθεί ένας «σάκος λέξεων» (Bag of Words). Από το «σάκο λέξεων» χρησιμοποιήθηκαν οι λέξεις από τις οποίες είναι πιο πιθανό και εφικτό να προσδιοριστεί πολικότητα ή συναίσθημα. Συνήθως, οι λέξεις που χρησιμοποιούνται κατά τη δημιουργία λεξικών συναισθήματος είναι: τα επίθετα, τα επιρρήματα, τα ουσιαστικά και τα ρήματα.

Κάθε λέξη κατηγοριοποιήθηκε σε πρώτη φάση ως προς το απόλυτο συναίσθημα που έβγαζε, ενώ σε δεύτερη φάση εντοπίστηκε το σχετικό συναίσθημα που προέκυπτε κατά την ανάγνωση του κειμένου στο οποίο ανήκε. Το σχετικό συναίσθημα είναι κι αυτό που επικράτησε τελικά και χρησιμοποιήθηκε στην ανάλυση. Το σχετικό συναίσθημα απεικονίζεται ως (-1) στην περίπτωση αρνητικής πολικότητας της λέξης και (1) στην περίπτωση θετικής πολικότητας της λέξης.

Στη συνέχεια, με τη βοήθεια και τη χρήση του ελληνικού λεξικού συναισθήματος που είναι διαθέσιμο στην πλατφόρμα Github, έγινε προσπάθεια διασταύρωσης του συναισθήματος των λέξεων που εντοπίστηκαν αρχικά.

Το πρώτο ενδιαφέρον συμπέρασμα είναι ότι οι λέξεις που εμφανίζονται τόσο στο λεξικό που δημιουργήθηκε όσο και στο υπάρχον λεξικό του διαδικτύου, συμφωνούν ως προς το

συναίσθημα και την πολικότητα τους. Το δεύτερο ενδιαφέρον συμπέρασμα είναι ότι η πλειοψηφία των λέξεων που περιλαμβάνονται στο λεξικό που δημιουργήθηκε για τους σκοπούς της παρούσας ανάλυσης, δεν έχει εξεταστεί ως προς την πολικότητα των λέξεων και δεν συμπεριλαμβάνεται στο υπάρχον λεξικό.

Με σκοπό την εξέταση της αποδοτικότητας του λεξικού συναισθήματος που δημιουργήθηκε, έγινε εκ νέου έλεγχος του αθροίσματος της τελικής πολικότητας του εκάστοτε κειμένου, που προέκυπτε από το άθροισμα της πολικότητας των λέξεων που εμπεριέχονται σε αυτό.

4.5 Τελική επεξεργασία δεδομένων

Τελευταίο στάδιο της διαδικασίας είναι η τελική επεξεργασία των δεδομένων με σκοπό την ανάλυση συναισθήματος σε επίπεδο κειμένου συνολικά. Έχοντας προετοιμάσει τη βάση δεδομένων και το λεξικό συναισθήματος, τα κείμενα κατηγοριοποιήθηκαν ως προς τον τίτλο τους και η πολικότητά τους καθορίστηκε από το άθροισμα της πολικότητας των λέξεων που περιλαμβάνονται στο κείμενο και στο λεξικό.

Ο έλεγχος δείγματος για την εξέταση της λειτουργικότητας και αποδοτικότητας του συστήματος και σε αυτό το στάδιο κρίθηκε απαραίτητος. Η απεικόνιση των αποτελεσμάτων έγινε σε ένα τελικό πίνακα που περιλαμβάνει το αρχείο με όλα τα βασικά στοιχεία κάθε κειμένου και την τελική τους πολικότητα.

Ωστόσο, για την καλύτερη ανάλυση των αποτελεσμάτων, χρησιμοποιήθηκε αυτόματος κατηγοριοποιητής (AutoBinner) μέσω του συστήματος KNIME, σύμφωνα με τον οποίο τα κείμενα διαχωρίζονται σε επίπεδα (bins) τα οποία προσδιορίζονται από τη μέγιστη και την ελάχιστη τιμή του αθροίσματος πολικότητας των κειμένων. Ο διαχωρισμός γίνεται σε όσα επίπεδα είναι χρήσιμα για τη διεξαγωγή της ανάλυσης, στην προκειμένη περίπτωση τρία (3), και το εύρος τους είναι περίπου ίσο.

5 Εμπειρικά αποτελέσματα

Ιδιαίτερη σημασία παρουσιάζει το κεφάλαιο αυτό, δεδομένου ότι παρουσιάζονται τα εμπειρικά αποτελέσματα που προέκυψαν κατά την ανάλυση και συμπληρώνει τη μεθοδολογία που έχει ήδη αναφερθεί. Αρχικά, αναλύονται τα στάδια της συλλογής δεδομένων, με σαφή αναφορά στο χρόνο κατά τον οποίο έγινε η ανάλυση, των γεγονότων, σχετικών με το υπό διερεύνηση θέμα, που έλαβαν χώρα και των περιορισμών που προέκυψαν κατά τη συλλογή, ενώ παρουσιάζονται και τα αποτελέσματα που προέκυψαν σε αυτό το στάδιο. Στη συνέχεια, αναλύονται τα στάδια της επεξεργασίας και υλοποίησης του λεξικού συναισθήματος και της επεξεργασίας των δεδομένων, ενώ παρουσιάζονται και τα τελικά ευρήματα.

5.1 Επιλογή δεδομένων και περίοδος αναφοράς

Τα πολιτικά γεγονότα παρουσιάζουν εξαιρετικό ενδιαφέρον για την επικαιρότητα και την ειδησεογραφία και κατ' επέκταση, μπορούν να αποτελέσουν την αφορμή για περαιτέρω ανάλυσή τους. Ειδικότερα, η ανάλυση συναισθήματος για ένα πολιτικό γεγονός, μπορεί να δείξει θετικές ή αρνητικές τάσεις που πιθανώς να έχουν κι επιρροή στην κοινή γνώμη.

Ως σημείο αναφορά της ανάλυσης, επιλέχθηκε η «Συμφωνία των Πρεσπών», που υπεγράφη μεταξύ των χωρών της Ελλάδος και της πρώην Γιουγκοσλαβικής Δημοκρατίας της Μακεδονίας (π.Γ.Δ.Μ.) την 12^η Ιουνίου του έτους 2018. Η συμφωνία αυτή, όπως ήδη αναφέρθηκε, ήταν καθοριστική στην επίλυση ενός χρόνιου προβλήματος μεταξύ των δύο χωρών.

Την υπογραφή της συμφωνίας ακολούθησε η τελική επικύρωση από το Κοινοβούλιο της π.Γ.Δ.Μ. τον Ιούλιο του 2018 και στη συνέχεια, στις 30 Ιουλίου, η ανακοίνωση της κυβέρνησης της γειτονικής χώρας για τη διεξαγωγή δημοψηφίσματος κατά την 30^η Σεπτεμβρίου του 2018.

Κατά το φθινόπωρο του 2018, έλαβαν χώρα σημαντικά γεγονότα που σχετίζονταν άμεσα με τη «Συμφωνία των Πρεσπών» και κρίθηκαν σημαντικά για το πλαίσιο της ανάλυσης. Στις 30 Σεπτεμβρίου 2018 ολοκληρώθηκε το δημοψήφισμα στη γειτονική χώρα, με το αποτέλεσμα να αναδεικνύει την υποστήριξη της συμφωνίας από τους ψηφίσαντες. Στις 15 Οκτωβρίου ξεκίνησαν οι συζητήσεις στο Κοινοβούλιο της π.Γ.Δ.Μ. για συνταγματική αναθεώρηση με σκοπό την ενσωμάτωση των όρων της συμφωνίας στο Σύνταγμα της π.Γ.Δ.Μ., που κατέληξαν

σε ψήφιση της πρότασης τροποποίησης του Συντάγματος της γειτονικής χώρας στις 19 Οκτωβρίου 2018. Παράλληλα, στην Ελλάδα την 17^η Οκτωβρίου 2018, ανακοινώνεται η παραίτηση του Υπουργού Εξωτερικών Υποθέσεων, κυρίου Νίκου Κοτζιά, ενός εκ των ένθερμων υποστηρικτών της συμφωνίας.

Αρχικά, θεωρήθηκε σημαντική περίοδος για την ανάλυση το διάστημα του Σεπτεμβρίου γύρω από την ημερομηνία διεξαγωγής του δημοψηφίσματος. Η συλλογή δεδομένων ξεκίνησε στις 27 Σεπτεμβρίου 2018. Παρόλα αυτά και δεδομένων των ανωτέρω, επεκτάθηκε το διάστημα συλλογής δεδομένων έως και την 31^η Οκτωβρίου, επειδή κρίθηκε σημαντικό κι ενδιαφέρον να συμπεριληφθεί αρθρογραφία που ακολούθησε τα γεγονότα που αναφέρθηκαν.

5.2 Ανάλυση συναισθήματος με τη χρήση λεξικού

5.2.1 Συλλογή δεδομένων

Για τους σκοπούς της ανάλυσης και λαμβάνοντας υπόψη ότι το υπό διερεύνηση ερώτημα είναι η ανάλυση συναισθήματος στην επικαιρότητα, το αρχικό σημείο για την συλλογή δεδομένων ήταν ο εντοπισμός των ειδησεογραφικών ιστοσελίδων από τις οποίες θα γίνει η εξόρυξη των κειμένων μέσω του RSS.

Κατόπιν έρευνας στο διαδίκτυο και εντοπισμού σημαντικού αριθμού ειδησεογραφικών ιστοσελίδων, επιλέχθηκαν 17 ιστοσελίδες που παρείχαν τη δυνατότητα εξόρυξης κειμένων μέσω RSS και επιλέχθηκαν για κάθε μία διαφορετικές κατηγορίες ενδιαφέροντος, που απεικονίζονται στον Πίνακα 5.1. Οι κατηγορίες που προτιμήθηκαν, καθώς το θέμα είναι κυρίως πολιτικού ενδιαφέροντος, ήταν σχετικές με την πολιτική σε όσες ιστοσελίδες παρείχαν την επιλογή αυτή, ενώ στις ιστοσελίδες που δεν παρείχαν αντίστοιχη κατηγορία επιλέχθηκε η κατηγορία της επικαιρότητας γενικά.

Ωστόσο, κατά την εισαγωγή των σχετικών συνδέσμων στο σύστημα ανάλυσης KNIME, με σκοπό την ανάκτηση δεδομένων, προέκυψε ότι δεν ήταν δυνατή η ορθή λειτουργία όλων των ιστοσελίδων. Σε κάποιες από αυτές, δεν ήταν συμβατό το RSS των ιστοσελίδων με τη μορφή ή τις εκδόσεις του RSS που υποστηρίζει το σύστημα KNIME, με αποτέλεσμα να μην είναι εφικτή η ανάκτηση στοιχείων των σελίδων και η περαιτέρω επεξεργασία, ενώ σε κάποιες άλλες ήταν αδύνατη η ανάκτηση και των πέντε βασικών στοιχείων των κειμένων, με αποτέλεσμα να μην μπορεί να ολοκληρωθεί πλήρως και σωστά η μετέπειτα ανάλυση.

Ως εκ τούτου, τελικά επιλέχθηκαν οι πρώτες οκτώ ιστοσελίδες που περιλαμβάνονται στον Πίνακα 5.1 και από τις οποίες ήταν επιτυχής η εξόρυξη δεδομένων.

Πίνακας 5.1

Επιλεγμένες ιστοσελίδες

A/A	ΙΣΤΟΣΕΛΙΔΑ	ΕΠΙΛΕΓΜΕΝΗ ΚΑΤΗΓΟΡΙΑ
1	Kathimerini.gr	Όλη η επικαιρότητα
2	Avgi.gr	-
3	Euro2day.gr	Επικαιρότητα
4	Skai.gr	Πολιτική
5	Naftemporiki.gr	Newsroom/ Πολιτική/ Οικονομία & Αγορές
6	Parapolitika.gr	-
7	Timesnews.gr	Ελλάδα
8	Topontiki.gr	-
9	Capital.gr	Όλες οι ειδήσεις
10	News.gr	Ροή ειδήσεων/ Πολιτική/ Σημαντικές ειδήσεις
11	Newsbeast.gr	Πολιτική
12	Tokarfi.gr	-
13	Tokerdos.gr	Διεθνή/ Πολιτική - Κοινωνία
14	Rizospastis.gr	Πολιτική
15	Insider.gr	-
16	Efsyn.gr	-
17	Protothema.gr	-

Επόμενο στάδιο της διαδικασίας κατά τη συλλογή δεδομένων ήταν η αφαίρεση των κειμένων που εμφανίζονται περισσότερες από μία φορές, καθώς και η σωστή μορφοποίηση του πεδίου της ημερομηνίας, με σκοπό η μορφή να είναι ενιαία για όλα κείμενα και να παρέχεται η δυνατότητα επεξεργασίας της από το σύστημα KNIME σε επόμενα στάδια της διαδικασίας ανάλυσης. Από τις οκτώ ιστοσελίδες, μόνο 2 είχαν μορφή που επέτρεπε άμεσα τη ομαλή συνέχιση της διαδικασίας της ανάλυσης. Για την ανάλυση στοιχείων από τις υπόλοιπες έξι,

χρειάστηκε επεξεργασία των δεδομένων μέχρι να απεικονιστούν στην επιθυμητή μορφή που μπορούσε να αναγνωρίσει το σύστημα του KNIME και να συνεχιστεί η ανάλυση.

Πίνακας 5.2

Αριθμός ανακτηθέντων κειμένων πριν και μετά την προεπεξεργασία

A/A	ΙΣΤΟΣΕΛΙΔΑ	ΑΡΙΘΜΟΣ ΚΕΙΜΕΝΩΝ ΑΡΧΙΚΑ	ΑΡΙΘΜΟΣ ΚΕΙΜΕΝΩΝ ΤΕΛΙΚΑ
1	Kathimerini.gr	259	35
2	Avgi.gr	160	10
3	Euro2day.gr	1100	8
4	Skai.gr	420	28
5	Naftemporiki.gr	155	6
6	Parapolitika.gr	427	6
7	Timesnews.gr	140	3
8	Topontiki.gr	200	6
ΣΥΝΟΛΟ		2861	102

Συνεχίζοντας τη διαδικασία συλλογής δεδομένων, το επόμενο βήμα ήταν η αποθήκευση των δεδομένων με τη μορφή πίνακα σε καθημερινή βάση για όλο το διάστημα της περιόδου αναφοράς. Μετά την ολοκλήρωση της διαδικασίας συλλογής, οι πίνακες που είχαν συγκεντρωθεί επεξεργάστηκαν μέσω του KNIME και αθροίστηκαν με χρήση κατάλληλου κόμβου του συστήματος KNIME σε ένα τελικό πίνακα δεδομένων για κάθε ιστοσελίδα. Τα αποτελέσματα των κειμένων που συγκεντρώθηκαν πριν το στάδιο της προεπεξεργασίας παρουσιάζονται στον Πίνακα 5.2.

Τελευταίο στάδιο της διαδικασίας συλλογής δεδομένων είναι ο περιορισμός των κειμένων ως προς τον όρο ενδιαφέροντος, δηλαδή τα κείμενα που αναφέρουν τον όρο ενδιαφέροντος στο κυρίως κείμενό τους. Η επιλογή του τελικού όρου που χρησιμοποιήθηκε, δηλαδή τη λέξη «Πρέσπες», δεν έγινε αυθαίρετα. Με σκοπό την τελική ορθότερη επιλογή, στο στάδιο αυτό δοκιμάστηκαν κι άλλοι όροι, όπως : Μακεδονικό, Σκοπιανό, Σκόπια, ονομασία, είτε μεμονωμένα οι όροι αυτοί είτε σε εκφράσεις, όπως «ζήτημα ονομασίας», «Μακεδονικό ζήτημα» κι εξετάστηκε δειγματοληπτικά κάθε φορά η συνάφεια των κειμένων που προέκυψαν ως προς το γεγονός της «Συμφωνίας των Πρεσπών». Το συμπέρασμα ήταν ότι οι άλλοι όροι σε πολλά κείμενα αποτελούσαν απλή αναφορά, χωρίς να αποτελεί μέρος της ανάλυσής τους η

«Συμφωνία των Πρεσπών». Αντίθετα, ο όρος «Πρέσπες» συνεισέφερε στην ανάκτηση περισσότερων κειμένων που είχαν ως θέμα αναφοράς και τη «Συμφωνία των Πρεσπών».

5.2.2 Δημιουργία εξειδικευμένου λεξικού

Κατά την προσπάθεια δημιουργίας εξειδικευμένου λεξικού συναισθήματος, όπως έχει ήδη αναφερθεί, είναι σημαντική η χρήση λέξεων που χρησιμοποιούνται σε κείμενα που περιλαμβάνονται στη βάση δεδομένων που δημιουργείται για την εκάστοτε ανάλυση. Για το λόγο αυτό, μετά τον περιορισμό των κειμένων που συλλέχθηκαν με τη χρήση του όρου ενδιαφέροντος, έγινε τυχαία επιλογή μερικών κειμένων, τα οποία και αναλύθηκαν στις λέξεις από τις οποίες αποτελούνται, μία προς μία, δημιουργώντας έτσι ένα «σάκο λέξεων».

Από το «σάκο λέξεων» διατηρήθηκαν αρχικά μόνο τα ρήματα, τα επίθετα, τα ουσιαστικά και τα επιρρήματα. Στη συνέχεια, ακολούθησε η προσπάθεια εντοπισμού του απόλυτου συναισθήματος των λέξεων που συγκεντρώθηκαν και κατά συνέπεια εξαιρέθηκαν, αφού εντοπίστηκαν και σε επίπεδο έκφρασης μέσα στο κείμενο, οι λέξεις που ήταν ουδέτερες σε απόλυτο συναίσθημα, π.χ. ρήματα όπως το ξέρω, μπορώ, και οι οποίες δεν συνέβαλαν στην διεξαγωγή της ανάλυσης, ενώ πιθανώς στην πορεία να διευκόλυναν την ανάπτυξη αστοχιών του συστήματος. Ακολούθησε η επεξεργασία της λίστας των λέξεων που διατηρήθηκαν με σκοπό των εντοπισμό, πέραν του απόλυτου, και του σχετικού συναισθήματος. Για να επιτευχθεί αυτό, οι λέξεις εντοπίστηκαν μέσα σε φράσεις των κειμένων στα οποία περιέχονταν και χαρακτηρίστηκαν τελικά με βάση το συναίσθημα που προέκυπτε κατά την ανάγνωση του κειμένου. Το τελικό αποτέλεσμα είναι η δημιουργία ενός εξειδικευμένου λεξικού που περιλαμβάνει συνολικά 296 λέξεις: 148 λέξεις αρνητικής πολικότητας και 148 λέξεις θετικής πολικότητας.

Στο σημείο αυτό, θεωρήθηκε ότι θα συνέβαλε θετικά στην αποτελεσματικότητα του λεξικού μία προσπάθεια συσχετισμού του λεξικού που δημιουργήθηκε με κάποιο υφιστάμενο ελληνικό λεξικό συναισθήματος. Η διαδικασία αυτή θεωρήθηκε ιδανική για τον έλεγχο συναισθήματος των λέξεων που επιλέχθηκαν, καθώς δεν υπήρχε η δυνατότητα εντοπισμού συναισθήματος των λέξεων μέσα σε φράσεις από δεύτερο αναγνώστη. Κατόπιν έρευνας στο διαδίκτυο, βρέθηκε διαθέσιμο ένα γενικευμένο ελληνικό λεξικό συναισθήματος στην πλατφόρμα GitHub, δημιουργημένο από την ομάδα των : Α. Τσακαλίδης, Σ. Παπαδόπουλος, Ουρ. Βοσκάκη, Κ. Ιωαννίδου και Χρ. Βοϊδίδου (Tsakalidis, Papadopoulos, & Kompatsiaris, 2014).

Ο αριθμός των λέξεων που περιλαμβάνονται στο συγκεκριμένο λεξικό ανέρχεται στις 2317 και παρέχει πληροφόρηση ως προς το μέρος του λόγου κάθε λέξης, την υποκειμενικότητα, την πολικότητα και το συναίσθημα που απορρέει από την κάθε λέξη. Τα αποτελέσματα προέρχονται από τις υποκειμενικές αξιολογήσεις τεσσάρων βαθμολογητών για κάθε κατηγορία.

Για να ολοκληρωθεί ο συσχετισμός του εξειδικευμένου λεξικού με το γενικευμένο, έγινε αναζήτηση όλων των λέξεων του πρώτου ώστε να εντοπιστούν, αν αυτό ήταν εφικτό, στο γενικευμένο λεξικό και να συγκριθεί η πολικότητά τους. Το αποτέλεσμα της σύγκρισης των λέξεων ήταν η παρατήρηση ότι πολλές από τις λέξεις που είχαν επιλεγεί για το εξειδικευμένο λεξικό να μην περιλαμβάνονται στο γενικευμένο. Συγκεκριμένα, από τις 148 λέξεις αρνητικής

πολικότητας του εξειδικευμένου λεξικού περιλαμβάνονται μόνο οι 35 στο γενικευμένο, ενώ από τις 148 λέξεις θετικής πολικότητας μόνο οι 30 περιλαμβάνονται στο γενικευμένο. Επιπλέον, είναι σημαντικό να αναφερθεί ότι κατά τη σύγκριση της πολικότητας των λέξεων που ήταν κοινές στο εξειδικευμένο και το γενικευμένο λεξικό, συνολικά 75, μόνο σε μία λέξη εντοπίστηκε διαφορετική πολικότητα στο γενικευμένο λεξικό, σε σχέση με την πολικότητα με την οποία χαρακτηρίστηκε κατά τη δημιουργία του εξειδικευμένου λεξικού. Η πολικότητα της συγκεκριμένης λέξης επανεξετάστηκε σε σχέση με το κείμενο στο οποίο περιλαμβάνεται και το συναίσθημα που απορρέει από τα συμφραζόμενα και διατηρήθηκε η πολικότητά της.

Μία ενδιαφέρουσα πρόκληση, που έπρεπε να αντιμετωπιστεί μετά τη δημιουργία του λεξικού, ήταν ότι δεν ήταν εφικτή η χρήση εφαρμογής διατήρησης μόνο της ρίζας κάθε λέξης, μία διαδικασία γνωστή στην ανάλυση ως stemming και για την οποία έχει γίνει ήδη αναφορά. Για την αντιμετώπιση του συγκεκριμένου προβλήματος και για να μην παρουσιάσει αστοχίες η ανάλυση κατά το στάδιο της τελικής επεξεργασίας, κρίθηκε σκόπιμο να εντοπιστούν δειγματοληπτικά από κείμενα οι συνηθέστερα χρησιμοποιούμενοι τύποι, χρόνοι και πρόσωπα ρημάτων, αριθμοί και πρόσωπα επιθέτων και ουσιαστικών, στοιχεία τα οποία συμπεριλήφθηκαν στο λεξικό.

Το λεξικό ελέγχθηκε, ως προς την ακρίβεια της πολικότητας που προσδίδει, με την εφαρμογή ανάλυσης συναισθήματος σε τυχαίο δείγμα κειμένων και τη διασταύρωση της πολικότητας του δείγματος σε σχέση με το συναίσθημα που προκύπτει κατά την ανάγνωση των κειμένων, ενώ χρειάστηκαν και μικρές παρεμβάσεις βελτίωσης οι οποίες καθόρισαν τον τελικό αριθμό που αναφέρθηκε πιο πάνω.

5.3 Επεξεργασία δεδομένων και αποτελέσματα

Έχοντας ολοκληρώσει τα στάδια της συλλογής και της δημιουργίας του λεξικού συναισθήματος, ακολουθεί το στάδιο της ανάλυσης συναισθήματος, το οποίο περιλαμβάνει μια σειρά διαδικασιών που πρέπει να ακολουθηθούν για να προκύψουν ορθά και μετρήσιμα αποτελέσματα.

Βασικό στάδιο της ανάλυσης ήταν η προεπεξεργασία των δεδομένων στα κείμενα που συγκεντρώθηκαν μετά το φιλτράρισμα της αρχικής βάσης δεδομένων με τον όρο ενδιαφέροντος που είχε επιλεγεί. Αφαιρέθηκαν από τα κείμενα όλοι οι αριθμοί και τα σημεία στίξης, αφαιρέθηκαν όλες οι λέξεις που είχαν έως και τέσσερα ψηφία και τα κεφαλαία γράμματα μετατράπηκαν σε πεζά. Μετά την προεπεξεργασία, το αποτέλεσμα ήταν ένα σύνολο κειμένων από το οποίο δημιουργήθηκε ο τελικός «σάκος λέξεων» που αποτέλεσε τη βάση στην οποία επρόκειτο να πραγματοποιηθεί η ανάλυση συναισθήματος.

Στη συνέχεια, οι λέξεις που περιλαμβάνονταν τόσο στο «σάκο λέξεων» όσο και στο εξειδικευμένο λεξικό, ορίστηκαν μέσω συγκεκριμένης διαδικασίας του συστήματος ως «αρνητικές» (NEG) ή «θετικές» (POS) και χαρακτηρίστηκαν αντίστοιχα από την πολικότητα που τους είχε αποδοθεί. Οι λέξεις του «σάκου» που δεν περιλαμβάνονταν στο λεξικό δεν λήφθηκαν υπόψη κατά την υπόλοιπη διαδικασία της ανάλυσης. Ακολούθησε η διαδικασία προσδιορισμού του αθροίσματος της πολικότητας των λέξεων κάθε κειμένου. Η διαδικασία αυτή καθορίστηκε με βάση τον τίτλο του εκάστοτε κειμένου, ώστε το άθροισμα τελικά να

αφορά ολόκληρο το κείμενο. Οι πολικότητες που προέκυψαν τελικά ήταν (9) η μέγιστη και (-5) η ελάχιστη, σε απολύτους αριθμούς του αθροίσματος για κάθε κείμενο.

Επόμενο βήμα της διαδικασίας είναι η κατηγοριοποίηση και ο τελικός χαρακτηρισμός των κειμένων με βάση το άθροισμα πολικότητας του καθενός, χρησιμοποιώντας τη δυνατότητα αυτόματης κατηγοριοποίησης που παρέχει το σύστημα KNIME. Η επιθυμητή επιλογή ήταν η κατηγοριοποίηση των κειμένων σε 3 επίπεδα, με σκοπό το διαχωρισμό της πολικότητας των κειμένων σε αρνητική, θετική και σχεδόν ουδέτερη. Η κατηγοριοποίηση γίνεται αυτόματα και χωρίζει τα δεδομένα σε ίσα διαστήματα, που αναφέρονται ως bins (κάδοι). Το επίπεδο 1 (bin 1), εκφράζει την αρνητική πολικότητα και αφορά το διάστημα (-5) έως (-1), το επίπεδο 2 (bin 2), εκφράζει την ουδέτερη πολικότητα και αφορά το διάστημα 0 έως 4 και το επίπεδο 3 (bin 3), εκφράζει τη θετική πολικότητα και αφορά το διάστημα 5 έως 9. Η τελική κατανομή των πλέον κατηγοριοποιημένων κειμένων είναι η εξής: 5 κείμενα στο επίπεδο 1 (bin 1), 76 κείμενα στο επίπεδο 2 (bin 2) και 21 κείμενα στο επίπεδο 3 (bin 3), όπως δείχνει και ο Πίνακας 5.3.

Πίνακας 5.3

Αποτελέσματα κατηγοριοποίησης

A/A	ΙΣΤΟΣΕΛΙΔΑ	BIN 1	BIN 2	BIN 3
1	Euro2day.gr	1	5	2
2	Avgi.gr	-	6	4
3	Kathimerini.gr	1	24	10
4	Naftemporiki.gr	-	6	-
5	Parapolitika.gr	-	6	-
6	Skai.gr	2	21	5
7	Timesnews.gr	-	3	-
8	Topontiki.gr	1	5	-
ΣΥΝΟΛΟ		5	76	21

Ανάγοντας τους αριθμούς σε ποσοστά, προκύπτουν τα εξής: το 74,5% των κειμένων παρουσιάζει ουδέτερη πολικότητα, το 20,5% των κειμένων παρουσιάζει θετική πολικότητα και το 5% παρουσιάζει αρνητική πολικότητα. Το συμπέρασμα λοιπόν που προκύπτει είναι η ουδετερότητα της αρθρογραφίας σύμφωνα με την παρούσα ανάλυση, κατά σημαντικό ποσοστό, για τη συγκεκριμένη περίοδο ως προς το θέμα της «Συμφωνίας των Πρεσπών» με

ακόλουθη τη θετική τάση προς το θέμα, ενώ η κατηγορία κειμένων με αρνητική πολικότητα παρουσιάζει πολύ μικρό ποσοστό.

Σε μία δεύτερη προσπάθεια ανάγνωσης των δεδομένων, παρατηρούνται οι τάσεις ανά ειδησεογραφική ιστοσελίδα. Έτσι, παρατηρείται ότι από τρεις μόνο ιστοσελίδες, δηλαδή ποσοστό 37,5% των πηγών: Naftemporiki.gr, Parapolitika.gr και Timesnews.gr, συλλέχθηκαν μόνο κείμενα που χαρακτηρίστηκαν από ουδέτερη πολικότητα. Από μία ιστοσελίδα, δηλαδή ποσοστό 12,5% των πηγών: Avgi.gr, συλλέχθηκαν κείμενα που χαρακτηρίστηκαν από ουδέτερη και θετική πολικότητα, με ποσοστά 60% και 40% αντίστοιχα και δεν προέκυψαν καθόλου κείμενα αρνητικής πολικότητας. Αντίστοιχα, από μία ιστοσελίδα, δηλαδή ποσοστό 12,5% των πηγών: Topotiki.gr, συλλέχθηκαν κείμενα που χαρακτηρίστηκαν από ουδέτερη και αρνητική πολικότητα σε ποσοστά 83,3% και 16,7% αντίστοιχα, χωρίς να προκύπτουν καθόλου κείμενα θετικής πολικότητας.

Τέλος, από τρεις ιστοσελίδες, δηλαδή 37,5% των πηγών: Euro2day.gr, Kathimerini.gr και Skai.gr, συλλέχθηκαν κείμενα που χαρακτηρίζονταν από ουδέτερη, θετική και αρνητική πολικότητα. Στην πρώτη τα ποσοστά ουδέτερης, θετικής και αρνητικής πολικότητας ανέρχονται σε 62,5%, 25% και 12,5% αντίστοιχα. Στη δεύτερη τα αντίστοιχα ποσοστά κειμένων για κάθε κατηγορία πολικότητας ανέρχονται σε 68,6%, 28,6% και 2,8%, ενώ στην περίπτωση της τρίτης τα ποσοστά κειμένων για κάθε κατηγορία πολικότητας ανέρχονται σε 75%, 17,6% και 7,4%.

6 Συμπεράσματα

Συνοψίζοντας κι έχοντας ολοκληρώσει τη συγκεκριμένη ανάλυση συναισθήματος που έγινε με τη χρήση δεδομένων από ειδησεογραφικές ιστοσελίδες για το θέμα της «Συμφωνίας των Πρεσπών» κατά το διάστημα μεταξύ της 27ης Σεπτεμβρίου και της 31^{ης} Οκτωβρίου 2018, με τη χρήση εξειδικευμένου λεξικού που δημιουργήθηκε επί τούτου, προκύπτουν μερικά χρήσιμα συμπεράσματα.

Κατά πρώτον, προκύπτει τελικά ότι η γενική τάση για το εν λόγω θέμα στην αρθρογραφία διαδικτύου ήταν ουδέτερη με σημαντική διαφορά σε σχέση με τις άλλες δύο τάσεις.

Κατά δεύτερον, η ακόλουθη τάση, σε ποσοστό διόλου αμελητέο, είναι θετική υπέρ του συγκεκριμένου θέματος, ενώ η αρνητική τάση εμφανίζει σημαντικά μικρό ποσοστό στο σύνολο των κειμένων που συλλέχθηκαν.

Κατά τρίτον, η παρατήρηση των δεδομένων αναλυτικά σε επίπεδο πηγών επιβεβαίωσε τη βασική τάση ουδέτερης πολικότητας, δεδομένου ότι σε όλες τις ιστοσελίδες τα κείμενα που συλλέχθηκαν και είχαν ουδέτερη πολικότητα εμφάνιζαν τα μεγαλύτερα ποσοστά.

Κατά τέταρτον, κείμενα μόνο από δύο ιστοσελίδες από τις οκτώ παρουσίασαν τάσεις προς μία κατεύθυνση πολικότητας πέραν της ουδέτερης. Στη μία περίπτωση, η τάση που ακολουθούσε την ουδέτερη ήταν η αρνητική, ενώ στην άλλη περίπτωση ή τάση που ακολουθούσε την ουδέτερη ήταν η θετική.

Κατά πέμπτον, ως προς τη διαδικασία ανάλυσης συναισθήματος με τη χρήση λεξικού, ένα ουσιαστικό συμπέρασμα που προκύπτει είναι ότι η παρέμβαση του αναλυτή σε όλα τα στάδια της διαδικασίας κρίνεται σημαντική και απαραίτητη για τη μεγιστοποίηση της αποδοτικότητας της συνολικής διαδικασίας.

Επιπλέον, για την επιτυχή διεξαγωγή της ανάλυσης ήταν απαραίτητο να αντιμετωπιστούν πολλές ιδιαιτερότητες και προκλήσεις. Οι προκλήσεις αυτές σχετίζονται τόσο με τις πηγές άντλησης των δεδομένων όσο και με τη διαδικασία ανάπτυξης του μοντέλου με βάση το οποίο πραγματοποιήθηκε η ανάλυση. Η διαδικασία δημιουργίας του εξειδικευμένου λεξικού αποτελεί μεμονωμένα μία πρόκληση που απαιτεί σημαντικό χρόνο και ουσιαστική προσπάθεια μέχρι την ολοκλήρωσή της.

Ωστόσο, η παρούσα έρευνα αποτελεί μία αρχική βάση και θα μπορούσε να επεκταθεί περαιτέρω σε πολλαπλά επίπεδα. Αμέσως επόμενο επίπεδο ανάλυσης θα μπορούσε να είναι το χρονικό διάστημα συλλογής των δεδομένων και η επέκταση της δημιουργηθείσας βάσης δεδομένων σε αριθμό κειμένων, με σκοπό να διερευνηθεί αν από τα ευρήματα με χρήση

μεγαλύτερης βάσης δεδομένων θα προκύψουν διαφορετικά τελικά συμπεράσματα. Σε επίπεδο θέματος, η έρευνα προσφέρεται για σύγκριση των αποτελεσμάτων με την αντίστοιχη επικαιρότητα της γειτονικής χώρας, με σκοπό την εξέταση των τάσεων της επικαιρότητας μεταξύ των δύο χωρών παράλληλα. Ενδιαφέροντα αποτελέσματα πιθανώς να προέκυπταν κατά την επέκταση της ανάλυσης σε επίπεδο πηγών, αναλύοντας την επικαιρότητα για το συγκεκριμένο θέμα σε κείμενα που θα προκύψουν από μέσα κοινωνικής δικτύωσης.

Μεγάλο εύρος επιλογών ανάλυσης και περαιτέρω επέκτασης προσφέρει επίσης το λεξικό συναισθήματος που δημιουργήθηκε, στο οποίο μπορούν τόσο να προστεθούν διαφορετικοί βαθμοί πολικότητας στις λέξεις που περιλαμβάνει όσο και να προστεθούν συνδυασμοί φράσεων ευρέως χρησιμοποιούμενων σε διπλωματικά και πολιτικά θέματα, ενώ ιδιαίτερη πρόκληση αποτελεί και η εφαρμογή του εξειδικευμένου λεξικού που δημιουργήθηκε σε άλλα θέματα πολιτικής και διεθνών σχέσεων, ώστε να εξεταστεί εις βάθος η αποτελεσματικότητά του.

Κατάλογος πηγών

Ξενόγλωσσες πηγές

- Agathangelou, P., Katakis, I., Rori, L., Gunopulos, D., & Richards, B. (2017). Understanding online political networks: The case of the far-right and far-left in Greece. In G. Ciampaglia, A. Mashhadi, & T. Yasseri, *Social informatics - Part I* (pp. 162-177). Springer.
- Benedetto, F., & Tedeschi, A. (2016). Big data sentiment analysis for brand monitoring in social media streams by cloud computing. In W. Pedrycz, & S.-M. Chen, *Sentiment analysis and ontology engineering* (pp. 341-377). Springer.
- Biagioni, R. (2016). Sentiment analysis. In R. Biagioni, *The SenticNet sentiment lexicon: exploring semantic richness in multi-word concepts* (pp. 7-16). Springer.
- Bogdanovski, A. (2015). Macedonia - back in the global spotlight. *Norwegian Institute of International Affairs* (23).
- Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). Affective computing and sentiment analysis. In E. Cambria, D. Das, S. Bandyopadhyay, & A. Feraco, *A practical guide to sentiment analysis* (pp. 1-8). Springer.
- Ceka, B. (2018, 4). Macedonia: A new beginning? *Journal of Democracy* , 29 (2), pp. 142-157.
- Chen, H. (2012). Sentiment analysis. In H. Chen, *Darkweb: exploring and data mining the dark side of the web* (pp. 171-201). Springer.
- Dabrowski, M., & Myachenkova, Y. (2018, 2). The Western Balkans on the road to the European Union. *Policy Contribution* (04), pp. 1-23.
- Fernandez-Gavilanez, M., Juncal-Martinez, J., Garcia-Mendez, S., Costa-Montenegro, E., & Gonzalez-Castano, F. J. (2018, 2 26). Creating emoji lexica from unsupervised sentiment analysis of their descriptions. *Expert systems with applications* , 103, pp. 74-91.
- Fidanovski, K. (2018). What's in a name? Possible ways forward in the Macedonian name dispute. *School of Slavonic and Eastern European Studies* , 31 (1), pp. 18-44.
- Guidi, M., Ruiz-Agundez, I., & Canga-Sanchez, I. (2012). Knowledge mining from the twitter social network: the case of Barack Obama. In A. Abraham, *Computational social networks: mining and visualization* (pp. 211-229). Springer.
- Illik, G., & Gjurovski, M. (2014). Modernism as an obstacle: the postmodern nature of the European Union and the Republic of Macedonia. In C. T. Mojanoski, *Macedonia and the Balkans, a hundred years after the World War I - Security and Euro-Atlantic integrations* (pp. 373-384). Bitola Skopje: University "St. Kliment Ogridski".
- Ivanovski, H. (2013). The Macedonia-Greece dispute/ difference over the name issue: mitigating the inherently unsolvable. *New Balkan Politics* (14), pp. 48-80.

- Janev, K., Dimeski, S., & Ristov, I. (2014). Republic of Macedonia - Regional zone of conflict of interests and geopolitics. In A. Pajaziti, J. Jakimovski, H. Jashari, & J. Abdullai, *The Balkans in the new millenium: From balkanization to eutopia* (pp. 426-436). Skopje: Balkan Sociological Forum - Universitas Europae Orientalis Meridionalis.
- Kranjc, J., Smailovic, J., Podpecan, V., Grcar, M., Znidarsic, M., & Lavrac, N. (2015). Active learning for sentiment analysis on data streams: methodology and workflow implementation in the cloudflows platform. *Information Processing and Management* , 51, pp. 187-203.
- Liu, B. (2017). Many facets of sentiment analysis. In E. Cambria, D. Das, S. Bandyopadhyay, & A. Feraco, *A practical guide to sentiment analysis* (pp. 11-38). Springer.
- Luo, T., Chen, S., Xu, G., & Zhou, Z. (2013). Sentiment analysis. In T. Luo, S. Chen, G. Xu, & J. Zhou, *Trust-based collective view prediction* (pp. 53-68). Springer.
- Mavromatidis, F. (2010, 3 22). The role of the European Union in the name dispute between Greece and FYR Macedonia. *Journal of Contemporary European Studies* , 18 (1), pp. 46-62.
- McKenna, B., Myers, M. D., & Newman, M. (2017, 3 17). Social media in qualitative research: challenges and recommendations. *Information and Organization* , 27, pp. 87-99.
- Mohammad, S. M. (2017). Challenges in sentiment analysis. In E. Cambria, D. Das, S. Bandyopadhyay, & A. Feraco, *A practical guide to sentiment analysis* (pp. 61-77). Springer.
- Nikodinovska-Stefanovska, S. (2014). EU and regional cooperation in Western Balkans. In C. T. Mojanoski, *Macedonia and the Balkans, a hundred years after the World War I - Security and Euro-Atlantic Integrations - Vol. II* (pp. 339-353). Bitola Skopje: University "St. Kliment Ohridski".
- Ryan, J. (2013). *Sentiment analysis and text visualization within Irish politics*. Institute of Technology Blanchardstown, Dublin, Ireland.
- Sjoeberg, E. (2011). *Battlefields of memory: The Macedonian conflict and Greek historical culture*. Umea University.
- Tanev, I. (1999, 1). Legal aspects of the use of a provisional name for Macedonia in the United Nations system. *The American Journal of International Law* , 93 (1), pp. 155-160.
- Taylor, E. M., Rodriguez, C. O., Velasquez, J. D., Ghosh, G., & Banerjee, S. (2013). Web opinion mining and sentimental analysis. In J. D. Velasquez, V. Palade, & L. Jain, *Advanced techniques ing web intelligence - 2* (pp. 105-125). Springer.
- Tsakalidis, A., Papadopoulou, S., & Kompatsiaris, I. (2014). An ensemble model for cross-domain polarity classification on twitter. *International Conference on web information systems engineering* (pp. 168-177). Springer.
- Tziampiris, A. (2005). The name dispute in the Former Yugoslav Republic of Macedonia after the signing of the Interim Accord. In E. Kofos, *Athens-Skopje: An uneasy symbiosis (1995-2002)* (pp. 225-252). Athens: ELIAMEP.

Wang, H., & Zhai, C. (2017). Generative models for sentiment analysis and opinion mining. In E. Cambria, D. Das, S. Bandyopandhyay, & A. Feraco, *A practical guide to sentiment analysis* (pp. 107-130). Springer.

Zadrozny, P., & Kodali, R. (2013). Sentiment analysis. In P. Zadrozny, & R. Kodali, *Big data analytics using Splunk* (pp. 255-282). Springer.

Διαδικτυακοί τόποι/ Ιστοσελίδες

<https://github.com/MKLab-ITI/greek-sentiment-lexicon>

<https://www.knime.com>

Παράρτημα Α

Πίνακας Α-1

Τελικός Πίνακας κειμένων κειμένων μετά την ανάλυση συναισθήματος

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
1	Όλγα Γεροβασίλη: Η Συμφωνία των Πρεσπών η καλύτερη δυνατή	4	Bin 2	AVGI.GR
2	Αλ. Φλαμπουράρης: Αυξημένο επίδομα θέρμανσης για τα αδύνατα οικονομικά στρώματα	2	Bin 2	AVGI.GR
3	Γ.Σταθάκης: Δε θα υπάρξουν αυξήσεις στα τιμολόγια της ΔΕΗ	6	Bin 3	AVGI.GR
4	Κ. Γαβρόγλου: Να είμαστε ευγνώμονες για όσους θυσιάστηκαν	0	Bin 2	AVGI.GR
5	Κ. Παπακώστα από Κοζάνη: Να υψώσουμε τείχη απέναντι σε κείνους που επιχειρούν να διχάσουν τους Έλληνες	3	Bin 2	AVGI.GR
6	Μπουγιάρ Οσμάνι: Ενισχύεται η στήριξη στη Συμφωνία των Πρεσπών	4	Bin 2	AVGI.GR
7	Πάνος Καμμένος: Μην περιμένετε να πέσει η κυβέρνηση	4	Bin 2	AVGI.GR
8	Ρ. Σβίγκου: Η αντιπολίτευση επένδυσε στη μείωση των συντάξεων αλλά τελικά	7	Bin 3	AVGI.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	θα εφαρμοστεί σειρά θετικών μέτρων			
9	Ρ.Σβίγκου: Υπάρχει πολιτική ευθύνη για τις υποθέσεις διαφθοράς	9	Bin 3	AVGI.GR
10	Χ. Καφαντάρη: Η αποδοχή του σχεδίου προϋπολογισμού από την Ε.Ε. σημαντικό βήμα για την επιστροφή στην κανονικότητα (Video)	8	Bin 3	AVGI.GR
11	Ασημακοπούλου: Το υπουργικό συμβούλιο έχει καταντήσει πολιτικό surrínivor	4	Bin 2	EURO2DAY.GR
12	Γκιόκας: Υλοποιούνται τα πιο επικίνδυνα σχέδια των ΗΠΑ-NATO στα Βαλκάνια	-2	Bin 1	EURO2DAY.GR
13	Δημοσιογράφοι σχολιάζουν στον realfm 97,8 την παραίτηση Κοτζιά	0	Bin 2	EURO2DAY.GR
14	Δημοσκόπηση: Δείτε τη διαφορά ΣΥΡΙΖΑ - ΝΔ	2	Bin 2	EURO2DAY.GR
15	Νοτοπούλου: Το έργο σύνδεσης της Δράμας με την Εγνατία Οδό εξελίσσεται εντός χρονοδιαγράμματος	6	Bin 3	EURO2DAY.GR
16	ΠΓΔΜ: Εισαγγελική έρευνα σε βάρος του Ιβανόφ για τη μη υπογραφή της Συμφωνίας των Πρεσπών	1	Bin 2	EURO2DAY.GR
17	ΠτΔ: Ν' αφήσουμε πίσω τις πολιτικές υπερβολικής και αδικαιολόγητης λιτότητας	5	Bin 3	EURO2DAY.GR
18	Πώς είδε ο διεθνής Τύπος την επικύρωση της Συμφωνίας των Πρεσπών από τη βουλή της ΠΓΔΜ	2	Bin 2	EURO2DAY.GR
19	Βορίδης: Για το δεξιό η συνεργασία με τους αριστερούς είναι αμάρτημα	1	Bin 2	PARAPOLITIKA.GR
20	Κακλαμάνης: Ο Τσίπρας και ο Καμμένος παίζουν	3	Bin 2	PARAPOLITIKA.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	θέατρο σκιών			
21	Καμμένος: Την μια μέρα «αγκαλιάζει» το VMRO την άλλη αποστασιοποιείται	2	Bin 2	PARAPOLITIKA.GR
22	Καρατζαφέρης: Το υβρίδιο των Πρεσπών δεν έχει καμία προοπτική	1	Bin 2	PARAPOLITIKA.GR
23	Πώς ερμηνεύει ο γερμανικός Τύπος το «Ναι» της Βουλής της ΠΓΔΜ	1	Bin 2	PARAPOLITIKA.GR
24	Σκόπια: Προδικαστική έρευνα κατά του προέδρου Ιβανόφ	1	Bin 2	PARAPOLITIKA.GR
25	Κουμουτσάκος: Οι κυβερνώντες μετέτρεψαν σε μικροκομματική διαφωνία ένα εθνικό θέμα	1	Bin 2	TIMESNEWS.GR
26	Πλησιάζει το πολιτικό τέλος Καμμένου με «σφραγίδα» από το Μαξίμου	0	Bin 2	TIMESNEWS.GR
27	Τζανακόπουλος για ΠΓΔΜ: Εγινε το πρώτο μεγάλο βήμα	1	Bin 2	TIMESNEWS.GR
28	Βούτσης κατά Καμμένου: Θεωρώ την παρέμβαση του, αποσταθεροποιητική για την κυβερνητική πολιτική	-3	Bin 1	TOPONTIKI.GR
29	Γ. Κουμουτσάκος: Η χώρα δεν χρειάζεται υπουργό Εξωτερικών μερικής απασχόλησης και ειδικού σκοπού	4	Bin 2	TOPONTIKI.GR
30	Ιταλική συγγνώμη από τον Πρόεδρο της Ιταλίας Σέρτζιο Ματαρέλα για τον πόλεμο του '40	1	Bin 2	TOPONTIKI.GR
31	Καραθανασόπουλος: Στο πόδι αντιμετωπίζουν την επέκταση της αιγιαλίτιδας ζώνης	1	Bin 2	TOPONTIKI.GR
32	Ξενογιαννακοπούλου: Ενδυνάμωση της Δημόσιας Διοίκησης με το αναγκαίο προσωπικό και	2	Bin 2	TOPONTIKI.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	τον ψηφιακό μετασχηματισμό της			
33	Το περιεχόμενο της επιστολής παραίτησης Κοτζιά - Αιχμές για Καμμένο	4	Bin 2	TOPONTIKI.GR
34	Pulse: Προβάδισμα 11 μονάδων για τη Ν.Δ.	1	Bin 2	NAFTEMPORIKI.GR
35	Βουλή ΠΓΔΜ: Την Παρασκευή και πάλι για το Σύνταγμα	1	Bin 2	NAFTEMPORIKI.GR
36	Κυρ. Μητσοτάκης: Εκβιαζόμενος Πρωθυπουργός ο Αλ. Τσίπρας	4	Bin 2	NAFTEMPORIKI.GR
37	Σκοπιανά ΜΜΕ για την παραίτηση: Τορπιλίζεται η συμφωνία	0	Bin 2	NAFTEMPORIKI.GR
38	Αλ. Φλαμπουράρης: Σχέδια για αυξημένο επίδομα θέρμανσης	2	Bin 2	NAFTEMPORIKI.GR
39	Φ. Κουβέλης: Θα επεκτείνουμε το μεταφορικό ισοδύναμο, «παλεύουμε» τον ΦΠΑ στα νησιά	2	Bin 2	NAFTEMPORIKI.GR
40	Reuters: Εξασφαλισμένη η στήριξη για τη ψήφιση της συμφωνίας των Πρεσπών	2	Bin 2	SKAI.GR
41	Δημοσκόπηση ΣΚΑΪ: Διψήφιο προβάδισμα ΝΔ, μουρμούρα ΣΥΡΙΖΑ για τους ΑΝΕΛ	2	Bin 2	SKAI.GR
42	Κοτζιάς: Η ΠΓΔΜ να καταλάβει πως δεν υπάρχει λύση χωρίς αποδοχή των Πρεσπών	4	Bin 2	SKAI.GR
43	Πώς επηρεάζει η παραίτηση Κοτζιά την Συμφωνία των Πρεσπών	1	Bin 2	SKAI.GR
44	Στέιτ Ντιπάρτμεντ: Να εφαρμοστεί η Συμφωνία των Πρεσπών	4	Bin 2	SKAI.GR
45	Αποστάσεις προσπαθεί να πάρει ο Καμμένος από το VMRO	-2	Bin 1	SKAI.GR
46	Βαρουφάκης: Ο Σόρος ζήτησε την αποπομπή	-1	Bin 1	SKAI.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	μου . Νεοφασιστοειδές ο Καμμένος			
47	Δημοσκόπηση: Αυτοδυναμία της ΝΔ με 160 βουλευτές - Προβάδισμα 11,2%	6	Bin 3	SKAI.GR
48	Θεοδωράκης: «Δεν συμμετέχουμε σε σενάρια στήριξης της κυβέρνησης»	5	Bin 3	SKAI.GR
49	Θεοδωράκης: Το Ποτάμι αυτόνομο στις επόμενες εκλογές	2	Bin 2	SKAI.GR
50	Ικανοποίηση της Αθήνας. Τα συγχαρητήρια Τσίπρα σε Ζάεφ	5	Bin 3	SKAI.GR
51	Καμμένος: Αποστάτες οι οκτώ βουλευτές που στήριξαν τον Ζάεφ	2	Bin 2	SKAI.GR
52	Κατρούγκαλος: Η επέκταση της αιγιαλίτιδας νόμιμο και αναφαίρετο δικαίωμά μας	5	Bin 3	SKAI.GR
53	Μεγάλο χάσμα στους ANEL: Καμμένος κατά Υπουργών και Υφυπουργών που διόρισε	0	Bin 2	SKAI.GR
54	Μετρούν τα κουκιά στον ΣΥΡΙΖΑ μετά τις νέες απειλές Καμμένου	0	Bin 2	SKAI.GR
55	Μητσοτάκης: Μόνος τρόπος να μην κυρωθούν οι Πρέσπες είναι οι εκλογές	3	Bin 2	SKAI.GR
56	ΝΔ: Δεν στηρίζουμε συμφωνία που εκχωρεί μακεδονική γλώσσα και εθνότητα	2	Bin 2	SKAI.GR
57	ΝΔ: Τσίπρας - Καμμένος κοροϊδεύουν τους πολίτες για να μείνουν στην εξουσία	1	Bin 2	SKAI.GR
58	Ορκίστηκε ο Τσίπρας, είμαι ο πιο ευτυχισμένος παραιτηθείς λέει ο Κοτζιάς	1	Bin 2	SKAI.GR
59	Πάνο... κάτω για τον Καμμένο οι ANEL –	0	Bin 2	SKAI.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	Σγουρίδης: Να φύγουν οι διαφωνούντες			
60	Παπακώστα: Ψηφίζω τις Πρέσπες μόνο με προϋποθέσεις	3	Bin 2	SKAI.GR
61	Σε εξέλιξη η Πολιτική Γραμματεία του ΣΥΡΙΖΑ υπό τη σκιά του Καμμένου	1	Bin 2	SKAI.GR
62	Σπυράκη: Η ΝΔ ζητά εκλογές πριν πάει στη Βουλή η συμφωνία των Πρεσπών	3	Bin 2	SKAI.GR
63	Τζανακόπουλος: Αν θέλουμε δημοσιεύουμε τα πρακτικά, αλλά είναι απόρρητα	3	Bin 2	SKAI.GR
64	Τζανακόπουλος: Η κυβέρνηση θα έχει πλειοψηφία και χωρίς τους ΑΝΕΛ.	2	Bin 2	SKAI.GR
65	Την κατάθεση της επιστολής παραίτησης Κοτζιά στη Βουλή ζητά η ΝΔ	1	Bin 2	SKAI.GR
66	Τσίπρας για Κοτζιά: Δεν δημοσιοποιώ αυστηρά προσωπικές επιστολές	8	Bin 3	SKAI.GR
67	Υπονοούμενα Καμμένου αλλά VMRO για χρηματισμούς στα Σκόπια	1	Bin 2	SKAI.GR
68	«Είπε και εγένετο»;	4	Bin 2	KATHIMERINI.GR
69	«Με πόνο καρδιάς»	1	Bin 2	KATHIMERINI.GR
70	Αμεση Ανάλυση: Τι σηματοδοτεί η παραίτηση του Ν. Κοτζιά	4	Bin 2	KATHIMERINI.GR
71	Αντιπρόεδρος κυβέρνησης Σκοπίων: Η Συμφωνία των Πρεσπών έχει ευρύτερη υποστήριξη και στην ΠΓΔΜ και στην Ελλάδα	4	Bin 2	KATHIMERINI.GR
72	Απάντηση Μίτσκοσκι σε ΗΠΑ: «Ναι» στην ένταξη σε ΝΑΤΟ και Ε.Ε. αλλά όχι ως χώρα ταπεινωμένη	9	Bin 3	KATHIMERINI.GR
73	Βατερλώ	3	Bin 2	KATHIMERINI.GR
74	Γ. Σταθάκης: τα	7	Bin 3	KATHIMERINI.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	Χριστούγεννα θα είναι έτοιμο να ξεκινήσει το έργο στο Ελληνικό			
75	Γιοχάνες Χαν: Είναι ζωτικής σημασίας να διατηρηθεί η πολιτική κληρονομιά του Ν. Κοτζιά	3	Bin 2	KATHIMERINI.GR
76	Δήλωση Αντ. Σαμαρά για τις εξελίξεις στην ΠΓΔΜ	0	Bin 2	KATHIMERINI.GR
77	Δημοσκόπηση: Δυσaréσκεια για τη συγκυβέρνηση, στις 10 μονάδες η διαφορά ΝΔ - ΣΥΡΙΖΑ	5	Bin 3	KATHIMERINI.GR
78	Εξήλθεν ο σπείρων	3	Bin 2	KATHIMERINI.GR
79	Επιβεβαίωσε ότι είναι όμηρος...	1	Bin 2	KATHIMERINI.GR
80	Επικρίσεις από την ελάσσονα αντιπολίτευση για το προσχέδιο του προϋπολογισμού 2019	3	Bin 2	KATHIMERINI.GR
81	Η νέμεσις του ΣΥΡΙΖΑ	4	Bin 2	KATHIMERINI.GR
82	Η πτώση του κ. Κοτζιά	3	Bin 2	KATHIMERINI.GR
83	Ιταλική συγγνώμη από Ματαρέλα για τον πόλεμο του '40	1	Bin 2	KATHIMERINI.GR
84	Κ. Γαβρόγλου: Να είμαστε ευγνώμονες για όσους θυσιάστηκαν	0	Bin 2	KATHIMERINI.GR
85	Κουμουτσάκος: Να δοθεί στη δημοσιότητα η επιστολή παραίτησης του Κοτζιά	0	Bin 2	KATHIMERINI.GR
86	Μήνυμα Πρ. Παυλόπουλου σε Αγκυρα: Δεν απεμπολούμε ούτε κεραία από τα δικαιώματά μας	5	Bin 3	KATHIMERINI.GR
87	Μητσοτάκης: Δρομολογούνται εξελίξεις που είναι εις βάρος των εθνικών συμφερόντων	2	Bin 2	KATHIMERINI.GR
88	Μητσοτάκης: Εκβιαζόμενος πρωθυπουργός ο Τσίπρας	6	Bin 3	KATHIMERINI.GR
89	Μόσχα κατά ΗΠΑ μετά το «Ναι» στην Αναθεώρηση	1	Bin 2	KATHIMERINI.GR
90	Ν.Δ.: Έρευνα για τις	3	Bin 2	KATHIMERINI.GR

A/A	ΤΙΤΛΟΣ ΚΕΙΜΕΝΟΥ	ΣΥΝΟΛΙΚΗ ΠΟΛΙΚΟΤΗΤΑ	ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΕ ΤΡΙΑ ΕΠΙΠΕΔΑ	ΠΗΓΗ
	αναφορές του κ. Ν. Κοτζιά			
91	Ο Αλέξης Τσίπρας αναλαμβάνει το υπουργείο Εξωτερικών	0	Bin 2	KATHIMERINI.GR
92	Οι ΗΠΑ απαντούν στη Μόσχα για την ψηφοφορία στην ΠΓΔΜ: Παράλογες και ψευδείς οι κατηγορίες	5	Bin 3	KATHIMERINI.GR
93	Οι ΗΠΑ διαψεύδουν τις αιτιάσεις της Μόσχας	4	Bin 2	KATHIMERINI.GR
94	Πάνος Καμμένος: Ημερολόγια	2	Bin 2	KATHIMERINI.GR
95	Παπαγγελόπουλος στη Βουλή: Πρώτη φορά ακούω ότι οι εισαγγελείς έχουν κύκλους	9	Bin 3	KATHIMERINI.GR
96	Στ. Θεοδωράκης: Το Ποτάμι δεν συμμετέχει σε σενάρια στήριξης της κυβέρνησης	6	Bin 3	KATHIMERINI.GR
97	Στο στόχαστρο Ν.Δ. ο Καμμένος	-5	Bin 1	KATHIMERINI.GR
98	Σχέδιο εθνικής συμφιλίωσης από Ζάεφ	6	Bin 3	KATHIMERINI.GR
99	Τζανακόπουλος για ΠΓΔΜ: Εγινε το πρώτο μεγάλο βήμα	2	Bin 2	KATHIMERINI.GR
100	Το απόλυτο σενάριο φρίκης επανέρχεται	4	Bin 2	KATHIMERINI.GR
101	Το ρευστό πλαίσιο της εξωτερικής πολιτικής	4	Bin 2	KATHIMERINI.GR
102	Τσίπρας: Δεν θα ανεχθώ καμία διγλωσσία και προσωπική στρατηγική στην εθνική γραμμή της χώρας	5	Bin 3	KATHIMERINI.GR

Παράρτημα Β

Πίνακας Β-1

Εξειδικευμένο λεξικό συναισθήματος

A/A	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ
1	αβέβαιος-η-ο	ΑΡΝ	αγαπώ	ΘΕΤ
2	αγανακτώ	ΑΡΝ	αγκαλιάζω	ΘΕΤ
3	αγανάκτηση	ΑΡΝ	αδελφικός-η-ο	ΘΕΤ
4	άγουρος-η-ο	ΑΡΝ	αδιάρρηκτος-η-ο	ΘΕΤ
5	άδικος-η-ο	ΑΡΝ	άθλος	ΘΕΤ
6	αδυναμία	ΑΡΝ	ακλόνητος-η-ο	ΘΕΤ
7	αδύνατος-η-ο	ΑΡΝ	ακούραστος-η-ο	ΘΕΤ
8	αιχμηρός-η-ο	ΑΡΝ	αμέριμνος-η-ο	ΘΕΤ
9	ακύρωση	ΑΡΝ	αναγνώριση	ΘΕΤ
10	αμαρτία	ΑΡΝ	ανάπτυξη	ΘΕΤ
11	αμφίβολος-η-ο	ΑΡΝ	ανέλπιστος-η-ο	ΘΕΤ
12	αμφιβολία	ΑΡΝ	ανθρώπινος-η-ο	ΘΕΤ
13	αναγκάζομαι	ΑΡΝ	ανταμοιβή	ΘΕΤ
14	ανακόπτω	ΑΡΝ	αξία	ΘΕΤ
15	ανακόπτομαι	ΑΡΝ	αξίζω	ΘΕΤ
16	αναμενόμενος-η-ο	ΑΡΝ	αξιολογώ	ΘΕΤ
17	αναρωτιέμαι	ΑΡΝ	απαλλάσσω	ΘΕΤ
18	ανατρέπω	ΑΡΝ	απαλλάσσομαι	ΘΕΤ
19	ανεπιθύμητος-η-ο	ΑΡΝ	απαράμιλλος-η-ο	ΘΕΤ
20	ανεργία	ΑΡΝ	αποδοχή	ΘΕΤ
21	άνευρος-η-ο	ΑΡΝ	αποκτώ	ΘΕΤ
22	ανησυχία	ΑΡΝ	αποφασιστικότητα	ΘΕΤ
23	αντίθετος-η-ο	ΑΡΝ	αρέσω	ΘΕΤ
24	αντιμετωπίζω	ΑΡΝ	αρχή	ΘΕΤ
25	αντίρρηση	ΑΡΝ	αύξηση	ΘΕΤ
26	αντίσταση	ΑΡΝ	αυτοδυναμία	ΘΕΤ
27	αντισυστημικός-η-ο	ΑΡΝ	δεξιότητα	ΘΕΤ
28	απαιτώ	ΑΡΝ	δεσμός	ΘΕΤ
29	απαιτείται	ΑΡΝ	δημιουργώ	ΘΕΤ
30	απειλούμαι	ΑΡΝ	δημιουργία	ΘΕΤ
31	απειλώ	ΑΡΝ	διατηρώ	ΘΕΤ
32	απελπισμένος-η-ο	ΑΡΝ	διερευνώ	ΘΕΤ
33	απέχω	ΑΡΝ	διευθέτηση	ΘΕΤ
34	αποκλείομαι	ΑΡΝ	διευθετώ	ΘΕΤ
35	απορρίπτω	ΑΡΝ	διευκρίνιση	ΘΕΤ

A/A	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ
36	αποστασία	ΑΡΝ	διεύρυνση	ΘΕΤ
37	αποσύρω	ΑΡΝ	δικαιοσύνη	ΘΕΤ
38	αποχή	ΑΡΝ	δίκιο	ΘΕΤ
39	αποχώρηση	ΑΡΝ	διπλωματικός-η-ο	ΘΕΤ
40	απρόβλεπτος-η-ο	ΑΡΝ	δρομολογούμαι	ΘΕΤ
41	απώτερος-η-ο	ΑΡΝ	δυναμικός-η-ο	ΘΕΤ
42	ασοιχειώτος-η-ο	ΑΡΝ	δυνατός-η-ο	ΘΕΤ
43	ασφυκτικός-η-ο	ΑΡΝ	εγγυώμαι	ΘΕΤ
44	αυτοκαταστρέφομαι	ΑΡΝ	εγγύηση	ΘΕΤ
45	αφορμή	ΑΡΝ	εγκρίνω	ΘΕΤ
46	βασανιστικός-η-ο	ΑΡΝ	έγκυρος-η-ο	ΘΕΤ
47	δειλός-η-ο	ΑΡΝ	ειρηνευτικός-η-ο	ΘΕΤ
48	δειλία	ΑΡΝ	εισακούω	ΘΕΤ
49	δεσμευτικός-η-ο	ΑΡΝ	εκτιμώμαι	ΘΕΤ
50	διακινδυνεύω	ΑΡΝ	εκτιμώ	ΘΕΤ
51	διαλύομαι	ΑΡΝ	ελπίδα	ΘΕΤ
52	διαλύω	ΑΡΝ	εμπιστοσύνη	ΘΕΤ
53	διαμελίζομαι	ΑΡΝ	ενδιαφέρον	ΘΕΤ
54	διάλογος	ΑΡΝ	ενθουσιασμός	ΘΕΤ
55	διασπασμένος-η-ο	ΑΡΝ	ενσρκώνω	ΘΕΤ
56	διατρέχω	ΑΡΝ	εξασφαλίζω	ΘΕΤ
57	διαφορά	ΑΡΝ	εξασφαλίζομαι	ΘΕΤ
58	διαφωνία	ΑΡΝ	εξελίσσομαι	ΘΕΤ
59	διάψευση	ΑΡΝ	επανεκκίνηση	ΘΕΤ
60	δικαιολογούμαι	ΑΡΝ	επάρκεια	ΘΕΤ
61	διπλωμάτης	ΑΡΝ	επενδύω	ΘΕΤ
62	δουλεία	ΑΡΝ	επιδιόρθωση	ΘΕΤ
63	δύσκολος-η-ο	ΑΡΝ	επιθυμώ	ΘΕΤ
64	δυσφορώ	ΑΡΝ	επικαιρος-η-ο	ΘΕΤ
65	δυσφορία	ΑΡΝ	επικρατώ	ΘΕΤ
66	εγλωτισμός	ΑΡΝ	επίλυση	ΘΕΤ
67	ένταση	ΑΡΝ	επιταχύνομαι	ΘΕΤ
68	έντονος-η-ο	ΑΡΝ	επιταχύνω	ΘΕΤ
69	εξαέρωση	ΑΡΝ	επίτευξη	ΘΕΤ
70	εξαρτώμαι	ΑΡΝ	ερμηνεύομαι	ΘΕΤ
71	επιβάλλω	ΑΡΝ	ευσεβής-ες	ΘΕΤ
72	επέρχομαι	ΑΡΝ	ευχάριστος-η-ο	ΘΕΤ
73	επίπτωση	ΑΡΝ	ηγεσία	ΘΕΤ
74	ερωτηματικό	ΑΡΝ	ήρωας	ΘΕΤ
75	ευθύνη	ΑΡΝ	ηρωικός-η-ο	ΘΕΤ
76	ευτελίζω	ΑΡΝ	θαυμάζω	ΘΕΤ
77	ζορίζομαι	ΑΡΝ	θαυμασμός	ΘΕΤ
78	θρίλερ	ΑΡΝ	θαυμαστός-η-ο	ΘΕΤ
79	καθυστερημένος-η-ο	ΑΡΝ	θέληση	ΘΕΤ

A/A	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ
80	καθυστερώ	ΑΡΝ	θερμός-η-ο	ΘΕΤ
81	καταγγελτικός-η-ο	ΑΡΝ	θεσμικός-η-ο	ΘΕΤ
82	καταπίνω	ΑΡΝ	θετικός-η-ο	ΘΕΤ
83	κατάρα	ΑΡΝ	θεωρούμαι	ΘΕΤ
84	κίνδυνος	ΑΡΝ	θρυλικός-η-ο	ΘΕΤ
85	κουράζομαι	ΑΡΝ	θυμάμαι	ΘΕΤ
86	κυριαρχία	ΑΡΝ	ιδανικό	ΘΕΤ
87	κύρωση	ΑΡΝ	ιδανικός-η-ο	ΘΕΤ
88	λεγόμενος-η-ο	ΑΡΝ	κάθαρση	ΘΕΤ
89	λιγότερος-η-ο	ΑΡΝ	καλοσύνη	ΘΕΤ
90	λιτότητα	ΑΡΝ	κατακτώ	ΘΕΤ
91	μαστιγώνομαι	ΑΡΝ	καταλαβαίνω	ΘΕΤ
92	μαστιγώνω	ΑΡΝ	λογική	ΘΕΤ
93	μισός-η-ο	ΑΡΝ	λογικός-η-ο	ΘΕΤ
94	μομφή	ΑΡΝ	λύση	ΘΕΤ
95	μοναξιά	ΑΡΝ	μαλακός-η-ο	ΘΕΤ
96	μοναχός-η-ο	ΑΡΝ	ομάδα	ΘΕΤ
97	μπάλωμα	ΑΡΝ	ορθολογικός-η-ο	ΘΕΤ
98	παγιδεύομαι	ΑΡΝ	ουσιαστικός-η-ο	ΘΕΤ
99	παραπομπή	ΑΡΝ	όψιμος-η-ο	ΘΕΤ
100	περιορίζω	ΑΡΝ	παραμένω	ΘΕΤ
101	περιορίζομαι	ΑΡΝ	πατρίδα	ΘΕΤ
102	πίεση	ΑΡΝ	πατριωτικός-η-ο	ΘΕΤ
103	πολυδιασπασμένος-η-ο	ΑΡΝ	πειθαρχία	ΘΕΤ
104	προβληματικός-η-ο	ΑΡΝ	πιστεύω	ΘΕΤ
105	πρόκληση	ΑΡΝ	πιστός-η-ο	ΘΕΤ
106	προσυμφωνούμαι	ΑΡΝ	πελειοψηφία	ΘΕΤ
107	προσυμφωνώ	ΑΡΝ	πνευματικός-η-ο	ΘΕΤ
108	προσφυγή	ΑΡΝ	προβλέπομαι	ΘΕΤ
109	πρόωρος-η-ο	ΑΡΝ	προβλέπω	ΘΕΤ
110	πτώση	ΑΡΝ	προέρχομαι	ΘΕΤ
111	πυροδοτούμαι	ΑΡΝ	προοπτική	ΘΕΤ
112	πυροδοτώ	ΑΡΝ	προσδοκία	ΘΕΤ
113	σέρνω	ΑΡΝ	προσεκτικός-η-ο	ΘΕΤ
114	σιχαίνομαι	ΑΡΝ	προσέρχομαι	ΘΕΤ
115	σπάω	ΑΡΝ	προσφέρω	ΘΕΤ
116	σπεύδω	ΑΡΝ	προσήλωση	ΘΕΤ
117	στασιμότητα	ΑΡΝ	πρόσφορος-η-ο	ΘΕΤ
118	στεγνός-η-ο	ΑΡΝ	προτιμώ	ΘΕΤ
119	στοιχίζομαι	ΑΡΝ	προτροπή	ΘΕΤ
120	στοιχισμένος-η-ο	ΑΡΝ	προωθώ	ΘΕΤ
121	σύμπτωμα	ΑΡΝ	προωτόγνωρος-η-ο	ΘΕΤ
122	συμπτωματικός-η-ο	ΑΡΝ	πρωτοπόρος-α-ο	ΘΕΤ

A/A	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ	ΟΡΟΣ	ΠΟΛΙΚΟΤΗΤΑ
123	συνειδητός-η-ο	ΑΡΝ	ρεαλιστικός-η-ο	ΘΕΤ
124	συνωμοσία	ΑΡΝ	σηματοδοτώ	ΘΕΤ
125	τείχος	ΑΡΝ	σταθερός-η-ο	ΘΕΤ
126	τελευταίος-α-ο	ΑΡΝ	σταθερότητα	ΘΕΤ
127	τυφλός-η-ο	ΑΡΝ	στηρίζω	ΘΕΤ
128	τυφλώνομαι	ΑΡΝ	συγκίνηση	ΘΕΤ
129	τυφλώνω	ΑΡΝ	συμβολή	ΘΕΤ
130	τυχαίος-α-ο	ΑΡΝ	συμβολισμός	ΘΕΤ
131	υποθέτω	ΑΡΝ	συμβουλευτικός-η-ο	ΘΕΤ
132	υποχρέωση	ΑΡΝ	συμμετοχή	ΘΕΤ
133	υποχωρώ	ΑΡΝ	συμφωνία	ΘΕΤ
134	υστερία	ΑΡΝ	συνδρομή	ΘΕΤ
135	φοβάμαι	ΑΡΝ	συνείδηση	ΘΕΤ
136	φόβος	ΑΡΝ	συνταγματικός-η-ο	ΘΕΤ
137	φορτίζω	ΑΡΝ	σώζομαι	ΘΕΤ
138	φορτίο	ΑΡΝ	τέλειος-α-ο	ΘΕΤ
139	φταίω	ΑΡΝ	υπερασπίζομαι	ΘΕΤ
140	φτώχεια	ΑΡΝ	υπερτερώνω	ΘΕΤ
141	φτωχός-η-ο	ΑΡΝ	υπερψηφίζω	ΘΕΤ
142	χαλάω	ΑΡΝ	υπερψηφίζομαι	ΘΕΤ
143	χαλαρότητα	ΑΡΝ	υψηλός-η-ο	ΘΕΤ
144	χάλι	ΑΡΝ	υψηλόφρων	ΘΕΤ
145	χαμηλός-η-ο	ΑΡΝ	φαντάζω	ΘΕΤ
146	χαμηλώνω	ΑΡΝ	φίλος	ΘΕΤ
147	χρέος	ΑΡΝ	φίλη	ΘΕΤ
148	ψωραλέος-α-ο	ΑΡΝ	ωραίος-α-ο	ΘΕΤ