



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

Ανάλυση Συναισθήματος με μεθόδους Βαθιάς Μάθησης

Διονύσιος Καραμούζας - 8827 (A.M.)

Μια διπλωματική εργασία που εκπονήθηκε στα πλαίσια του
Προπτυχιακού διπλώματος

στο
Πολυτεχνείο ΑΠΘ
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Επίβλεψη: Καθηγ. Ιωάννης Πήτας
Συνεπίβλεψη: Μεταδιδακ. Ιωάννης Μαδεμλής

1η Αυγούστου, 2022

Δικαιώματα Πνευματικής Ιδιοκτησίας

Εγώ, ο Διονύσιος Καραμούζας, δηλώνω ότι αυτή η διπλωματική εργασία με τίτλο, "Ανάλυση Συναισθήματος με μενόδους Βαθιάς Μάθησης" και η μελέτη που παρουσιάζεται σε αυτήν είναι δικά μου. Επιβεβαιώνω ότι:

- Η εργασία αυτή έγινε εξ ολοκλήρου κατά την υποψηφιότητα για πτυχίο έρευνας στο πανεπιστήμιο αυτό.
- 'Οπου οποιοδήποτε μέρος αυτής της εργασίας έχει προηγουμένως αναφερθεί σε οποιουδήποτε είδους δημοσιευμένη μελέτη, αυτό έχει δηλωθεί σαφώς.
- 'Οπου έχω παραθέσει από έργα άλλων, η πηγή δίνεται πάντα. Με εξαίρεση τέτοια αποσπάσματα, αυτή η εργασία είναι εξ ολοκλήρου δική μου δουλειά.
- Αναγνώρισα όλες τις κύριες πηγές βοήθειας.
- 'Οπου η εργασία βασίζεται σε δουλειά που έκανα εγώ από κοινού με άλλους, έχω καταστήσει σαφές τι ακριβώς έγινε από άλλους και τι έχω συνεισφέρει ο ίδιος.

Eυχαριστίες

Θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντά μου, Ιωάννη Πήτα, που με εμπιστεύτηκε και με καθοδήγησε σε όλο αυτό το έργο. Θα ήθελα επίσης να ευχαριστήσω τον συνεπιβλέποντά μου, τον Ιωάννη Μαδεμλή, που απαντούσε σε τακτική βάση στις ερωτήσεις και τα μηνύματά μου και με βοήθησε να ολοκληρώσω αυτό το έργο. Τέλος, θέλω να ευχαριστήσω την οικογένεια και τους φίλους μου που με στήριξαν σε δύσκολες στιγμές.

Περιεχόμενα

Δικαιώματα Πνευματικής Ιδιοκτησίας	i
Ευχαριστίες	ii
Κατάλογος Σχημάτων	viii
Κατάλογος Πινάκων	xii
1 Εισαγωγή	1
1.1 Ορισμός Προβλημάτων	1
1.1.1 Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet	2
1.1.2 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα	2
1.2 Σχετική Βιβλιογραφία	3
1.2.1 Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet	3
1.2.2 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα	4
1.2.2.1 Ανάλυση Συναισθήματος σε Μεταφορική Γλώσσα	4
1.2.2.2 Μεταφορά Νευρωνικής Γνώσης	5
1.3 Προτεινόμενες Μέθοδοι	5
1.3.1 Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet	6
1.3.2 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα	6
1.4 Δομή Περιεχομένου	6
2 Επεξεργασία Φυσικής Γλώσσας	8
2.1 Τί είναι η ΕΦΓ;	8
2.1.1 Σύντομοι Ορισμοί	8
2.1.2 Φυσική Γλώσσα	9
2.1.3 Πρόκληση	9
2.1.4 Από τη γλωσσολογία στην ΕΦΓ	9
2.1.4.1 Γλωσσολογία	9

2.1.4.2	Υπολογιστική Γλωσσολογία	10
2.1.4.3	ΥΓ vs ΕΦΓ	10
2.2	Ιστορική Αναδρομή	11
2.2.1	Ρίζες	11
2.2.2	Τρεις Περίοδοι	11
2.2.2.1	Συμβολική ΕΦΓ (κλασικός προγραμματισμός)	11
	Επιτεύγματα	11
2.2.2.2	Στατιστική ΕΦΓ – Μηχανική Μάθηση	12
	Μηχανική Μάθηση	12
	Εποπτευόμενη Μάθηση	13
	Μη-Εποπτευόμενη Μάθηση	13
	Γιατί Στατιστική ΕΦΓ;	14
	Επιτεύγματα	14
2.2.2.3	Νευρωνική ΕΦΓ	15
	Τεχνητό Νευρωνικό Δίκτυο (ΤΝΔ)	15
	Επιτεύγματα	15
2.3	Μέθοδοι: Κανόνες, Στατιστική, Νευρωνικά Δίκτυα	18
2.3.1	Κανόνες	18
2.3.2	Στατιστική	18
2.3.2.1	Γιατί Μηχανική Μάθηση;	19
2.3.2.2	Αλγόριθμοι Μηχανικής Μάθησης για ΕΦΓ	19
	Εποπτευόμενη Μάθηση	19
	Μη-Εποπτευόμενη Μάθηση	20
2.3.3	Νευρωνικά Δίκτυα	21
2.3.3.1	Γιατί Νευρωνικά Δίκτυα;	22
	Νευρωνικά Δίκτυα - Αυτόματη Μάθηση Χαρακτηριστικών . .	22
	Νευρωνικά Δίκτυα - Μοντέλα 'Άκρο-σε-'Άκρο	22
2.3.3.2	Είδη Νευρωνικών Δικτύων για ΕΦΓ	23
	Επίπεδο Ενσωμάτωσης	23
	Πολυεπίπεδο Perceptron (MLP)	24
	Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ)	25
	Επαναλαμβανόμενα Νευρωνικά Δίκτυα (ΕΝΔ)	26
	Δίκτυα Μακρυπρόθεσμης και Βραχυπρόθεσμης Μνήμης (ΔΜΒΜ)	27
	ΕΝΔ vs ΣΝΔ	28
	Υβριδικά Νευρωνικά Δίκτυα	29
	ΣΝΔ + ΔΜΒΜ	30
	ΔΜΒΜ + ΣΝΔ	30
	ΣΝΔ // ΔΜΒΜ	31
2.4	Αναπαραστάσεις Λέξεων	32
2.4.1	Σταθερές Αναπαραστάσεις	32
2.4.1.1	One-hot Κωδικοποίηση	33
2.4.1.2	Σύνολα Λέξεων (ΣΛ)	33
2.4.1.3	Συχνότητα 'Όρου - Ανάστροφη Συχνότητα Εγγράφου (ΣΟ-ΑΣΕ)	34
2.4.1.4	Σταθερές Αναπαραστάσεις - Πλεονεκτήματα & Μειονεκτήματα	35

2.4.2	Κατανεμημένες Αναπαραστάσεις (Ενσωματώσεις)	35
2.4.2.1	Κλασικές Ενσωματώσεις Λέξεων	36
	Word2Vec	36
	Διαδικασία εκπαίδευσης Skip-gram	37
	Καθολικά Διανύσματα (GloVe)	39
	FastText	39
	Πρόβλημα με τις Κλασικές Ενσωματώσεις Λέξεων	40
2.4.2.2	Πλαισιωμένες Ενσωματώσεις Λέξεων	40
	Διανύσματα Περιβάλλοντος (CoVe)	41
	ELMo	41
	Μετασχηματιστής	42
	Μετασχηματιστής vs ENΔ	42
	Οπτική περιγραφή Μετασχηματιστή	43
	Προεκπαιδευμένα γλωσσικά μοντέλα που βασίζονται στον Μετασχηματιστή	50
	Γενετικός Προεκπαιδευμένος Μετασχηματιστής OpenAI (GPT)	50
	Αναπαραστάσεις Αμφίδρομου Κωδικοποιητή από Μετασχηματιστή (BERT)	51
2.5	Προβλήματα ΕΦΓ	54
2.5.1	Επεξεργασία κειμένου και ομιλίας	54
2.5.2	Μορφολογική Ανάλυση	54
2.5.3	Συντακτική Ανάλυση	54
2.5.4	Σημασιολογία Λέξεων	55
2.5.5	Σχεσιακή Σημασιολογία	55
2.5.6	Συζήτηση	55
2.5.7	Εφαρμογές ΕΦΓ υψηλού επιπέδου	55
3	Ανάλυση Συναισθήματος	57
3.1	Σχετικά με την ΑΣ	57
3.1.1	Τί είναι η ΑΣ;	57
3.1.2	Εφαρμογές της ΑΣ	58
3.2	Ένα πρόβλημα ταξινόμησης κειμένου	58
3.2.1	Επίπεδα ταξινόμησης	58
3.2.2	Βασικοί Αλγόριθμοι	59
3.2.2.1	Μέθοδοι βάσει Κανόνων (λεξικού)	59
	TextBlob	59
	VADER	60
	SentiWordNet	60
3.2.2.2	Μέθοδοι Εποπτευόμενης Μάθησης	61
3.2.3	Προεπεξεργασία Κειμένου	61
3.2.3.1	Κοινά βήματα προεπεξεργασίας για ΑΣ	62
	Στημονοποίηση/Λημματοποίηση	62
	Συμβολοποίηση	63
	Διανυσματική Αναπαράσταση Κειμένου	64
3.2.4	Ανάλυση Συναισθήματος με Νευρωνικά Δίκτυα (ENΔ, ΣΝΔ, BERT)	64
3.2.4.1	Βασική Ιδέα	64

3.2.4.2	Επαναλαμβανόμενα Νευρωνικά Δίκτυα (ΕΝΔ)	66
	Απλό ΕΝΔ	66
	Πρόβλημα με τα απλά ΕΝΔ – Βραχυπρόθεσμη Μνήμη	67
	Δίκτυα Μακρυπρόθεσμης-Βραχυπρόθεσμης Μνήμης (ΔΜΒΜ)	67
	Παράδειγμα ταξινόμησης κειμένου με ΔΜΒΜ	69
	Στοιβαγμένο ΔΜΒΜ	70
	Αμφίδρομο ΔΜΒΜ	71
3.2.4.3	Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ)	72
	Συνελικτικοί Πυρήνες	72
	Συνέλιξη σε ακολουθίες λέξεων	72
	Αναγνώριση γενικών προτύπων	73
	1Δ Συνελίξεις	74
	Πολλαπλοί πυρήνες	75
	Max-pooling	75
	Παράδειγμα ταξινόμησης κειμένου με το ΣΝΔ	76
3.2.4.4	Υβριδικό ΣΝΔ+ΔΜΒΜ	77
3.2.4.5	Ταξινόμηση κειμένου με BERT	78
4	Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet	80
4.1	Περίληψη	80
4.2	Εισαγωγή	81
4.3	Σχετική Βιβλιογραφία	83
4.3.1	Περιγραφή Κοινής Γνώμης	83
4.3.1.1	Μη σημασιολογικές μέθοδοι	83
4.3.1.2	Σημασιολογικές μέθοδοι χωρίς Συνάθροιση	84
4.3.1.3	Σημασιολογικές μέθοδοι με Συνάθροιση	85
4.3.1.4	Διαστάσεις και Αλγόριθμοι Σημασιολογικής Ανάλυσης	86
4.3.2	Πρόβλεψη Χρονοσειρών	87
4.4	Προτεινόμενος Μηχανισμός	88
4.4.1	Βήμα 1: Επιλέγοντας το επιθυμητό σύνολο tweets	89
4.4.2	Βήμα 2: Εξαγωγή μεμονωμένου περιγραφέα ανά tweet	89
4.4.2.1	Δεδομένα Εκπαίδευσης	90
4.4.2.2	Προεπεξεργασία Κειμένου	90
4.4.2.3	Νευρωνικά Μοντέλα	91
4.4.2.4	Υπερπαράμετροι	92
4.4.3	Βήμα 3: Συνάθροιση	93
4.5	Αξιολόγηση	94
4.5.1	Δεδομένα	94
4.5.2	1η Ανάλυση: Πρόβλεψη Χρονοσειράς	95
4.5.2.1	Υλοποίηση	95
4.5.2.2	Μετρικές και Αποτελέσματα	97
4.5.3	2η Ανάλυση: Οπτικοποιήσεις και Ποιοτική Αξιολόγηση	100
4.6	Συζήτηση	110
4.7	Συμπεράσματα	111

5 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα	113
5.1 Περίληψη	113
5.2 Εισαγωγή	113
5.3 Σχετική Βιβλιογραφία	114
5.3.1 Ανάλυση συναισθήματος σε Μεταφορική Γλώσσα	115
5.3.2 Μεταφορά Νευρωνικής Γνώσης	115
5.4 Προτεινόμενη Μέθοδος	116
5.5 Ποσοτική Αξιολόγηση	119
5.5.1 Λεπτομέρειες Υλοποίησης	119
5.5.2 Πλαίσιο Αξιολόγησης	120
5.6 Συμπεράσματα	120
Βιβλιογραφία	122

Κατάλογος Σχημάτων

2.1	Τα βασικά επίπεδα της γλωσσικής δομής	10
2.2	Διάγραμμα Ροής - Η διαγραμματική αναπαράσταση ενός αλγορίθμου.	11
2.3	Μηχανική Μάθηση vs Προσέγγιση βάσει Κανόνων	13
2.4	Ιστορία Νευρωνικής ΕΦΓ.	18
2.5	Παράδειγμα χρήστης επιπέδου ενσωμάτωσης σε ένα νευρωνικό δίκτυο.	23
2.6	Παράδειγμα αρχιτεκτονικής MLP με 5 κρυφά επίπεδα.	25
2.7	Παράδειγμα ΣΝΔ σε μια εφαρμογή ταξινόμησης ιριτικών ταινιών.	26
2.8	Αναπαράσταση του “ξεδιπλωμένου” ΕΝΔ.	26
2.9	Η δομή μιας μονάδας ΔΜΒΜ.	27
2.10	Παράδειγμα ΔΜΒΜ για ταξινόμηση ιριτικών ταινιών.	28
2.11	Παράδειγμα αρχιτεκτονικής ΣΝΔ + ΔΜΒΜ.	30
2.12	Παράδειγμα αρχιτεκτονικής ΔΜΒΜ + ΣΝΔ.	31
2.13	Παράδειγμα αρχιτεκτονικής ΣΝΔ // ΔΜΒΜ.	32
2.14	Παράδειγμα One-hot κωδικοποίησης.	33
2.15	Παράδειγμα κωδικοποίησης με ΣΛ.	34
2.16	Παράδειγμα κωδικοποίησης ΣΟ-ΑΣΕ.	35
2.17	Skip-gram vs CBOW.	36
2.18	Η βασική ιδέα της προσέγγισης Skip-gram.	37
2.19	Παράδειγμα αρνητικής δειγματοληψίας.	37
2.20	Τα δεδομένα εκπαίδευσης του μοντέλου Skip-gram.	38
2.21	Οι πίνακες ενσωματώσεων του μοντέλου Skip-gram.	39
2.22	Η διαδικασία εκπαίδευσης του μοντέλου Skip-gram.	39
2.23	Οι στοίβες κωδικοποιητών και αποκωδικοποιητών του Μετασχηματιστή.	43
2.24	Η εσωτερική δομή των μπλοκ Κωδικοποιητή-Αποκωδικοποιητή.	43
2.25	Ροή δεδομένων στον Κωδικοποιητή.	44
2.26	Διανύσματα Query, Key και Value.	45
2.27	Τα βήματα υπολογισμού της αυτο-προσοχής.	45
2.28	Τα βήματα υπολογισμού της αυτο-προσοχής σε μορφή πινάκων.	46
2.29	Υπολογισμός Προσοχής πολλαπλών Κεφαλών.	47
2.30	Πρόσθεση κωδικοποιήσεων θέσης για να ληφθεί υπόψη η σειρά εμφάνισης των λέξεων στην ακολουθία.	48
2.31	Αυτόματη μετάφραση από άκρο σε άκρο με τον Μετασχηματιστή.	49
2.32	Το γραμμικό και το softmax επίπεδο μετατρέπουν την έξοδο της στοίβας του αποκωδικοποιητή σε λέξη.	50
2.33	Γλωσσικό Μοντέλο Μάσκας.	52
2.34	Πρόβλεψη επόμενης πρότασης.	53
2.35	Η στοίβα κωδικοποιητών του BERT παράγει πλαισιωμένες ενσωματώσεις λέξεων.	53

3.1	Αλγόριθμοι που χρησιμοποιούνται στην ΑΣ.	59
3.2	Στημονοποίηση vs Λημματοποίηση.	63
3.3	Παράδειγμα Συμβολοποίησης.	64
3.4	Τρεις ενσωματώσεις λέξεων.	64
3.5	Γενικό σχήμα Ανάλυσης Συναισθήματος (ταξινόμηση κειμένου).	65
3.6	Επεξεργασία ακολουθίας με ένα απλό ENΔ.	66
3.7	Υπολογισμοί μέσα στη μονάδα απλού ENΔ.	67
3.8	Πρόβλημα βραχυπρόθεσμης μνήμης των απλών ENΔ.	67
3.9	Η μονάδα ΔMBM.	68
3.10	Ανάλυση συναισθήματος κριτικών ταινιών με χρήση ΔMBM.	69
3.11	Υπολογισμοί που εκτελούνται από τη μονάδα ΔMBM σε κάθε χρονικό βήμα.	69
3.12	Παράδειγμα ανάλυσης συναισθήματος με Στοιβαγμένο ΔMBM.	70
3.13	Παράδειγμα ανάλυσης συναισθήματος με το AM-ΔMBM.	71
3.14	Ένας συνελικτικός πυρήνας.	72
3.15	Λειτουργία συνέλιξης που εφαρμόζεται σε κείμενο.	73
3.16	Παράδειγμα συνελικτικού πυρήνα που εντοπίζει θετικά χαρακτηριστικά.	74
3.17	Έξοδος 1Δ συνέλιξης.	75
3.18	Εφαρμογή max-pooling στην έξοδο της 1Δ συνέλιξης.	76
3.19	Παράδειγμα ανάλυσης συναισθήματος (ταξινόμηση κειμένου) με χρήση ΣΝΔ.	77
3.20	Παράδειγμα ανάλυσης συναισθήματος χρησιμοποιώντας αρχιτεκτονική ΣΝΔ+ΔMBM.	78
3.21	Παράδειγμα ταξινόμησης κειμένου χρησιμοποιώντας το προεκπαιδευμένο BERT.	78
4.1	Ποσοτική πρόβλεψη κοινής γνώμης με χρήση του προτεινόμενου μηχανισμού/σημασιολογικού περιγραφέα.	88
4.2	Καθημερινός αριθμός tweets για Δημοκρατικούς και Ρεπουμπλικάνους στο σύνολο δεδομένων του 2016. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	99
4.3	Καθημερινός αριθμός tweets για Δημοκρατικούς και Ρεπουμπλικάνους στο σύνολο δεδομένων του 2020. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	99
4.4	Οι καθημερινές τιμές ανά κανάλι των χρονοσειρών 4Δ που κατασκευάστηκαν από τα δεδομένα του 2016 χρησιμοποιώντας τον προτεινόμενο περιγραφέα και τη μέση στρατηγική συγκέντρωσης, ξεχωριστά για Δημοκρατικούς και Ρεπουμπλικάνους. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	100
4.5	Οι καθημερινές τιμές ανά κανάλι των χρονοσειρών 4Δ που κατασκευάστηκαν από τα δεδομένα του 2020 χρησιμοποιώντας τον προτεινόμενο περιγραφέα και τη μέση στρατηγική συγκέντρωσης, ξεχωριστά για Δημοκρατικούς και Ρεπουμπλικάνους. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	101

4.6 Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Δημοκρατικούς, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2016 (163 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	103
4.7 Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Ρεπουμπλικάνους, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2016 (163 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	104
4.8 Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Δημοκρατικούς, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2020 (25 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	104
4.9 Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Ρεπουμπλικάνους, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2020 (25 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).	105
4.10 Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Δημοκρατικούς στις 9 Νοεμβρίου 2016 (την επομένη των εκλογών).	106
4.11 Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Ρεπουμπλικάνους στις 9 Νοεμβρίου 2016 (την επομένη των εκλογών).	106
4.12 Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Δημοκρατικούς στις 4 Νοεμβρίου 2020 (την επομένη των εκλογών).	107
4.13 Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Ρεπουμπλικάνους στις 4 Νοεμβρίου 2020 (την επομένη των εκλογών).	107
4.14 Αριθμός tweets που αφορούν τους Δημοκρατικούς, ξεχωριστά για κάθε ικάση των τεσσάρων διαστάσεων περιγραφής, στις 9 Νοεμβρίου 2016. Τα δύο χρώματα διακρίνουν τις αντίθετες ικάσεις κάθε σημασιολογικής διάστασης.	109
4.15 Αριθμός tweets που αφορούν τους Ρεπουμπλικάνους, ξεχωριστά για κάθε ικάση των τεσσάρων διαστάσεων περιγραφής, στις 9 Νοεμβρίου 2016. Τα δύο χρώματα διακρίνουν τις αντίθετες ικάσεις κάθε σημασιολογικής διάστασης.	109
4.16 Αριθμός tweets που αφορούν τους Δημοκρατικούς, ξεχωριστά για κάθε ικάση των τεσσάρων διαστάσεων περιγραφής, στις 4 Νοεμβρίου 2020. Τα δύο χρώματα διακρίνουν τις αντίθετες ικάσεις κάθε σημασιολογικής διάστασης.	109

4.17 Αριθμός tweets που αφορούν τους Ρεπουμπλικάνους, ξεχωριστά για κάθε κλάση των τεσσάρων διαστάσεων περιγραφής, στις 4 Νοεμβρίου 2020. Τα δύο χρώματα διακρίνουν τις αντίθετες κλάσεις κάθε σημασιολογικής διάστασης.	110
5.1 Η προτεινόμενη αρχιτεκτονική εκπαίδευσης δασκάλου-μαθητή.	117

Κατάλογος Πινάκων

4.1	Επιτευχθείσα ακρίβεια καθενός από τους τέσσερις ταξινομητές γνώμης στο σύνολο δοκιμής του αντίστοιχου συνόλου δεδομένων εκπαίδευσης.	91
4.2	Υπερπαράμετροι που χρησιμοποιούνται για την εκπαίδευση των ταξινομητών συναισθημάτων.	92
4.3	Υπερπαράμετροι που χρησιμοποιούνται για την εκπαίδευση του μοντέλου πρόβλεψης.	96
4.4	Πρόβλεψη αποτελεσμάτων σχετικά με το σύνολο δεδομένων tweet για τις προεδρικές εκλογές των ΗΠΑ 2016 για τις έξι κατασκευασμένες χρονοσειρές. Το 'Dem' αναφέρεται στους Δημοκρατικούς, το 'Rep' στους Ρεπουμπλικάνους, ενώ τα 'mean', 'med', 'trim' στις αντίστοιχες 3 στρατηγικές συνάθροισης: μέσου όρου, διαμέσου και περικομμένου μέσου όρου. Σε κάθε περίπτωση, οι μετρικές SMAPE/MASE έχουν υπολογιστεί παίρνοντας μια ενιαία τιμή και για τα τέσσερα κανάλια περιγραφής χρησιμοποιώντας τόσο τον μέσο όρο όσο και τον διάμεσο τελεστή. Μια χαμηλότερη τιμή είναι καλύτερη και για τις δύο μετρικές, ενώ το SMAPE είναι ένα ποσοστό.	97
4.5	Πρόβλεψη αποτελεσμάτων σχετικά με το σύνολο δεδομένων tweet για τις προεδρικές εκλογές των ΗΠΑ 2020 για τις έξι κατασκευασμένες χρονοσειρές. Το 'Dem' αναφέρεται στους Δημοκρατικούς, το 'Rep' στους Ρεπουμπλικάνους, ενώ τα 'mean', 'med', 'trim' στις αντίστοιχες 3 στρατηγικές συνάθροισης: μέσου όρου, διαμέσου και περικομμένου μέσου όρου. Σε κάθε περίπτωση, οι μετρικές SMAPE/MASE έχουν υπολογιστεί παίρνοντας μια ενιαία τιμή και για τα τέσσερα κανάλια περιγραφής χρησιμοποιώντας τόσο τον μέσο όρο όσο και τον διάμεσο τελεστή. Μια χαμηλότερη τιμή είναι καλύτερη και για τις δύο μετρικές, ενώ το SMAPE είναι ένα ποσοστό.	97
5.1	Αποτελέσματα αξιολόγησης στο σύνολο δεδομένων S15-T11. Το υψηλότερο/χαμηλότερο είναι καλύτερο για τη μέτρηση COS/MTΣ, αντίστοιχα. Τα καλύτερα αποτελέσματα είναι με έντονους χαρακτήρες.	119

*Αφιερωμένη στη γιαγιά μου, Μαρία, που με φιλοξένησε και με στήριξε
κατά τη διάρκεια αυτού του πρότζεκτ.*

Κεφάλαιο 1

Εισαγωγή

Αυτό το κεφάλαιο έχει ως στόχο να δώσει μια εικόνα στον αναγνώστη για το αντικείμενο έρευνας αυτής της διπλωματικής εργασίας. Με σκοπό τη διεύρυνση του κοινού στο οποίο απευθύνεται το παρόν έργο, έχουν συμπεριληφθεί κάποια εισαγωγικά, θεωρητικά κεφάλαια επιπρόσθετα των πειραμάτων που πραγματοποιήθηκαν. Εκεί εξηγούνται βασικές έννοιες που βοηθούν στην βαθιά κατανόηση των πειραμάτων που παρουσιάζονται μετέπειτα αλλά και στην μύηση ενδιαφερόμενων ερευνητών στον ιλάδο της Επεξεργασίας Φυσικής Γλώσσας (ΕΦΓ). Σε αυτό το εισαγωγικό κεφάλαιο θα ορίσουμε συνοπτικά τα προβλήματα των οποίων η λύση αποτέλεσε το βασικό-πειραματικό κομμάτι αυτής της μελέτης. Θα δούμε ενδεικτικά την υπάρχουσα σχετική βιβλιογραφία για τα προβλήματα αυτά καθώς και τις προτεινόμενες από την παρούσα έρευνα μεθόδους επίλυσής τους. Τέλος, θα παρουσιάσουμε τη δομή του τί πρόκειται να ακολουθήσει στα επόμενα κεφάλαια.

1.1 Ορισμός Προβλημάτων

Στο πλαίσιο της κατηγορίας προβλημάτων ΕΦΓ που ασχολείται με την Ανάλυση Συναισθήματος (ΑΣ) κειμένου, ερευνήθηκαν δύο διαφορετικές εφαρμογές. Η πρώτη αφορά στην παρακολούθηση κοινής γνώμης μέσω συλλογικής σημασιολογικής ανάλυσης tweet. Η δεύτερη αφορά στην μεταφορά νευρωνικής γνώσης για βελτιωμένη ανάλυση συναισθημάτων σε κείμενα με μεταφορική γλώσσα. Περισσότερες λεπτομέρειες μπορούν να βρεθούν στις ενότητες 4.2, 5.2, αντίστοιχα.

1.1.1 Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet

Τα μέσα κοινωνικής δικτύωσης επιτρέπουν στους ανθρώπους να εκφράσουν άμεσα τις απόψεις και τα συναισθήματά τους χρησιμοποιώντας μια ποικιλία μέσων, όπως κείμενο, εικόνες, βίντεο κ.λπ. Κατά συνέπεια, αυτές οι πλατφόρμες μπορούν να χρησιμοποιηθούν για την παρακολούθηση της κοινής γνώμης που σχετίζεται με ένα θέμα ιδιαίτερου ενδιαφέροντος. Η κοινή γνώμη “αντιπροσωπεύει τις απόψεις και τις επιθυμίες της πλειοψηφίας ενός πληθυσμού σχετικά με ένα συγκεκριμένο θέμα, είτε πολιτικό, εμπορικό, κοινωνικό ή άλλο”[1]. Το twitter φαίνεται να είναι το μέσο επιλογής για τη διατύπωση απόψεων σχετικά με κοινωνικοπολιτικά ζητήματα όπως ο COVID-19, οι εκλογές, ο ρατσισμός κ.λπ. Ο τεράστιος αριθμός ανθρώπων που χρησιμοποιούν το twitter για να ενημερώνονται και να εκφράζουν τις απόψεις τους έχει προσφέρει στους πολιτικούς την ευκαιρία να μεταδώσουν το μήνυμά τους γρήγορα και φθηνά, χωρίς να περάσουν στις παραδοσιακές τεχνικές μέσων ενημέρωσης και συνεντεύξεις τύπου [2]. Δεδομένης αυτής της πολύ μεγάλης προοπτικής του twitter, η εξόρυξη γνώμης μέσω tweets μπορεί να μας δώσει πολύτιμες πληροφορίες. Τα αυτοματοποιημένα εργαλεία σημασιολογικής ανάλυσης κειμένου, που βασίζονται σε σύγχρονους αλγόριθμους ΕΦΓ που βασίζονται στην Τεχνητή Νοημοσύνη, μπορούν να προσδιορίσουν τη διάθεση ενός tweet (π.χ. πολικότητα/συναίσθημα [3]) με αξιοσημείωτη ακρίβεια. Τα βαθιά νευρωνικά δίκτυα (BNΔ) έχουν εξελίξει πολύ τις σύγχρονες μεθόδους αντιμετώπισης τέτοιων προβλημάτων, ειδικά στην ΑΣ. Αυτή η μελέτη επιχειρεί να διερευνήσει τις σχετικές ανεξερεύνητες δυνατότητες πολυδιάστατης συλλογικής ανάλυσης tweet μέσω υπερσύγχρονων BNΔ. Έτσι, προτείνεται ένας νέος, αυτοματοποιημένος μηχανισμός παρακολούθησης της κοινής γνώμης, που αποτελείται από έναν σύνθετο, ποσοτικό, σημασιολογικό περιγραφέα tweet που βασίζεται σε αλγόριθμους ΕΦΓ με BNΔ.

1.1.2 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναίσθημάτων σε Κείμενα με Μεταφορική Γλώσσα

Τα κείμενα που περιέχουν μεταφορική γλώσσα (π.χ. που περιλαμβάνουν σαρκασμό, ειρωνεία, μεταφορές κ.λπ.) εξακολουθούν να αποτελούν πρόκληση για τις υπάρχουσες μεθόδους ΑΣ [4]. Αυτό οφείλεται σε δυσκολίες στον προσδιορισμό του εάν μια μεταφορική φράση υπονοεί πράγματι αρνητική ή θετική γνώμη, δεδομένης της εγγενούς σημασιολογικής της ασάφειας. Το ζήτημα είναι ιδιαίτερα σημαντικό λόγω της πανταχού παρούσας φύσης της μεταφορικής γλώσσας στην ανθρώπινη επικοινωνία. Για παράδειγμα, ο σαρκασμός και η ειρωνεία είναι γνωστό ότι προσελκύουν μεγαλύτερη προσοχή στις διαδικτυακές κοινότητες [5], με πολλούς ανθρώπους να είναι έμπειροι στη χρήση τους και να τα εκμεταλλεύονται για να αυξήσουν τον αντίκτυπο των αναρτήσεών τους στα μέσα κοινωνικής δικτύωσης (π.χ. στο Twitter). Σε

αυτή τη μελέτη, προτείνεται μια νέα ρύθμιση νευρωνικής απόσταξης γνώσης [6] με σκοπό την αύξηση της ακρίβειας της ΑΣ σε κείμενα με μεταφορική γλώσσα [4].

1.2 Σχετική Βιβλιογραφία

Παρακάτω βλέπουμε κάποιες ενδεικτικές προσεγγίσεις προηγούμενων ερευνών πάνω στα δύο αυτά προβλήματα. Για μια πιο ενδελεχή παρουσίαση της υπάρχουσας βιβλιογραφίας των δύο αυτών προβλημάτων ανατρέξτε στις ενότητες 4.3, 5.3, αντίστοιχα.

1.2.1 Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet

Οι σημασιολογικές μέθοδοι που συνήθως εκτελούν έναν τύπο ανάλυσης συναισθήματος/εξόρυξης απόψεων με δυνατότητα τεχνητής νοημοσύνης σε tweet είναι οι πιο προηγμένες και παρέχουν πιο ακριβή αποτελέσματα. Η πλειοψηφία τέτοιων μεθόδων λειτουργούν στο twitter, αλλά δεν αντιμετωπίζουν τα σχετικά tweets συλλογικά και συνήθως λαμβάνουν υπόψη τον όγκο των tweets ανά κατηγορία συναισθήματος. Για παράδειγμα, το [1] παρουσίασε ένα πλαίσιο για την παρακολούθηση της εξέλιξης (16 μήνες) της κοινής γνώμης σχετικά με το θέμα της κλιματικής αλλαγής. Ο προτεινόμενος μηχανισμός είναι σε θέση να εντοπίζει και να παρακολουθεί on-the-fly το συναίσθημα σε ένα επιθυμητό σύνολο μεμονωμένων tweets, πιθανώς κατά τη δημοσίευσή τους, αλλά πραγματοποιείται μόνο στοιχειώδης ανάλυση σε αυτά τα αποτελέσματα ως συλλογή. Έτσι, η κοινή γνώμη στο σύνολό της ελάχιστα εξετάζεται. Έχοντας κατά νου έναν διαφορετικό στόχο, η [7] διεξήγαγε ΑΣ των tweets για την πρόβλεψη του εκλογικού αποτελέσματος. Χρησιμοποιήθηκε ένα απλό μοντέλο Συνελικτικού Νευρωνικού Δικτύου (ΣΝΔ) για αυτήν την εργασία, ενώ ο συνολικός όγκος των tweets (μη σημασιολογική ιδιότητα) συγκρίθηκε με την πολικότητα του συναισθήματος (όγκος θετικών tweets) για να διαπιστωθεί ποιος είναι καλύτερος προγνωστικός δείκτης των εκλογικών αποτελεσμάτων. Ο σημασιολογικός περιγραφέας βρέθηκε να είναι πιο ακριβής. Προχωρώντας σε αυτήν τη γραμμή έρευνας, το [8] είχε ως στόχο να προβλέψει όχι μόνο τον νικητή αλλά και το ποσοστό ψήφου κάθε υποψηφίου στις ισπανικές προεδρικές εκλογές του 2019, λαμβάνοντας υπόψη τον όγκο των θετικών tweets ανά υποψήφιο. Ωστόσο, τα ίδια τα σημασιολογικά χαρακτηριστικά δεν συναρμοίστηκαν στο σύνολο όλων των tweets.

Η συντριπτική πλειονότητα των σημασιολογιών μεθόδων που παρουσιάζονται παραπάνω χρησιμοποιεί μόνο το συναίσθημα του tweet, επομένως μπορούν να θεωρηθούν ότι εκμεταλλεύονται μονοδιάστατη σημασιολογία κειμένου. Αυτό είναι πιο προφανές σε περιπτώσεις όπου η σημασιολογική ανάλυση εξάγει μια πολικότητα (π.χ. δυαδική ή τριαδική ταξινόμηση σε θετικά/αρνητικά tweets ή σε θετικά/ουδέτερα/αρνητικά). Αυτή η περιορισμένη

προσέγγιση είναι η πιο κυρίαρχη [7] [8] [9] [10] [11]. Ωστόσο, η μονοδιάστατη ανάλυση συναισθημάτων μπορεί να θεωρηθεί ότι ισχύει ακόμη και όταν χρησιμοποιούνται πολυσυναισθηματικοί ταξινομητές αντί της απλής πολικότητας. Π.χ., στο [12] (προσδοκία, χαρά, αγάπη, έκπληξη, άγχος, λύπη, θυμός και μίσος), [13] (ευτυχία, αρέσκεια, λύπη, αηδία, έκπληξη, θυμός και φόβος) και [1] (χαρά, έμπνευση, θυμός, ρατσισμός, υποστήριξη). Αν και είναι μια πιο διαφοροποιημένη προσέγγιση, αυτά τα συναισθήματα εξακολουθούν να εμπίπτουν στη γενική ομπρέλα του συναισθήματος, επομένως αυτές οι μέθοδοι συνεχίζουν να αγνοούν άλλα χαρακτηριστικά σημασιολογικού κειμένου. Η μόνη περίπτωση όπου λαμβάνεται υπόψη η πολυδιάστατη σημασιολογία είναι η [14], όπου λαμβάνονται υπόψη 2 διαφορετικές διαστάσεις γνώμης (πολικότητα και προκατάληψη). Αντίθετα, αυτή η μελέτη προτείνει έναν 4-διάστατο περιγραφέα που εξετάζει από κοινού την πολικότητα, την μεροληψία, την μεταφορικότητα και την προσβλητικότητα, τα οποία είναι όλα διαφορετικά χαρακτηριστικά κειμένου και επαληθεύει πειραματικά τη χρησιμότητά τους.

1.2.2 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα

1.2.2.1 Ανάλυση Συναισθήματος σε Μεταφορική Γλώσσα

Η ΑΣ σε μεταφορική γλώσσα προτάθηκε αρχικά στο διαγωνισμό Semantic Evaluation 2015 - Task 11 (S15-T11) [4]. Δεδομένου ενός συνόλου tweets που είναι πλούσια σε μεταφορά, σαρκασμό και ειρωνεία, ο στόχος ήταν να προσδιοριστεί εάν ένας χρήστης έχει ευφράσει θετικό, αρνητικό ή ουδέτερο συναίσθημα σε καθένα από αυτά.

Οι πρώτες μέθοδοι της βιβλιογραφίας για την αντιμετώπιση του προβλήματος αυτού βασίζονταν σε Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ), Δέντρα Αποφάσεων ή μοντέλα παλινδρόμησης, που λειτουργούσαν σε χειροποίητα χαρακτηριστικά και λεξικά (π.χ. Senti-Wordnet, Depeche Mood, American National Corpus, κ.λπ.) [15] [16]. Αργότερα, το [17] προσέγγισε το πρόβλημα ως πρόβλημα παλινδρόμησης και εκμεταλλεύτηκε μια αρχιτεκτονική ΣΝΔ. Οι λέξεις με περισσότερες από δύο εμφανίσεις χρησιμοποιήθηκαν ως κειμενικά χαρακτηριστικά εισόδου, ενώ τα hashtags που προσδιορίζουν το είδος της μεταφορικής έκφρασης αντικαταστάθηκαν από δυαδικούς δείκτες. Το DESC [18] ενσωμάτωσε ένα αμφίδρομο δίκτυο μακρυπρόθεσμης-βραχυπρόθεσμης μνήμης (ΑΜ-ΔΜΒΜ), ένα ΔΜΒΜ με μηχανισμό προσοχής και ένα MLP, με τα δύο πρώτα να τροφοδοτούνται με προεκπαιδευμένες ενσωματώσεις λέξεων GloVe [19] και το MLP να τροφοδοτείται με χειροποίητα χαρακτηριστικά tweet, όπως ζεύγη λέξεων. Το DESC επεκτάθηκε στο [20] χρησιμοποιώντας πλαισιωμένες ενσωματώσεις εισόδου από ένα προεκπαιδευμένο μοντέλο RoBERTa [21], οι οποίες τροφοδοτούνται σε ένα επαναλαμβανόμενο συνελικτικό νευρωνικό δίκτυο (ΕΣΝΔ). Το πρόβλημα προσεγγίστηκε ως ταξινόμηση κειμένου.

1.2.2.2 Μεταφορά Νευρωνικής Γνώσης

Η απόσταξη γνώσης είναι η πιο διαδεδομένη μορφή νευρωνικής μεταφοράς γνώσης. Συνεπάγεται την εκπαίδευση ενός μαθητή BNΔ σύμφωνα με την απόκριση ενός προεκπαιδευμένου BNΔ (δάσκαλος) σε κάποια δεδομένα εισόδου (το λεγόμενο «σύνολο δεδομένων μεταφοράς»), που συνήθως είναι μη επισημασμένα. Συνήθως, ο μαθητής θεωρείται ότι είναι αρχικά μη εκπαιδευμένος, ενώ τα δύο δίκτυα μπορεί να έχουν διαφορετικές νευρωνικές αρχιτεκτονικές [6]. Το μαθητικό BNΔ εκπαιδεύεται με μια αντιειμενική συνάρτηση κόστους που μετρά την απόκλιση της εξόδου του από την αντίστοιχη έξοδο του αμετάβλητου δασκάλου BNΔ στο σύνολο μεταφοράς [6]. Ο στόχος είναι η μεταφορά της λεγόμενης σκοτεινής γνώσης από τον δάσκαλο στον μαθητή.

Αρκετές παραλλαγές αυτού του βασικού συστήματος νευρωνικής απόσταξης έχουν εμφανιστεί όλα αυτά τα χρόνια. Για παράδειγμα, σε περίπτωση που το σύνολο μεταφοράς είναι γνωστό/επισημασμένο, οι παράμετροι/βάρη του εκπαιδευμένου δασκάλου μπορούν να χρησιμοποιηθούν για την έξυπνη αρχικοποίηση του μαθητή πριν προχωρήσουμε στην τακτική εποπτευόμενη εκπαίδευση [22]. Πληροφορία/χαρακτηριστικά από ενεργοποιήσεις ενδιάμεσων κρυφών επιπέδων στον δάσκαλο κατά την επεξεργασία του συνόλου μεταφοράς μπορούν επίσης να χρησιμοποιηθούν για την καθοδήγηση της εκπαίδευσης των μαθητών, εκτός από το τελικό αποτέλεσμα του επιπέδου εξόδου του δασκάλου [23]. Μπορεί να χρειαστεί μείωση διαστάσεων όταν οι διαστάσεις του επιπέδου διαφέρουν μεταξύ μαθητή και δασκάλου. Για να παρακαμφθεί αυτός ο περιορισμός, η εκπαίδευση των μαθητών στο [24] καθοδηγείται από τις ομοιότητες μεταξύ των αναπαραστάσεων δειγμάτων δεδομένων συνόλου μεταφοράς που κατασκευάζονται από τον δάσκαλο. Ενδιάμεσα δίκτυα «βιοηθού δασκάλου», στο πλαίσιο μιας διαδικασίας πολλαπλών βημάτων, προτάθηκαν επίσης στο [25] για σενάρια όπου τα δύο μοντέλα διαφέρουν ως προς την πολυπλοκότητα. Τέλος, παρουσιάστηκε απόσταξη από έναν βαθύ γραμμικό δάσκαλο για δυαδική ταξινόμηση στο πλαίσιο μιας καθαρά θεωρητικής ανάλυσης στο [26], ενώ παρόμοια έρευνα πραγματοποιήθηκε στο [27] για ρηχά νευρωνικά δίκτυα.

1.3 Προτεινόμενες Μέθοδοι

Προκειμένου να επιλυθούν τα παραπάνω προβλήματα διερευνήθηκαν και υλοποιήθηκαν μέθοδοι που αξιοποιούν σύγχρονες τεχνολογίες βαθιάς μάθησης. Περισσότερες λεπτομέρειες σχετικά με την υλοποίηση των δύο μεθόδων είναι διαθέσιμες στις ενότητες 4.4, 5.4, αντίστοιχα.

1.3.1 Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet

Ο προτεινόμενος νέος αυτοματοποιημένος μηχανισμός παρακολούθησης της κοινής γνώμης αποτελείται από έναν σύνθετο, ποσοτικό, σημασιολογικό περιγραφέα που βασίζεται σε ταξινομητές που βασίζονται σε ΕΦΓ/ΒΝΔ. Με τη χρησιμοποίησή τους, ένα τετραδιάστατο διάνυσμα, δηλαδή, ένα δείγμα του προτεινόμενου περιγραφέα, εξάγεται πρώτα για νάθε tweet ανεξάρτητα, ποσοτικοποιώντας έτσι την πολικότητα, προσβλητικότητα, μεροληψία και μεταφορικότητα. Στη συνέχεια, οι υπολογισμένοι περιγραφείς συναθροίζονται για πολλά tweets, σύμφωνα με μια επιθυμητή στρατηγική συνάθροισης (π.χ. αριθμητικός μέσος όρος) και στόχο συνάθροισης (π.χ. μια συγκεκριμένη χρονική περίοδο). Οι συνοπτικοί περιγραφείς μπορούν να αξιοποιηθούν με διάφορους τρόπους: για παράδειγμα, με τη συνάθροισή τους σε καθημερινή βάση μας επιτρέπει να κατασκευάσουμε μια πολυμεταβλητή χρονοσειρά που μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός ΒΝΔ για την πρόβλεψη μελλοντικών περιγραφέων κοινής γνώμης.

1.3.2 Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα

Η προτεινόμενη μέθοδος εκμεταλλεύεται την απόσταξη γνώσης [6] για την αύξηση της ακρίβειας ενός ΒΝΔ κατά την ΑΣ σε μεταφορικά κείμενα, που περιέχουν σαρκασμό, ειρωνεία ή/και μεταφορά. Έτσι, κατά τη διάρκεια της εκπαίδευσης, χρησιμοποιείται μια αρχιτεκτονική δασκάλου-μαθητή για να εμπλουτιστεί το μοντέλο του μαθητή με τη γνώση ενός προεκπαιδευμένου αναγνωριστή μεταφορικής γλώσσας. Ο τελευταίος είναι ένας δυαδικός ταξινομητής, ενώ ο πρώτος είναι ένας ταξινομητής πολλαπλών κλάσεων που έχει ως αποστολή την αναγνώριση συναισθημάτων στα εισερχόμενα κείμενα. Έτσι, λόγω της φύσης του δασκάλου και της διαφορετικής εργασίας που επιλύει σε σύγκριση με τον μαθητή, προτείνεται ένα άτυπο είδος απόσταξης που δεν έχει χρησιμοποιηθεί προηγουμένως σε κανονικά ΒΝΔ για πραγματικά προβλήματα.

1.4 Δομή Περιεχομένου

Το υπόλοιπο αυτής της εργασίας οργανώνεται με τον ακόλουθο τρόπο. Στο κεφάλαιο 2 γίνεται μια αρκετά εκτενή εισαγωγή στον ευρύ κλάδο της ΕΦΓ. Εκεί δίνονται βασικοί ορισμοί 2.1, γίνεται μια σύντομη ιστορική αναδρομή 2.2, περιγράφονται βασικές μέθοδοι που συναντώνται στη βιβλιογραφία 2.3 και βλέπουμε πως μετατρέπεται το κείμενο σε αριθμούς 2.4. Στη συνέχεια, στο κεφάλαιο 3 αναλύεται σε βάθος το πρόβλημα της Ανάλυσης Συναισθημάτων. Συγκεκριμένα, ορίζεται το πρόβλημα 3.1 και περιγράφονται διάφοροι τρόποι επίλυσής

του 3.2. Τέλος, στα κεφάλαια 4, 5 παρουσιάζεται το κυρίως πειραματικό κομμάτι αυτής της εργασίας, όπου περιγράφονται αναλυτικά τα δύο διαφορετικά προβλήματα 4.2, 5.2, οι υπάρχουσες σχετικές προσεγγίσεις της βιβλιογραφίας 4.3, 5.3, οι προτεινόμενες μέθοδοι 4.4, 5.4 καθώς και η πειραματική αξιολόγηση και επαλήθευση της λειτουργικότητάς τους 4.5, 5.5.

Κεφάλαιο 2

Επεξεργασία Φυσικής Γλώσσας

Σε αυτό το πρώτο κεφάλαιο θα εξερευνήσουμε τον ευρύ κλάδο της Επεξεργασίας Φυσικής Γλώσσας (ΕΦΓ). Θα δούμε κάποιους ορισμούς για την ΕΦΓ και θα κάνουμε μία ιστορική αναδρομή. Επιπλέον, συζητάμε σχετικά με το κομμάτι της υλοποίησης, διαφορετικές μεθόδους και αλγόριθμους που χρησιμοποιούνται στην ΕΦΓ και πώς η γλώσσα μοντελοποιείται σαν επεξεργάσιμα δεδομένα. Τέλος, παρουσιάζουμε σύντομα μερικά συνήθη προβλήματα της ΕΦΓ.

2.1 Τί είναι η ΕΦΓ;

2.1.1 Σύντομοι Ορισμοί

Μπορούμε να βρούμε πολλούς ορισμούς στο διαδίκτυο για την ΕΦΓ. Προφανώς όλοι είναι παρόμοιοι αφού περιγράφουν το ίδιο πράγμα, ωστόσο εγώ επέλεξα τους παρακάτω σαν πιο απλοϊκούς και ακριβείς.

Ορισμός 1 : Η Επεξεργασία Φυσικής Γλώσσας (ΕΦΓ) μπορεί να οριστεί ως η αυτόματη διαχείριση φυσικής γλώσσας, όπως η ομιλία και το κείμενο από κάποιου είδους λογισμικό [28].

ή,

Ορισμός 2 : Η Επεξεργασία Φυσικής Γλώσσας αποτελείται από αυτόματες μεθόδους που δέχονται φυσική γλώσσα σαν είσοδο ή παράγουν φυσική γλώσσα σαν έξοδο [28].

Μερικοί άνθρωποι μπορεί να σκεφτούν ότι η ΕΦΓ είναι ένας τελείως καινούργιος κλάδος αλλά στην πραγματικότητα υπάρχει για παραπάνω από 50 χρόνια. Η ΕΦΓ αναδύθηκε από τον κλάδο της γλωσσολογίας με την ανάπτυξη των υπολογιστών και έχει σημειώσει τεράστια

πρόοδο την τελευταία δεκαετία χάρη στις εκπληκτικές ανακαλύψεις που έγιναν από τους ερευνητές.

2.1.2 Φυσική Γλώσσα

Η φυσική γλώσσα μπορεί απλά να οριστεί ως ο τρόπος με τον οποίο εμείς οι άνθρωποι επικοινωνούμε μεταξύ μας. Η λέξη φυσική υποδεικνύει ότι αυτή η μορφή γλώσσας αναπτύχθηκε με φυσικό τρόπο κατά τη χρήση, σε αντίθεση με την τεχνητή γλώσσα ή τον κώδικα των υπολογιστών. Όπως κάθε γλώσσα έτσι και η φυσική γλώσσα περιλαμβάνει ένα ευρύ σετ λέξεων και κανόνων που μας λένε πως να τις συνδυάσουμε προκειμένου να επικοινωνήσουμε το επιθυμητό μήνυμα. Αυτά τα δύο σετ διαφέρουν από χώρα σε χώρα δίνοντας μας όλες τις διαφορετικές γλώσσες που υπάρχουν στον κόσμο (Αγγλικά, Γαλλικά, Γερμανικά). Η φυσική γλώσσα μπορεί να χρησιμοποιηθεί με δύο τρόπους, προφορικά (ομιλία) και γραπτά (κείμενο). Και τα δύο χρησιμοποιούνται από τους ανθρώπους για να εκφράσουν τη σκέψη τους, καθένα στην περίπτωση που ταιριάζει περισσότερο. Συνεπώς γίνεται αντιληπτή η σημασία αυτού του είδους δεδομένων και η ανάγκη να βρεθούν μέθοδοι για να κατανοήσουμε και να επεξεργαστούμε τη φυσική γλώσσα.

2.1.3 Πρόκληση

Η φυσική γλώσσα είναι συχνά δυσνόητη και διφορούμενη. Υπάρχουν λίγοι κανόνες και η γλώσσα συνεχώς αλλάζει και εξελίσσεται. Οι άνθρωποι είναι πολύ καλοί στο να παράγουν γλώσσα και να την καταλαβαίνουν και είναι ικανοί να εκφράζουν, να αντιλαμβάνονται και να ερμηνεύουν πολύπλοκα νοήματα. Ταυτόχρονα ενώ εμείς οι άνθρωποι είμαστε σπουδαίοι χρήστες της γλώσσας, είμαστε επίσης ικανοί στο να διατυπώνουμε αυστηρά τους κανόνες πού εφαρμόζονται στη γλώσσα. Είναι λοιπόν δύσκολο να δουλέψεις με τέτοιους είδους δεδομένα.

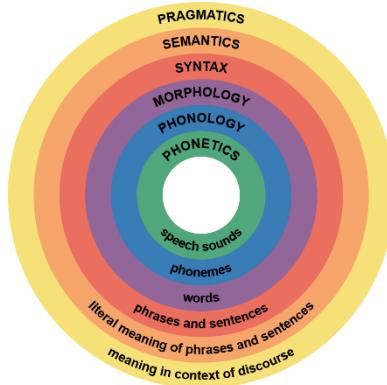
2.1.4 Από τη γλωσσολογία στην ΕΦΓ

2.1.4.1 Γλωσσολογία

Γλωσσολογία είναι η επιστημονική μελέτη της γλώσσας και της δομής της [29]. Τα πιο σημαντικά επίπεδα γλωσσικής δομής μπορούμε να τα δούμε στο σχήμα 2.1.

- *Πραγματολογία* είναι η μελέτη του πως τα συμφραζόμενα συνεισφέρουν στο νόημα.
- *Σημασιολογία* είναι η μελέτη του νοήματος των λέξεων και των προτάσεων.

- Η Σύνταξη αφορά την μελέτη της δομής μιας πρότασης και του πως αυτή αλληλεπιδρά με άλλες διαστάσεις της γλωσσικής πληροφορίας.
- Μορφολογία είναι η μελέτη των λέξεων, πώς αυτές σχηματίζονται και η σχέση τους με άλλες λέξεις της ίδιας γλώσσας.
- Φωνολογία είναι η μελέτη των μοτίβων ήχου σε μία γλώσσα.
- Φωνητική είναι η μελέτη των ήχων ομιλίας και της παραγωγής τους.



Σχήμα 2.1: Τα βασικά επίπεδα της γλωσσικής δομής.

2.1.4.2 Υπολογιστική Γλωσσολογία

Υπολογιστική γλωσσολογία (ΥΓ) είναι η σύγχρονη επιστήμη της γλωσσολογίας που χρησιμοποιεί τα εργαλεία της επιστήμης των υπολογιστών. Οι υπολογιστές μας βοηθούν να διαχειριστούμε μεγάλο όγκο δεδομένων κειμένου αποτελεσματικά και έτσι ανοίγουν νέες περιοχές έρευνας που ήταν αδύνατο να εξερευνηθούν πριν.

2.1.4.3 ΥΓ vs ΕΦΓ

Η υπολογιστική γλωσσολογία έχει επιστημονική και μηχανική πλευρά.

- Μηχανική πλευρά - Επεξεργασία Φυσικής Γλώσσας (ΕΦΓ): Αφορά τη δημιουργία υπολογιστικών εργαλείων που κάνουν χρήσιμα πράγματα με τη γλώσσα.
- Επιστημονική πλευρά - Κατανόηση Φυσικής Γλώσσας (ΚΦΓ): Αναζητά τη κατανόηση της γλώσσας χρησιμοποιώντας τους υπολογιστές και μεγάλο όγκο δεδομένων κειμένου.

Και οι δύο πλευρές χρησιμοποιούν τα ίδια μέσα αλλά με διαφορετικό σκοπό. Οι ερευνητές της ΕΦΓ θα φτιάχουν ένα χρήσιμο σύστημα και θα δείξουν ότι αυτό δουλεύει πολύ καλά. Οι ερευνητές υπολογιστικής γλωσσολογίας ενδιαφέρονται περισσότερο να μάθουν ποια γλωσσικά χαρακτηριστικά είναι χρήσιμοι δείκτες κάποιου νοήματος και γιατί.

2.2 Ιστορική Αναδρομή

2.2.1 Πίζες

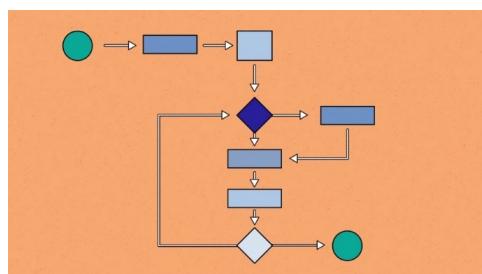
Η επεξεργασία φυσικής γλώσσας έχει τις ρίζες της στο 1950. Από το 1950, ο Alan Turing δημοσίευσε ένα άρθρο με τίτλο “Computing machinery and Intelligence” [30], το οποίο πρότεινε αυτό που σήμερα ονομάζεται Turing Test σαν κριτήριο νοημοσύνης. Με άλλα λόγια, αυτό το τεστ εξετάζει αν μία μηχανή παρουσιάζει ανθρώπινη νοημοσύνη. Το τεστ αυτό περιλαμβάνει την αυτόματη ερμηνεία και παραγωγή φυσικής γλώσσας.

2.2.2 Τρεις Περίοδοι

- Συμβολική ΕΦΓ (1950 - 1990)
 - Στατιστική ΕΦΓ (1990 - 2010)
 - Νευρωνική ΕΦΓ (2010 - τώρα)

2.2.2.1 Συμβολική ΕΦΓ (κλασικός προγραμματισμός)

Δοθείσας μιας συλλογής από χειρόγραφους κανόνες ο υπολογιστής προσομοιώνει την κατανόηση φυσικής γλώσσας εφαρμόζοντας αυτούς τους κανόνες στα δεδομένα με τα οποία τροφοδοτείται [31]. Ένα σετ τέτοιων κανόνων καλείται αλγόριθμος στον προγραμματισμό και μπορεί να οπτικοποιηθεί με ένα διάγραμμα ροής (σχήμα 2.2).



Σχήμα 2.2: Διάγραμμα Ροής - Η διαγραμματική αναπαράσταση ενός αλγορίθμου.

Επιτεύγματα

- **1950** : Το πείραμα του Georgetown το 1954 [32] περιλάμβανε την αυτόματη μετάφραση περισσότερων από 60 ρωσικών προτάσεων στα αγγλικά. Οι συγγραφείς ισχυρίστηκαν ότι μέσα σε 3 ή 5 χρόνια, η αυτόματη μετάφραση θα ήταν ένα λυμένο πρόβλημα. Ωστόσο η πραγματική πρόοδος ήταν πολύ πιο αργή.

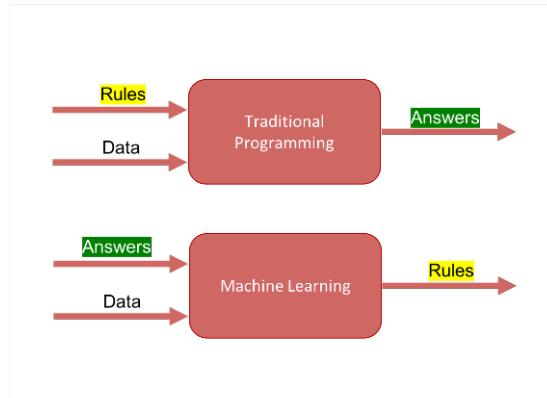
- **1960** : Ένα ιδιαίτερα επιτυχημένο σύστημα επεξεργασίας φυσικής γλώσσας ήταν η ELIZA [33], μία προσομοίωση ψυχοθεραπευτή εμπνευσμένη από τον Joseph Weizenbaum μεταξύ 1964 και 1966. Χρησιμοποιώντας σχεδόν καθόλου πληροφορία σχετικά με την ανθρώπινη σκέψη ή το συναίσθημα η ELIZA μερικές φορές παρείχε μία ανθρώπινη αλληλεπίδραση. Όταν ο ασθενής ξέφευγε από τα όρια της πολύ περιορισμένης γνώσης της, αυτή θα έδινε μία γενική απόκριση. Για παράδειγμα στην φράση “Το κεφάλι μου πονάει” θα απαντούσε “Γιατί λες ότι το κεφάλι σου πονάει;”
- **1970** : Κατά τη διάρκεια αυτής της περιόδου γράφτηκαν τα πρώτα chatbot (π.χ. PARRY [34]). Το chatbot είναι μία εφαρμογή λογισμικού που χρησιμοποιείται για τη διεξαγωγή μιας online συζήτησης μέσω κειμένου.
- **1980** : Η δεκαετία του 1980 και οι αρχές της δεκαετίας του 1990 σηματοδοτούν την αιμή των συμβολικών μεθόδων της ΕΦΓ. Δημοφιλείς περιοχές μελέτης της εποχής περιλάμβαναν την ανάλυση κειμένου βάσει κανόνων (π.χ., η ανάπτυξη του HPSG [35]) και τη μελέτη σημασιολογίας (π.χ., ο αλγόριθμος Lesk [36]).

2.2.2.2 Στατιστική ΕΦΓ – Μηχανική Μάθηση

Μέχρι τη δεκαετία του 1980, τα περισσότερα συστήματα επεξεργασίας φυσικής γλώσσας βασίζονταν σε πολύπλοκα σετ χειρόγραφων κανόνων. Στο τέλος της δεκαετίας του 1980 όμως έγινε μία στροφή προς τη στατιστική προσέγγιση του αντικειμένου με την εισαγωγή της μηχανικής μάθησης για επεξεργασία γλώσσας [31].

Μηχανική Μάθηση

Η μηχανική μάθηση είναι ένας ιλαράδος της τεχνητής νοημοσύνης που εστιάζει στη χρήση των δεδομένων και των αλγορίθμων για να μιμηθεί τον τρόπο με τον οποίο οι άνθρωποι μαθαίνουν, βελτιώνοντας σταδιακά την επίδοσή τους [37]. Η διαφορά μεταξύ μηχανικής μάθησης και των συμβατικών χειρόγραφων αλγορίθμων φαίνεται στο σχήμα 2.3.



Σχήμα 2.3: Μηχανική Μάθηση vs Προσέγγιση βάσει Κανόνων

Οι ολγόριθμοι μηχανικής μάθησης κατηγοριοποιούνται με βάση το είδος των δεδομένων που χρησιμοποιούν για την εκπαίδευσή τους. Υπάρχουν δύο είδη δεδομένων, τα επισημασμένα και τα μη επισημασμένα. Τα επισημασμένα δεδομένα περιλαμβάνουν τις παραμέτρους εισόδου και εξόδου σε μία μορφή απολύτως κατανοητή από τον αλγόριθμο αλλά χρειάζεται πολλή δουλειά για να επισημανθούν τα δεδομένα χειροκίνητα με την επιθυμητή έξοδο. Τα μη επισημασμένα δεδομένα περιλαμβάνουν μόνο τις παραμέτρους εισόδου. Αυτό μας γλιτώνει από την ανάγκη να δαπανήσουμε χρόνο και ενέργεια για την επισήμανση των δεδομένων αλλά απαιτεί πιο συνθέτες λύσεις. Οι προκύπτουσες κατηγορίες αλγορίθμων μηχανικής μάθησης είναι η εποπτευόμενη και μη εποπτευόμενη μάθηση αντίστοιχα [38].

Εποπτευόμενη Μάθηση

Στη εποπτευόμενη μάθηση, ο αλγόριθμος βρίσκει τις σχέσεις μεταξύ των παραμέτρων εισόδου και εξόδου. Στο τέλος της εκπαίδευσης ο αλγόριθμος αποκτά μία ιδέα του πως συνδέονται οι είσοδοι με τις αντίστοιχες εξόδους και είναι σε θέση να παράγει τη σωστή έξοδο για μια νέα είσοδο [39]. Τα προβλήματα της εποπτευόμενης μάθησης μπορούνε περαιτέρω να κατηγοριοποιηθούν σε προβλήματα Ταξινόμησης και Παλινδρόμησης [40].

- **Ταξινόμηση :** Ένα πρόβλημα ταξινόμησης έχουμε όταν η μεταβλητή εξόδου είναι μία κατηγορία. Για παράδειγμα "κόκκινο" ή "μπλε" (κατηγορίες χρωμάτων), "αρρώστια" ή "όχι αρρώστια" (κατηγορίες ασθένειας).
- **Παλινδρόμηση :** Ένα πρόβλημα παλινδρόμησης έχουμε όταν η μεταβλητή εξόδου είναι μία πραγματική τιμή, όπως δολάρια ή βάρος.

Μη-Εποπτευόμενη Μάθηση

Στην μη εποπτευόμενη μάθηση το εκπαιδευόμενο μοντέλο δεν έχει τις επιθυμητές εξόδους για κάθε είσοδο για να βρει τις σχέσεις μεταξύ τους και συνεπώς οδηγείται στη δημιουργία κρυφών δομών. Οι σχέσεις μεταξύ διαφορετικών σημείων των δεδομένων γίνονται αντιληπτές από τον αλγόριθμο με έναν αφηρημένο τρόπο. Η δημιουργία αυτών των κρυφών δομών είναι που κάνει την μη εποπτευόμενη μάθηση ευέλικτη. Αυτού του είδους οι αλγόριθμοι μπορούν να προσαρμόζονται στα δεδομένα αλλάζοντας δυναμικά τις κρυφές δομές [41]. Γενικά αυτό το πρόβλημα είναι πολύ πιο δύσκολο από την εποπτευόμενη μάθηση και τυπικά παράγει λιγότερο ακριβή αποτελέσματα για ένα συγκεκριμένο όγκο δεδομένων εισόδου. Ωστόσο η τεράστια ποσότητα από μη επισημασμένα δεδομένα μπορεί συχνά να αντισταθμίσει τις χαμηλότερες επιδόσεις εφόσον ο αλγόριθμος που χρησιμοποιείται έχει αρκετά χαμηλή πολυπλοκότητα για να είναι πρακτικός. Τα προβλήματα μη εποπτευόμενης μάθησης μπορούνε περαιτέρω να κατηγοριοποιηθούν σε προβλήματα Ομαδοποίησης και Συσχέτισης [40].

- **Ομαδοποίηση :** Σε αυτού του είδους τα προβλήματα παρόμοια δεδομένα ομαδοποιούνται υπολογίζοντας την απόσταση μεταξύ των σημείων.
- **Συσχέτιση :** Ένα πρόβλημα μάθησης κανόνων συσχέτισης είναι όταν θέλουμε να ανακαλύψουμε κανόνες που περιγράφουν μεγάλα τμήματα των δεδομένων μας όπως για παράδειγμα “άνθρωποι που αγοράζουν X επίσης τείνουν να αγοράσουν Ψ”.

Γιατί Στατιστική ΕΦΓ;

Οι κύριοι λόγοι που οδήγησαν σε αυτή την επανάσταση στην ΕΦΓ ήταν οι ακόλουθοι [31] :

- Αύξηση της υπολογιστικής δύναμης (νόμος του Μοορε).
- Σταδιακή μείωση της κυριαρχίας των Chomskyan θεωριών της γλωσσολογίας που αποθάρρυναν την προσέγγιση της μηχανικής μάθησης.
- Καλύτερα αποτελέσματα, ταχύτητα και ευρωστία.
- Αυξημένος όγκος δεδομένων λόγω του διαδικτύου.
- Απομυθοποίηση της συμβολικής προσέγγισης.

Επιτεύγματα

- **1990 :** Πολλές από τις αξιοσημείωτες επιτυχίες των στατιστικών μεθόδων στην ΕΦΓ συνέβησαν στον ιλαδό της Αυτόματης Μετάφρασης χάρη στην ερευνητική ομάδα της IBM. Σχεδιάστηκαν συστήματα που εκμεταλλεύονταν τα υπάρχοντα πολυ-γλωσσικά δεδομένα κειμένου. Επιπλέον σημαντική έρευνα έγινε για την δημιουργία μεθόδων που μαθαίνουν πιο αποτελεσματικά από περιορισμένο όγκο δεδομένων.

- **2000 :** Με την ανάπτυξη του διαδικτύου αυξανόμενος όγκος μη επισημασμένων γλωσσικών δεδομένων έγινε διαθέσιμος από τα μέσα της δεκαετίας του 1990. Η έρευνα λοιπόν εστίασε στην μη εποπτευόμενη και μερικώς εποπτευόμενη (συνδυασμός επισημασμένων και μη επισημασμένων δεδομένων) μάθηση.

2.2.2.3 Νευρωνική ΕΦΓ

Η νευρωνική ΕΦΓ αναφέρεται στη χρήση βαθιάς μάθησης (νευρωνικά δίκτυα) για την επεξεργασία φυσικής γλώσσας. Η βαθιά μάθηση είναι ένας κλάδος της μηχανικής μάθησης που ασχολείται με αλγόριθμος εμπνευσμένους από τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου και ονομάζονται τεχνητά νευρωνικά δίκτυα.

Τεχνητό Νευρωνικό Δίκτυο (TNΔ)

Ένα τεχνητό νευρωνικό δίκτυο βασίζεται σε μία συλλογή από συνδεδεμένους κόμβους που ονομάζονται τεχνητοί νευρώνες και μοντελοποιούν τους νευρώνες σε έναν βιολογικό εγκέφαλο. Κάθε σύνδεση όπως οι συνάψεις στον ανθρώπινο εγκέφαλο μπορεί να μεταφέρει ένα σήμα σε άλλους νευρώνες. Ένας τεχνητός νευρώνας δέχεται σήματα, μετά τα επεξεργάζεται και δίνει νέα σήματα σε νευρώνες που συνδέονται με αυτόν. Το σήμα/έξοδος κάθε νευρώνα είναι ένας πραγματικός αριθμός και υπολογίζεται από μία μη γραμμική συνάρτηση του αθροίσματος των εισόδων της. Οι συνδέσεις ονομάζονται ακμές. Οι ακμές έχουν ένα βάρος που προσαρμόζεται κατά τη μάθηση. Το βάρος αυτό αυξάνει ή μειώνει τη δύναμη του σήματος σε μία σύνδεση. Οι νευρώνες μπορεί να έχουν ένα κατώφλι έτσι ώστε ένα σήμα να στέλνεται σε επόμενους νευρώνες μόνο αν ξεπερνά αυτό το κατώφλι. Τυπικά οι νευρώνες ομαδοποιούνται σε επίπεδα. Διαφορετικά επίπεδα πραγματοποιούν διαφορετικούς μετασχηματισμούς στις εισόδους τους. Σήματα ταξιδεύουν από το πρώτο επίπεδο (επίπεδο εισόδου) μέχρι το τελευταίο επίπεδο (επίπεδο εξόδου) πολλαπλές φορές [42].

Η βαθιά μάθηση έχει τις ρίζες της πίσω στο 1962 όταν ο Frank Rosenblatt περιέγραψε όλα τα βασικά συστατικά των συστημάτων βαθιάς μάθησης στο βιβλίο του "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms" [43]. Ο πρώτος γενικός λειτουργικός αλγόριθμος βαθειάς εποπτευόμενης μάθησης δημοσιεύτηκε από τον Alexey Ivakhnenko το 1967 [44]. Ωστόσο τα νευρωνικά δίκτυα ξεκίνησαν να χρησιμοποιούνται αποτελεσματικά στην ΕΦΓ τη δεκαετία του 2000 και εκτοξεύτηκαν στη δεκαετία του 2010 με αξιοθαύμαστες ανακαλύψεις [45].

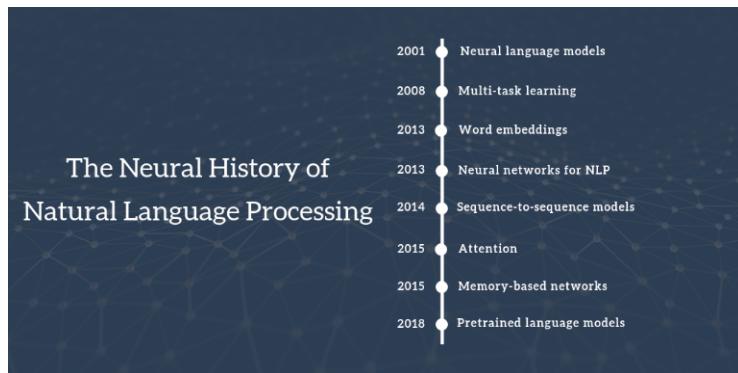
Επιτεύγματα

- **2001 - Νευρωνικά Γλωσσικά Μοντέλα :** Το πρώτο νευρωνικό γλωσσικό μοντέλο, ένα νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης προτάθηκε από τον Bengio [46]. Το πρόβλημα της μοντελοποίησης γλώσσας περιλαμβάνει την πρόβλεψη της επόμενης λέξης σε ένα κείμενο παίρνοντας σαν είσοδο τις προηγούμενες λέξεις. Τέτοια μοντέλα χρησιμοποιούνται για την μάθηση καλών διανυσμάτων αναπαράστασης λέξεων (ενότητα 2.4.2).
- **2008 - Μάθηση Πολλαπλών Προβλημάτων :** Η μάθηση πολλαπλών προβλημάτων εφαρμόστηκε πρώτα στα νευρωνικά δίκτυα για την ΕΦΓ από τους Collobert και Weston [47]. Στην αρχιτεκτονική τους οι πίνακες των διανυσμάτων λέξεων είναι κοινοί μεταξύ των δύο μοντέλων που εκπαιδεύονται σε διαφορετικά προβλήματα. Η μάθηση πολλαπλών προβλημάτων είναι μία γενική μέθοδος για τη χρήση κοινών παραμέτρων μεταξύ μοντέλων που εκπαιδεύονται σε διαφορετικά προβλήματα. Η κοινή χρήση διανυσμάτων λέξεων επιτρέπει στα μοντέλα να συνεργάζονται και να μοιράζονται γενική χαμηλού επιπέδου πληροφορία από τον πίνακα διανυσμάτων λέξεων, που συνήθως αποτελεί το μεγαλύτερο όγκο παραμέτρων σε ένα μοντέλο. Αυτό το paper αποδείχτηκε επιδραστικό και πέρα από τη χρήση του για μάθηση πολλαπλών προβλημάτων. Διέδωσε ιδέες όπως η προ-εκπαίδευση διανυσμάτων λέξεων και η χρήση συνελικτικών νευρωνικών δικτύων για κείμενο, τα οποία έχουν ευραίως υιοθετηθεί σε εφαρμογές ΕΦΓ τα τελευταία χρόνια.
- **2013 - Ενσωματώσεις Λέξεων :** Διανύσματα λέξεων χρησιμοποιούνται από το 2001 όπως είδαμε παραπάνω. Η κύρια καινοτομία που προτάθηκε το 2013 από τον Mikolov [48] ήταν να γίνει η εκπαίδευση αυτών των διανυσμάτων λέξεων πιο αποτελεσματική αφαιρώντας το κρυφό επίπεδο της αρχιτεκτονικής του Bengio [46] και να γίνει προσέγγιση του στόχου. Αυτές οι αλλαγές μαζί με την αποτελεσματική υλοποίηση του Word2Vec (βλέπε 2.4.2.1) επέτρεψαν την εκπαίδευση διανυσμάτων λέξεων σε μεγάλη κλίμακα. Ενώ αυτά τα διανύσματα δεν διαφέρουν σαν ιδέα από αυτά που μαθαίνει ένα νευρωνικό εμπρόσθιας τροφοδότησης, η εκπαίδευση με πολύ μεγάλα κείμενα τους επιτρέπει να προσεγγίσουν συγκεκριμένες σχέσεις μεταξύ λέξεων όπως το γένος, χρόνους ρημάτων, και να ενσωματώσουν το νόημα των λέξεων (ενσωματώσεις). Αυτό που καθιέρωσε τη χρήση διανυσμάτων λέξεων σε προβλήματα ΕΦΓ ήταν ότι χρησιμοποιώντας προ-εκπαίδευμένα διανύσματα κατα την αρχικοποίηση του μοντέλου, παρατηρούμε βελτιωμένη απόδοση σε ένα μεγάλο εύρος προβλημάτων.
- **2013 - Νευρωνικά Δίκτυα για ΕΦΓ :** Το 2013 και 2014 ήταν οι χρονιές που τα νευρωνικά δίκτυα ξεκίνησαν να υιοθετούνται στην ΕΦΓ. Δύο βασικοί τύποι νευρωνικών δικτύων γίναν οι πιο δημοφιλείς: Τα Επαναλαμβανόμενα Νευρωνικά Δίκτυα (ΕΝΔ) (βλέπε 2.3.3.2) και τα Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ) (βλέπε 2.3.3.2). Πριν το 2013 τα ΕΝΔ θεωρούνταν ακόμα δύσκολο να τα εκπαιδεύσουμε. Ο Ilya Sutskever στην

διδακτορική διατριβή του [49] άλλαξε αυτήν την φήμη. Τα ΣΝΔ πέρα από την ευρεία χρήση τους στην υπολογιστική όραση ξεκίνησαν να χρησιμοποιούνται και για την ΕΦΓ (Kalchbrenner 2014 [50]- Kim 2014 [51]).

- **2014 - Ακολουθία-σε-Ακολουθία μοντέλα :** Ο Sutskever [52] πρότεινε την Ακολουθία-σε-Ακολουθία μάθηση, ένα γενικό πλαίσιο αντιστοίχισης μιας ακολουθίας σε μία άλλη χρησιμοποιώντας ένα νευρωνικό δίκτυο. Σε αυτό το πλαίσιο, ένα νευρωνικό κωδικοποιητής επεξεργάζεται μία πρόταση σύμβολο προς σύμβολο και τη συμπιέζει σε μία διανυσματική αναπαράσταση. Στη συνέχεια ένα νευρωνικό αποκωδικοποιητής προβλέπει την έξοδο σύμβολο προς σύμβολο βασιζόμενο στην έξοδο του κωδικοποιητή, παίρνοντας σαν είσοδο σε κάθε βήμα το σύμβολο που προέβλεψε αμέσως πριν. Η αυτόματη μετάφραση αποδείχθηκε η κύρια εφαρμογή αυτού του είδους μοντέλων. Κωδικοποιητές ακολουθιών και αποκωδικοποιητές συνήθως βασίζονται σε επαναλαμβανόμενα νευρωνικά δίκτυα.
- **2015 - Προσοχή :** Ο μηχανισμός της Προσοχής (Bahdanau [53]) είναι μία από τις κύριες καινοτομίες στη Νευρωνική Αυτόματη Μετάφραση (NAM) και η βασική ιδέα που επέτρεψε στα μοντέλα βαθιάς μάθησης να ξεπεράσουν τα κλασικά συμβολικά συστήματα αυτόματης μετάφρασης. Η βασική αδυναμία της ακολουθίας-σε-ακολουθία μάθησης είναι ότι πρέπει να συμπιέσει όλη την ακολουθία εισόδου σε ένα διάνυσμα σταθερού μεγέθους. Ο μηχανισμός προσοχής αντιμετωπίζει αυτό το πρόβλημα επιτρέποντας στον αποκωδικοποιητή να κοιτάει πίσω στις κρυφές καταστάσεις της ακολουθίας εισόδου, οι οποίες στη συνέχεια παρέχονται σαν ένας σταθμισμένος μέσος σαν επιπρόσθετη είσοδος στον αποκωδικοποιητή. Πολλαπλά επίπεδα αυτο-προσοχής βρίσκονται στον πυρήνα της αρχιτεκτονικής του Μετασχηματιστή (Vaswani 2017 [54], βλέπε 2.4.2.2) που αποτελεί την τελευταία λέξη της τεχνολογίας στη NAM.
- **2015 - Δίκτυα Μνήμης :** Η Προσοχή αποτελεί ένα είδος μνήμης όπου η μνήμη αποτελείται από παρελθούσες κρυφές καταστάσεις στο μοντέλο το οποίο επιλέγει τι να ανακτήσει από τη μνήμη. Τα δίκτυα μνήμης που προτάθηκαν από τον Weston το 2015 [55] είναι μοντέλα με πιο ξεκάθαρη μνήμη. Η πρόσβαση στη μνήμη γίνεται συχνά με βάση την ομοιότητα με την τρέχουσα κατάσταση (παρόμοια με την Προσοχή) και συνήθως μπορούμε να γράψουμε σε αυτήν αλλά και να διαβάσουμε από αυτή. Τα δίκτυα μνήμης συνήθως εφαρμόζονται σε προβλήματα όπου είναι απαραίτητη η διατήρηση πληροφορίας σε βάθος χρόνου όπως η μοντελοποίηση γλώσσας και η κατανόηση κειμένου.
- **2018 – Προεκπαιδευμένα Γλωσσικά Μοντέλα (ΠΓΜ) :** Πρόκειται για μεγάλα νευρωνικά δίκτυα που χρησιμοποιούνται σε πληθώρα εφαρμογών ΕΦΓ. Αυτά τα μοντέλα εκπαιδεύονται πρώτα σε μεγάλα κείμενα, συνήθως σε προβλήματα πρόβλεψης λέξεων,

προκειμένου να μάθουν καλά διανύσματα λέξεων (δες 2.4.2.2). Στη συνέχεια φορτώνονται προ-εκπαιδευμένα και μερικές φορές βελτιώνονται με έξτρα εκπαίδευση σε κάποιο συγκεκριμένο πρόβλημα. Τα προ-εκπαιδευμένα γλωσσικά μοντέλα προτάθηκαν από τους Dai & Le [56] το 2015 και μόνο πρόσφατα αποδείχθηκαν αποτελεσματικά σε ένα μεγάλο εύρος προβλημάτων. Τα πιο δημοφιλή είναι τα CoVe (B. McCann, 2017 [57]), ELMo (M. Peters, 2018 [58]), GPT (OpenAI, 2018 [59]) και BERT (J. Devlin, 2018 [60]). Η χρήση διανυσμάτων λέξεων από προ-εκπαιδευμένα γλωσσικά μοντέλα μπορεί να βελτιώσει τη συνολική επίδοση αφού περιλαμβάνουν γνώση από μεγάλο όγκο δεδομένων. Επιπλέον, είναι προφανές ότι γλιτώνουμε χρόνο και χρειαζόμαστε λιγότερα δεδομένα για την εκπαίδευση ενός μοντέλου.



Σχήμα 2.4: Ιστορία Νευρωνικής ΕΦΓ.

2.3 Μέθοδοι: Κανόνες, Στατιστική, Νευρωνικά Δίκτυα

2.3.1 Κανόνες

Εδώ βρίσκονται οι μέθοδοι Συμβολικής ΕΦΓ. Αυτές οι μέθοδοι περιλαμβάνουν τη χειροκίνητη καθοικοποίηση (από τον προγραμματιστή) ενός συνόλου κανόνων που χειρίζονται τα δεδομένα εισόδου με αυστηρά ντετερμινιστικό τρόπο. Για παράδειγμα, ακολουθείται η λογική δομή IF-THEN. Όταν πληρούται μια γλωσσική συνθήκη IF, δημιουργείται η έξοδος THEN. Αυτό καθιστά εύκολη τη θέσπιση σαφών και εξηγήσιμων κανόνων, παρέχοντας πλήρη διαφάνεια στον τρόπο λειτουργίας του. Επίσης, συνήθως βρίσκουμε μια αναζήτηση λεξικού σε τέτοιους αλγόριθμους, όπως με τη σύνταξη γραμματικών ή την επινόηση ευρετικών κανόνων για την αποκοπή καταλήξεων.

2.3.2 Στατιστική

Η στατιστική επανάσταση στη δεκαετία του 1990 έκανε τους αλγόριθμους μηχανικής μάθησης να κυριαρχούν στον τομέα της ΕΦΓ. Σε αυτούς τους αλγόριθμους οι κανόνες δεν είναι

χειρόγραφοι όπως στη Συμβολική ΕΦΓ, αλλά μαθαίνονται αυτόματα μέσω της ανάλυσης μεγάλων κειμένων. Ένα μεγάλο σύνολο «χαρακτηριστικών» που παράγονται από τα δεδομένα εισόδου από τον προγραμματιστή (χειροκίνητα), τροφοδοτούνται ως είσοδος στον αλγόριθμο μηχανικής μάθησης. Στη συνέχεια, ο αλγόριθμος χρησιμοποιεί στατιστικά συμπεράσματα για να ανιχνεύσει τα μοτίβα στα χαρακτηριστικά εισόδου προκειμένου να παραχθεί το επιθυμητό αποτέλεσμα.

2.3.2.1 Γιατί Μηχανική Μάθηση;

Μειονεκτήματα χειροποίητων κανόνων:

- Δεν είναι καθόλου προφανές προς ποια κατεύθυνση πρέπει να δουλέψουμε.
- Ο χειρισμός άγνωστων ή λανθασμένων δεδομένων είναι εξαιρετικά δύσκολος.
- Τα συστήματα μπορούν να γίνουν πιο ακριβή μόνο αυξάνοντας την πολυπλοκότητα των κανόνων, κάτι που απαιτεί πολύ προσπάθεια.

2.3.2.2 Αλγόριθμοι Μηχανικής Μάθησης για ΕΦΓ

Εποπτευόμενη Μάθηση

- **Μηχανές Διανυσμάτων Υποστήριξης (MΔΥ) :** Δεδομένου ενός συνόλου παραδειγμάτων εκπαίδευσης, καθένα από τα οποία επισημαίνεται ότι ανήκει σε μία από τις δύο κατηγορίες, ένας αλγόριθμος εκπαίδευσης MΔΥ δημιουργεί ένα μοντέλο που εκχωρεί νέα παραδείγματα στη μία ή στην άλλη κατηγορία, καθιστώντας τον έναν μη πιθανολογικό δυαδικό γραμμικό ταξινομητή. Η MΔΥ αντιστοιχίζει παραδείγματα εκπαίδευσης σε σημεία στο χώρο, ώστε να μεγιστοποιήσει το πλάτος του χάσματος μεταξύ των δύο κατηγοριών. Στη συνέχεια, τα νέα παραδείγματα χαρτογραφούνται στον ίδιο χώρο και προβλέπεται ότι ανήκουν σε μια κατηγορία με βάση την πλευρά του κενού που πέφτουν [61].
- **Αφελής Bayes :** Οι ταξινομητές Αφελή Bayes είναι μια οικογένεια απλών «πιθανοτικών ταξινομητών» που βασίζονται στην εφαρμογή του θεωρήματος του Bayes με ισχυρές (αφελείς) υποθέσεις ανεξαρτησίας μεταξύ των χαρακτηριστικών, προκειμένου να ελαχιστοποιηθεί η πιθανότητα εσφαλμένης ταξινόμησης [62].
- **Μέγιστη Εντροπία :** Η αρχή της μέγιστης εντροπίας είναι ένας κανόνας δημιουργίας μοντέλου που απαιτεί την επιλογή της πιο απρόβλεπτης (μέγιστης εντροπίας) προηγούμενης υπόθεσης εάν είναι γνωστή μόνο μία παράμετρος για μια κατανομή πιθανότητας. Ο στόχος είναι να μεγιστοποιηθεί η «ομοιομορφία» ή η αβεβαιότητα όταν

γίνεται μια προηγούμενη υπόθεση πιθανότητας, έτσι ώστε η υποκειμενική προκατάληψη να ελαχιστοποιείται στα αποτελέσματα του μοντέλου [63].

- **Δέντρα Απόφασης :** Ο στόχος της χρήσης ενός Δέντρου Απόφασης είναι να δημιουργηθεί ένα μοντέλο εκπαίδευσης που μπορεί να χρησιμοποιηθεί για την πρόβλεψη της κλάσης (δέντρο ταξινόμησης) ή της τιμής (δέντρο παλινδρόμησης) της μεταβλητής στόχου μαθαίνοντας απλούς κανόνες απόφασης που συνάγονται από προηγούμενα δεδομένα (δεδομένα εκπαίδευσης). Στα Δέντρα Αποφάσεων, για την πρόβλεψη μιας ετικέτας κλάσης για μια εγγραφή ξεκινάμε από τη ρίζα του δέντρου. Συγκρίνουμε τις τιμές του χαρακτηριστικού ρίζας με το χαρακτηριστικό της εγγραφής. Με βάση τη σύγκριση, ακολουθούμε τον κλάδο που αντιστοιχεί σε αυτήν την τιμή και μεταπεδούμε στον επόμενο κόμβο. Σε ένα δέντρο απόφασης, τα φύλλα αντιπροσωπεύουν την τελική έξοδο (κλάση ή πραγματική τιμή) και οι κλάδοι αντιπροσωπεύουν συνδέσμους χαρακτηριστικών που οδηγούν σε αυτές τις τιμές εξόδου [64].
- **Τυχαία Δάση :** Τα Τυχαία Δάση ή τα Δάση Τυχαίας Απόφασης είναι μια μέθοδος εκμάθησης συνόλου για ταξινόμηση, παλινδρόμηση και άλλες εργασίες που λειτουργεί με την κατασκευή πολλών δέντρων απόφασης κατά το χρόνο εκπαίδευσης. Για εργασίες ταξινόμησης, η έξοδος του τυχαίου δάσους είναι η κλάση που επιλέγεται από τα περισσότερα δέντρα. Για εργασίες παλινδρόμησης, επιστρέφεται ο μέσος όρος ή ο μέσος όρος πρόβλεψης των μεμονωμένων δέντρων. Τα δάση τυχαίας απόφασης καταπολεμούν τη συνήθεια των δέντρων απόφασης να προσαρμόζονται υπερβολικά στο σετ εκπαίδευσής τους. Τα τυχαία δάση συνήθως υπερτερούν των δέντρων απόφασης [65].
- **K -Κοντινότεροι Γείτονες (κ -ΚΓ) :** Το κ -ΚΓ είναι μια μη παραμετρική μέθοδος που χρησιμοποιείται για ταξινόμηση και παλινδρόμηση. Και στις δύο περιπτώσεις, η είσοδος αποτελείται από τα κ πιο κοντινά παραδείγματα εκπαίδευσης σε ένα σύνολο δεδομένων. Η έξοδος εξαρτάται από το εάν το κ -NN χρησιμοποιείται για ταξινόμηση ή παλινδρόμηση [66] :
 - Στην ταξινόμηση κ -ΚΓ, η έξοδος είναι μια ιδιότητα μέλους κλάσης. Ένα αντικείμενο ταξινομείται με πλήθος ψηφοφοριών των γειτόνων του, με το αντικείμενο να εκχωρείται στην κλάση που είναι πιο κοινή μεταξύ των κ πλησιέστερων γειτόνων του (κ είναι ένας θετικός ακέραιος, συνήθως μικρός). Αν $\kappa = 1$, τότε το αντικείμενο απλώς εκχωρείται στην κλάση αυτού του απλού πλησιέστερου γειτόνα.
 - Στην παλινδρόμηση κ -ΚΓ, η έξοδος είναι η τιμή ιδιότητας για το αντικείμενο. Αυτή η τιμή είναι ο μέσος όρος των τιμών των κ κοντινότερων γειτόνων.

- **Iεραρχική Ομαδοποίηση:** Η iεραρχική ομαδοποίηση είναι μια μέθοδος που επιδιώκει να οικοδομήσει μια iεραρχία ομάδων. Οι στρατηγικές για την iεραρχική ομαδοποίηση γενικά χωρίζονται σε δύο τύπους:

- **Συγκεντρωτική :** Αυτή είναι μια προσέγγιση «από κάτω προς τα πάνω»: κάθε παρατήρηση ξεκινά στη δική της ομάδα και ζεύγη ομάδων συγχωνεύονται καθώς κάποιος προχωρά προς τα πάνω στην iεραρχία.
- **Διαιρετική :** Αυτή είναι μια προσέγγιση „από πάνω προς τα κάτω“: όλες οι παρατηρήσεις ξεκινούν σε μια ομάδα και οι διαχωρισμοί εκτελούνται αναδρομικά καθώς κάποιος κινείται προς τα κάτω στην iεραρχία.

Γενικά, οι συγχωνεύσεις και οι διασπάσεις καθορίζονται με άπληστο τρόπο. Τα αποτέλεσματα της iεραρχικής ομαδοποίησης παρουσιάζονται συνήθως σε ένα δενδρόγραμμα [67].

- **Ομαδοποίηση κ-μέσων :** Η ομαδοποίηση κ-μέσων είναι μια μέθοδος κβαντοποίησης διανυσμάτων, αρχικά από την επεξεργασία σήματος, που στοχεύει να χωρίσει ν παρατηρήσεις σε κ ομάδες. Κάθε παρατήρηση ανήκει στην ομάδα με τον πλησιέστερο μέσο όρο (κέντρα ομάδων), που χρησιμεύει ως πρωτότυπο της ομάδας. Αυτό έχει ως αποτέλεσμα την κατάτμηση του χώρου δεδομένων σε κελιά Voronoi. Η ομαδοποίηση κ-μέσων ελαχιστοποιεί τις διακυμάνσεις εντός της ομάδας (τετραγωνικές Ευκλείδειες αποστάσεις) [68].
- **Παραγοντοποίηση Πινάκων :** Η παραγοντοποίηση πινάκων είναι μια άλλη τεχνική για μη εποπτευόμενη μηχανική μάθηση στην EFG. Χρησιμοποιεί «λανθάνοντες παράγοντες» για να χωρίσει έναν μεγάλο πίνακα σε συνδυασμό δύο μικρότερων πινάκων. Οι λανθάνοντες παράγοντες είναι ομοιότητες μεταξύ των στοιχείων. Σκεφτείτε τη φράση: «Πέταξα την μπάλα πάνω από το βουνό». Η λέξη „πέταξα“ είναι πιο πιθανό να συσχετιστεί με τη „μπάλα“ παρά με το „βουνό“ [69].

2.3.3 Νευρωνικά Δίκτυα

Από τη νευρωνική στροφή (αρχές της δεκαετίας του 2010), οι στατιστικές μέθοδοι στην έρευνα της EFG έχουν αντικατασταθεί σε μεγάλο βαθμό από νευρωνικά δίκτυα, καθώς αποδείχθηκαν καλύτερα σε πολλά προβλήματα. Τα νευρωνικά δίκτυα απαιτούν περισσότερα δεδομένα αλλά λιγότερη γλωσσική εμπειρογνωμοσύνη για την εκπαίδευση και τη λειτουργία τους. Είναι λοιπόν προφανές γιατί έγιναν τόσο δημοφιλή από τις αρχές της δεκαετίας του 2010, όταν ο συνεχώς επεκτεινόμενος παγκόσμιος ιστός έλυσε το πρόβλημα της διαθεσιμότητας δεδομένων.

2.3.3.1 Γιατί Νευρωνικά Δίκτυα;

Μειονεκτήματα Στατιστικών Μεθόδων:

- Χειροκίνητη - χρονοβόρα κατασκευή χαρακτηριστικών εισόδου
- Βασίζονται σε μια σειρά από ξεχωριστές ενδιάμεσες εργασίες για την εκμάθηση μιας εργασίας υψηλότερου επιπέδου

Νευρωνικά Δίκτυα - Αυτόματη Μάθηση Χαρακτηριστικών

Οι μέθοδοι βαθιάς μάθησης έχουν τη δυνατότητα να μαθαίνουν αναπαραστάσεις χαρακτηριστικών, ενώ οι μέθοδοι μηχανικής μάθησης απαιτούν από τους ειδικούς να προσδιορίζουν και να εξάγουν με μη αυτόματο τρόπο χαρακτηριστικά από τη φυσική γλώσσα [70].

Μειονεκτήματα χειροποίητων χαρακτηριστικών

- Υπερκαθορισμένα ή ελλιπή
- Απαιτούν πολύ χρόνο για το σχεδιασμό και την αξιολόγησή τους
- Φτάνουν μόνο μέχρι σε ένα συγκεκριμένο όριο επίδοσης

Πλεονεκτήματα αυτόματης μάθησης χαρακτηριστικών

- Βελτιώνονται συνεχώς και με αυτόματο τρόπο
- Προσαρμόζονται εύκολα
- Εκπαιδεύονται/Μαθαίνονται γρήγορα

Νευρωνικά Δίκτυα - Μοντέλα Άκρο-σε-Άκρο

Με τη βαθιά μάθηση έχουμε τη δυνατότητα να αναπτύξουμε ένα νευρωνικό δίκτυο για να λύσουμε όλο το πρόβλημα από άκρο σε άκρο αντί να χρησιμοποιούμε αγωγούς εξειδικευμένων συστημάτων. Αυτό είναι επιθυμητό τόσο για την ταχύτητα όσο και για την απλότητα της ανάπτυξης καθώς και για τη βελτιωμένη απόδοση αυτών των μοντέλων. Για παράδειγμα, τα μεγάλα τμήματα ενός αγωγού αυτόματης αναγνώρισης ομιλίας είναι η επεξεργασία ομιλίας, τα ακουστικά μοντέλα, τα μοντέλα προφοράς και τα μοντέλα γλώσσας. Ωστόσο, οι ιδιότητες και κυρίως τα σφάλματα κάθε υποσυστήματος είναι διαφορετικά. Αυτό παρακινεί την ανάγκη να αναπτυχθεί ένα νευρωνικό δίκτυο για να μάθει ολόκληρο το πρόβλημα από άκρο σε άκρο [70].

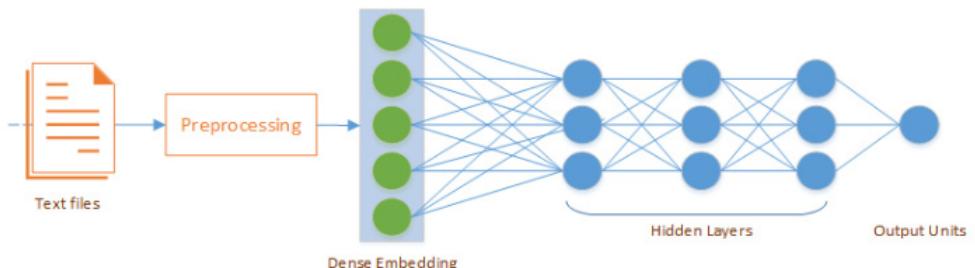
2.3.3.2 Είδη Νευρωνικών Δικτύων για ΕΦΓ

Υπάρχουν πέντε μέθοδοι βαθιάς μάθησης που αξίζουν τη μεγαλύτερη προσοχή για εφαρμογή στην ΕΦΓ.

- Επίπεδα Ενσωμάτωσης
- Πολυεπίπεδα Perceptrons (MLP)
- Συνελικτικά Νευρωνικά Δίκτυα ($\Sigma N \Delta$)
- Επαναλαμβανόμενα Νευρωνικά Δίκτυα (ENΔ)
- Υβριδικά – Συνδυαστικά Νευρωνικά Δίκτυα

Επίπεδο Ενσωμάτωσης

- Η δουλειά του είναι να αντιστοιχίζει τις one-hot κωδικοποιημένες λέξεις (δες 2.4.1.1) στα διανύσματα ενσωμάτωσής τους (δες 2.4.2).
- Χρησιμοποιείται ως επίπεδο εισόδου στις νευρωνικές αρχιτεκτονικές, όπως φαίνεται στο σχήμα 2.5.
- Μπορεί να χρησιμοποιηθεί μόνο του για να μάθει μια ενσωμάτωση λέξης που μπορεί να αποδημευτεί και να χρησιμοποιηθεί σε άλλο μοντέλο αργότερα.
- Μπορεί να χρησιμοποιηθεί ως μέρος ενός μοντέλου βαθιάς μάθησης όπου μαθαίνεται η ενσωμάτωση μαζί με τις παραμέτρους του μοντέλου που εκπαιδεύεται σε μια συγκεκριμένη εργασία ΕΦΓ.
- Μπορεί να χρησιμοποιηθεί για τη φόρτωση προεκπαίδευμένων ενσωματώσεων λέξεων, αρχικοποιώντας τα βάρη του επιπέδου με τις τιμές των προεκπαίδευμένων αυτών διανυσμάτων.



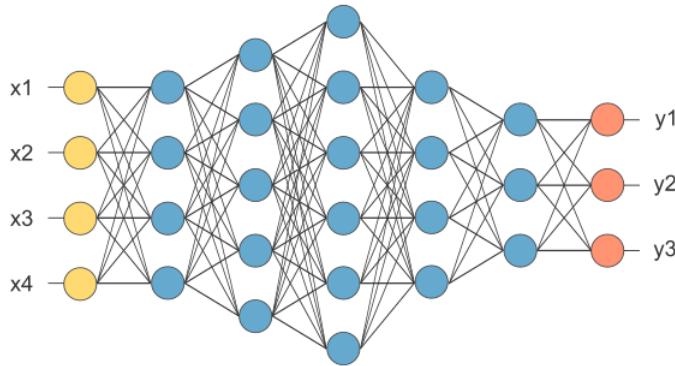
Σχήμα 2.5: Παράδειγμα χρήσης επιπέδου ενσωμάτωσης σε ένα νευρωνικό δίκτυο.

Τα διανύσματα λέξεων (ενσωματώσεις λέξεων) είναι στην πραγματικότητα τα βάρη του επιπέδου ενσωμάτωσης και γί' αυτό λέμε ότι οι ενσωματώσεις μαθαίνονται. Βελτιστοποιούνται η από την εκπαίδευση αυτού του επιπέδου σε ένα συγκεκριμένο πρόβλημα ΕΦΓ. Εφόσον μιλάμε για μοντελοποίηση γλώσσας, το πρόβλημα συνήθως σχετίζεται με την πρόβλεψη λέξεων σε μια ακολουθία. Ωστόσο, για να έχουμε ικανότερα αποτελέσματα σε ένα πρόβλημα ταξινόμησης κειμένου, οι ενσωματώσεις μπορούν να εκπαιδευτούν μαζί με ολόκληρο το μοντέλο σε αυτό το πρόβλημα. Η εκμάθηση μιας ενσωμάτωσης λέξης μπορεί να είναι πολύ χρονοβόρα, καθώς απαιτεί πολλά δεδομένα για να μάθουμε ικαλές, ουσιαστικές αναπαραστάσεις για όλες τις λέξεις [71]. Κατά την υλοποίηση του επιπέδου ενσωμάτωσης (βιβλιοθήκη keras) χρειάζονται τρία βασικά ορίσματα:

- Μέγεθος Εισόδου: Μέγεθος του λεξιλογίου εισόδου – πλήθος μοναδικών λέξεων που εμφανίζονται στο σύνολο δεδομένων εισόδου
- Μέγεθος Εξόδου: Μέγεθος κάθε ενσωμάτωσης λέξης (π.χ. 200-διάστατο διάνυσμα)
- Μήκος Εισόδου: Μήκος των ακολουθιών εισόδου (π.χ. Tweet – 15 λέξεις περίπου)

Πολυεπίπεδο Perceptron (MLP)

Ένα MLP είναι ένα πλήρως συνδεδεμένο τεχνητό νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης. Σε πλήρως συνδεδεμένα δίκτυα κάθε νευρώνας σε ένα επίπεδο συνδέεται με όλους τους νευρώνες στο επόμενο επίπεδο, περνώντας την είσοδο από το ένα επίπεδο στο άλλο με πολλαπλασιασμούς βάρους. Ένα MLP αποτελείται από τουλάχιστον τρία επίπεδα κόμβων: ένα επίπεδο εισόδου, ένα κρυφό επίπεδο και ένα επίπεδο εξόδου. Εκτός από τους κόμβους εισόδου, κάθε κόμβος είναι ένας νευρώνας που χρησιμοποιεί μια μη γραμμική συνάρτηση ενεργοποίησης. Το MLP χρησιμοποιεί μια εποπτευόμενη τεχνική μάθησης που ονομάζεται οπισθοδιάδοση, για την εκπαίδευση των βαρών του δικτύου. Τα πολλαπλά επίπεδα και η μη γραμμική ενεργοποίησή του διακρίνουν το MLP από ένα γραμμικό perceptron. Αυτός είναι ο λόγος για τον οποίο μπορεί να διακρίνει δεδομένα που δεν διαχωρίζονται γραμμικά [72]. Στο σχήμα 2.6 απεικονίζεται ένα MLP με μία είσοδο, μία έξοδο και πέντε κρυφά επίπεδα. Αφού λάβουμε τις ενσωματώσεις λέξεων από ένα μοντέλο γλώσσας ή από ένα επίπεδο ενσωμάτωσης, μπορούμε να τις τροφοδοτήσουμε σε ένα MLP για να εκτελέσει μια συγκεκριμένη εργασία ΕΦΓ, για παράδειγμα Ανάλυση Συναισθήματος (βλέπε 3.1.1).



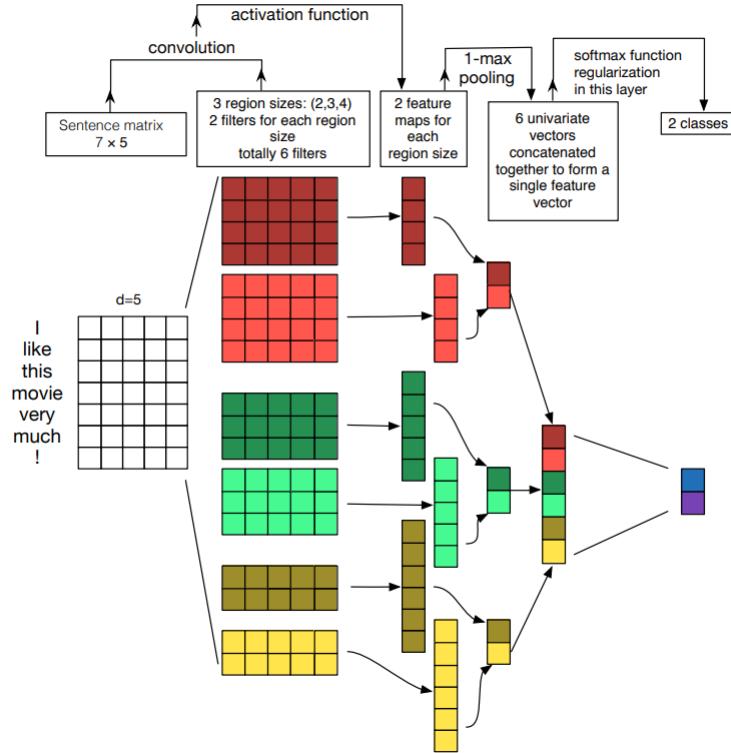
Σχήμα 2.6: Παράδειγμα αρχιτεκτονικής MLP με 5 κρυφά επίπεδα.

Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ)

Το ΣΝΔ είναι μια κατηγορία τεχνητών νευρωνικών δικτύων, που χρησιμοποιείται πιο συχνά για την ανάλυση εικόνων λόγω της ικανότητάς τους να αναγνωρίζουν χωρικά τοπικά μοτίβα εισόδου. Σε ένα ΣΝΔ, τα κρυφά επίπεδα περιλαμβάνουν επίπεδα που εκτελούν συνελιξεις - συνελικτικά επίπεδα. Αυτό το επίπεδο εκτελεί ένα δοτ γινόμενο του πυρήνα/φίλτρου συνέλιξης με τον πίνακα εισόδου του επιπέδου. Καθώς ο πυρήνας συνέλιξης ολισθαίνει κατά μήκος του πίνακα εισόδου για το επίπεδο, η λειτουργία συνέλιξης δημιουργεί έναν χάρτη χαρακτηριστικών, ο οποίος με τη σειρά του συμβάλλει στην είσοδο του επόμενου επιπέδου. Τα ΣΝΔ είναι επίσης γνωστά ως αμετάβλητα κατά την μετατόπιση, χάρη στην αρχιτεκτονική κοινών βαρών των πυρήνων/φίλτρων συνέλιξης που παρέχουν αποκρίσεις σταθερές για διάφορες μεταφράσεις της εισόδου. Αυτό σημαίνει ότι ο χάρτης χαρακτηριστικών που προκύπτει είναι αμετάβλητος κατά τις μετατοπίσεις των θέσεων των χαρακτηριστικών εισόδου στο οπτικό πεδίο. Με άλλα λόγια, μπορούμε να μάθουμε τον ίδιο εξαγωγέα χαρακτηριστικών (τη συνέλιξη) για κάθε θέση σε μια εικόνα, που είναι μια τεράστια εξοικονόμηση παραμέτρων σε σχέση με ένα πλήρως συνδεδεμένο δίκτυο (όχι αμετάβλητο στις μετατοπίσεις της εισόδου), το οποίο θα έπρεπε να ξαναμάθει τον ίδιο εξαγωγέα χαρακτηριστικών για κάθε θέση [73].

Ένα συνελικτικό επίπεδο ακολουθείται συνήθως από ένα επίπεδο max-pooling. Η λειτουργία pooling μειώνει τα μεγέθη των χαρτών χαρακτηριστικών (έξοδος συνελικτικού επιπέδου), επιλέγοντας μια μεμονωμένη τιμή από μια δεξαμενή (ορθογώνια περιοχή) τιμών που ολισθαίνει σαν συνελικτικό φίλτρο στην είσοδο. Εάν η επιλεγμένη τιμή είναι η μέγιστη, τότε η διαδικασία ονομάζεται max-pooling. Τα μαθημένα «φίλτρα» παράγουν την ισχυρότερη απόκριση σε ένα χωρικά τοπικό μοτίβο εισόδου και γι' αυτό το max-pooling είναι η πιο δημοφιλής επιλογή, καθώς βοηθά στην απομόνωση αυτών των τοπικών μοτίβων της εισόδου. Τα επίπεδα pooling παρέχουν επίσης αποκρίσεις που δεν μεταβάλλονται κατά τη μετάφραση. Ωστόσο, τα επίπεδα συνέλιξης ή pooling σε ένα ΣΝΔ που δεν έχουν βήμα μεγαλύτερο από ένα δεν είναι αμετάβλητα κατά τις μεταφράσεις της εισόδου.

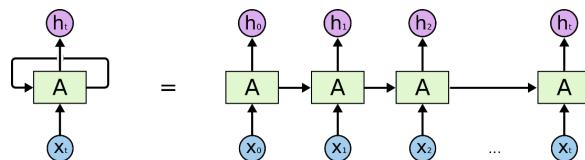
Πρόσφατα τα ΣΝΔ υιοθετήθηκαν επίσης στην ΕΦΓ. Συγκεκριμένα, χρησιμοποιούνται σε προβλήματα ταξινόμησης κειμένου όπου εκμεταλλεύονται όλες τις χρήσιμες ιδιότητές τους που εξηγήθηκαν παραπάνω. Η διαφορά είναι ότι αντί να ολισθαίνουν στον δισδιάστατο χώρο (εικόνα), τώρα οι συνελικτικοί πυρήνες ολισθαίνουν στον μονοδιάστατο χρόνο (λέξεις σε μια ακολουθία). Το σχήμα 2.7 δείχνει ένα παράδειγμα εφαρμογής ενός ΣΝΔ για την ταξινόμηση κριτικών ταινιών.



Σχήμα 2.7: Παράδειγμα ΣΝΔ σε μια εφαρμογή ταξινόμησης κριτικών ταινιών.

Επαναλαμβανόμενα Νευρωνικά Δίκτυα (ΕΝΔ)

Τα ΕΝΔ είναι ένας τύπος τεχνητού νευρωνικού δικτύου που βασίστηκαν στο έργο του Δαιδ Ρυμεληαρτ το 1986 [74]. Είναι δίκτυα με βρόχους μέσα τους, που επιτρέπουν στην πληροφορία να διατηρείται στο χρόνο.



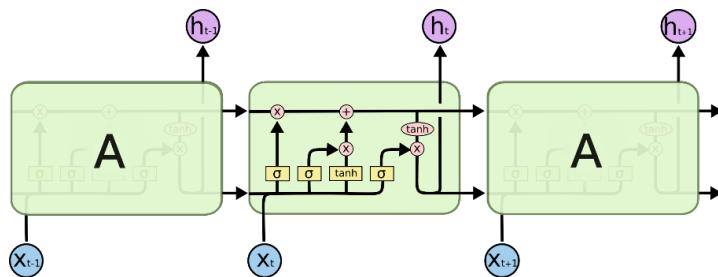
Σχήμα 2.8: Αναπαράσταση του "ξεδιπλωμένου" ΕΝΔ.

Το σχήμα 2.8 δείχνει το βασικό δομικό στοιχείο/επαναλαμβανόμενη μονάδα ενός ΕΝΔ. Ένα κοιμάτι νευρωνικού δικτύου, το A , κοιτάζει κάποια είσοδο x_t και βγάζει μια τιμή h_t . Ένας

βρόχος επιτρέπει τη μετάδοση πληροφοριών από το ένα βήμα του δικτύου στο επόμενο. Το ENΔ είναι στην πραγματικότητα πολλαπλά αντίγραφα του ίδιου δικτύου, καθένα από τα οποία περνά ένα μήνυμα σε έναν διάδοχο. Αυτή η αλυσιδωτή φύση αποκαλύπτει ότι τα ENΔ σχετίζονται στενά με ακολουθίες και λίστες. Είναι η φυσική αρχιτεκτονική του νευρωνικού δικτύου που χρησιμοποιείται για τέτοια δεδομένα. Τα τελευταία χρόνια, υπήρξε απίστευτη επιτυχία στην εφαρμογή των ENΔ σε μια ποικιλία προβλημάτων: Αναγνώριση Ομιλίας, Μοντελοποίηση Γλώσσας, Αυτόματη Μετάφραση, Υπότιτλοι Εικόνων. Ένα από τα ελκυστικά χαρακτηριστικά των ENΔ είναι ότι είναι σε θέση να συνδέουν προηγούμενες πληροφορίες με την παρούσα εργασία. Σε περιπτώσεις όπου το χάσμα μεταξύ των σχετικών πληροφοριών και του τόπου που χρειάζονται είναι μικρό, τα ENΔ μπορούν να μάθουν να χρησιμοποιούν τις προηγούμενες πληροφορίες (βραχυπρόθεσμη μνήμη). Δυστυχώς, καθώς αυτό το χάσμα μεγαλώνει (μακροπρόθεσμη μνήμη), τα απλά ENΔ δεν μπορούν να μάθουν να συνδέουν τις πληροφορίες [75].

Δίκτυα Μακρυπρόθεσμης και Βραχυπρόθεσμης Μνήμης (ΔΜΒΜ)

Τα ΔΜΒΜ είναι ένα ειδικό είδος ENΔ που εισήχθησαν από τους Ηοζηρειτερ & Σζημιδηβερ το 1997 [76]. Τα ΔΜΒΜ έχουν σχεδιαστεί ρητά για την αποφυγή του προβλήματος της μακροπρόθεσμης εξάρτησης. Όλα τα επαναλαμβανόμενα νευρωνικά δίκτυα έχουν τη μορφή μιας αλυσίδας επαναλαμβανόμενων μονάδων νευρωνικού δικτύου. Στα τυπικά ENΔ, αυτή η επαναλαμβανόμενη ενότητα θα έχει μια πολύ απλή δομή, όπως ένα επίπεδο τανη. Τα ΔΜΒΜ έχουν επίσης αυτή τη δομή σαν αλυσίδα, αλλά η επαναλαμβανόμενη μονάδα έχει διαφορετική δομή. Αντί να έχουμε ένα μόνο επίπεδο νευρωνικού δικτύου, υπάρχουν τέσσερα, που αλληλεπιδρούν μεταξύ τους με έναν πολύ ιδιαίτερο τρόπο που φαίνεται στο σχήμα 2.9.

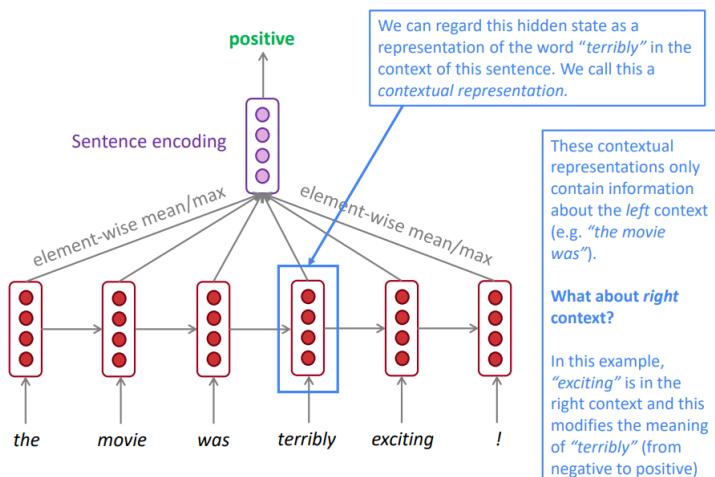


Σχήμα 2.9: Η δομή μιας μονάδας ΔΜΒΜ.

Το κλειδί για τα ΔΜΒΜ είναι η κατάσταση της κυψέλης, η οριζόντια γραμμή που διατρέχει την κορυφή του διαγράμματος. Τρέχει κατευθείαν σε ολόκληρη την αλυσίδα, αφήνοντας τις πληροφορίες να ρέουν κατά μήκος της. Το ΔΜΒΜ έχει την ικανότητα να αφαιρεί ή να προσθέτει πληροφορίες στην κατάσταση της κυψέλης, που ρυθμίζεται προσεκτικά από δομές που ονομάζονται πύλες. Οι πύλες είναι ένας τρόπος για να περάσουν προαιρετικά πληροφορίες. Αποτελούνται από ένα νευρωνικό επίπεδο σιγμοειδούς ενεργοποίησης και μια

πράξη πολλαπλασιασμού κατά σημείο. Ένα ΔMBM έχει τρεις από αυτές τις πύλες, για την προστασία και τον έλεγχο της κατάστασης της κυψέλης [77].

Θεωρητικά, τα κλασικά ENΔ μπορούν να παρακολουθούν αυθαίρετες μακροπρόθεσμες εξαρτήσεις στις ακολουθίες εισόδου. Το πρόβλημα με αυτά τα μοντέλα είναι υπολογιστικής (ή πρακτικής) φύσης: όταν εκπαιδεύεται ένα τυπικό ENΔ χρησιμοποιώντας οπισθοδιάδοση, οι μακροπρόθεσμες κλίσεις που διαδίδονται προς τα πίσω μπορούν να «εξαφανιστούν» (δηλαδή, μπορεί να τείνουν στο μηδέν) ή εκρήγνυνται» (δηλαδή, μπορούν να τείνουν στο άπειρο), λόγω των υπολογισμών που εμπλέκονται στη διαδικασία, οι οποίοι χρησιμοποιούν αριθμούς πεπερασμένης ακρίβειας. Τα ENΔ που χρησιμοποιούν μονάδες ΔMBM λύνουν εν μέρει το πρόβλημα της εξαφάνισης κλίσης, επειδή οι μονάδες ΔMBM επιτρέπουν στις κλίσεις να ρέουν επίσης αμετάβλητες. Ωστόσο, τα δίκτυα ΔMBM μπορεί ακόμα να υποφέρουν από το πρόβλημα της έκρηξης κλίσης. Το Σχήμα 2.10 απεικονίζει ένα παράδειγμα εφαρμογής του ΔMBM στην ταξινόμηση/ανάλυση συναισθημάτων κριτικής ταινίας. Θα δούμε περισσότερα για την ανάλυση συναισθήματος στο κεφάλαιο 3.



Σχήμα 2.10: Παράδειγμα ΔMBM για ταξινόμηση κριτικών ταινιών.

ΕΝΔ vs ΣΝΔ

ΕΝΔ

- Αναγνωρίζουν πρότυπα στο χρόνο.
- Αντιλαμβάνονται έννοιες μεγάλων φράσεων.
- Θυμούνται τη σειρά εμφάνισης των λέξεων στο κείμενο.
- Κατάλληλα για περιπτώσεις ακολουθιακών δεδομένων: Απάντηση Ερωτήσεων, Αυτόματη Μετάφραση, Αναγνώριση Μέρους του Λόγου, Περιγραφή Εικόνας, Μοντελοποίηση Γλώσσας κλπ.

- Αργή Εκπαίδευση – Δεν μπορούν να εκμεταλλευτούν την δυνατότητα παραλληλοποίησης που παρέχει η ΓΠΥ λόγω της σειριακής φύσης λειτουργίας τους.

ΣΝΔ

- Αναγνωρίζουν πρότυπα στο χώρο.
- Εντοπίζουν “γειτονιές” χαρακτηριστικών.
- Δεν λαμβάνουν υπόψη τη σειρά εμφάνισης των λέξεων στο κείμενο.
- Κατάλληλα μόνο για προβλήματα ταξινόμησης κειμένου (πέραν της υπολογιστικής όρασης): Ανάλυση Συναισθήματος, Ανίχνευση Ανεπιθύμητων Μηνυμάτων, Κατηγοριοποίηση Θέματος.
- Γρήγορη Εκπαίδευση (5ξ) – Αξιοποίηση της ΓΠΥ για τον παράλληλο υπολογισμό συνελίξεων.

Υβριδικά Νευρωνικά Δίκτυα

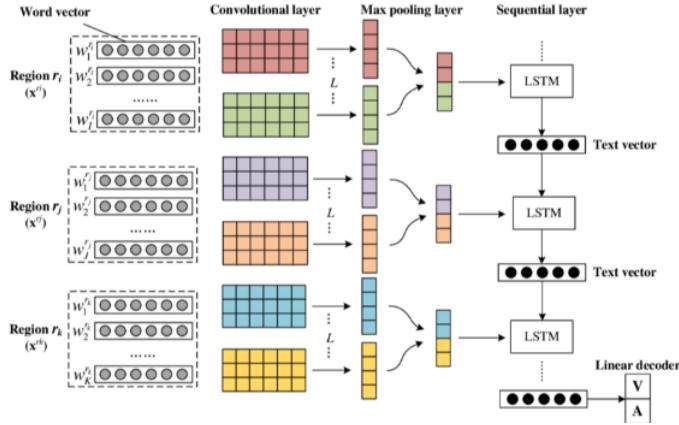
Το υβριδικό νευρωνικό δίκτυο είναι ένα δίκτυο με συνδυασμό διαφορετικών ειδών τεχνητών νευρωνικών δικτύων και προσεγγίσεων. Χρησιμοποιεί χαρακτηριστικά ποικίλων νευρωνικών δικτύων και προσεγγίσεων για την επίτευξη βέλτιστων αποτελεσμάτων. Ο πιο δημοφιλής συνδυασμός που βρέθηκε στη βιβλιογραφία είναι αυτός ενός ΣΝΔ και ενός ΔΜΒΜ. Τέτοια μοντέλα αναπτύχθηκαν αρχικά για προβλήματα πρόβλεψης οπτικών χρονοσειρών και για την εφαρμογή δημιουργίας περιγραφών κειμένου μεμονωμένης εικόνας ή ακολουθίας εικόνων (βίντεο). Ωστόσο, αυτή η αρχιτεκτονική έχει επίσης χρησιμοποιηθεί σε προβλήματα αναγνώρισης ομιλίας και ΕΦΓ, όπου τα ΣΝΔ χρησιμοποιούνται ως εξαγωγείς χαρακτηριστικών σε δεδομένα εισόδου ήχου και κειμένου τα οποία τροφοδοτούν στα ΔΜΒΜ [78]. Αυτός ο συνδυασμός ΣΝΔ και ΔΜΒΜ είναι κατάλληλος για προβλήματα που:

- Τα χαρακτηριστικά της εισόδου κατανέμονται χωρικά (χωρική δομή), όπως τα εικονοστοιχεία σε μια εικόνα (2Δ) ή η δομή των λέξεων σε μια πρόταση (1Δ), παράγραφο ή έγγραφο.
- Τα χαρακτηριστικά της εισόδου κατανέμονται στο χρόνο (χρονική δομή), όπως η σειρά των εικόνων σε ένα βίντεο ή των λέξεων σε κείμενο, ή απαιτούν τη δημιουργία εξόδου με χρονική δομή, όπως λέξεις σε μια περιγραφή κειμένου.

Μπορούν να συνδυαστούν με τρεις διαφορετικούς τρόπους όπως φαίνεται παρακάτω.

ΣΝΔ + ΔMBM

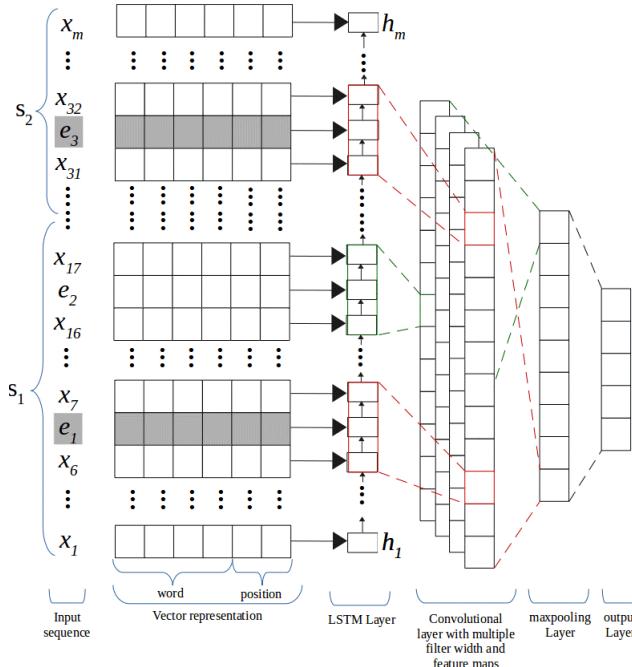
Μια σειριακή αρχιτεκτονική όπου οι ενσωματώσεις λέξεων εισόδου τροφοδοτούνται πρώτα στο τμήμα ΣΝΔ και στη συνέχεια τα εξαγόμενα χαρακτηριστικά περνούν μέσω του ΔMBM. Ένα τελικό επίπεδο εξόδου χρησιμοποιείται συνήθως μετά το επίπεδο ΔMBM. Ένα παράδειγμα αυτής της αρχιτεκτονικής απεικονίζεται στο σχήμα 2.11.



Σχήμα 2.11: Παράδειγμα αρχιτεκτονικής ΣΝΔ + ΔMBM.

ΔMBM + ΣΝΔ

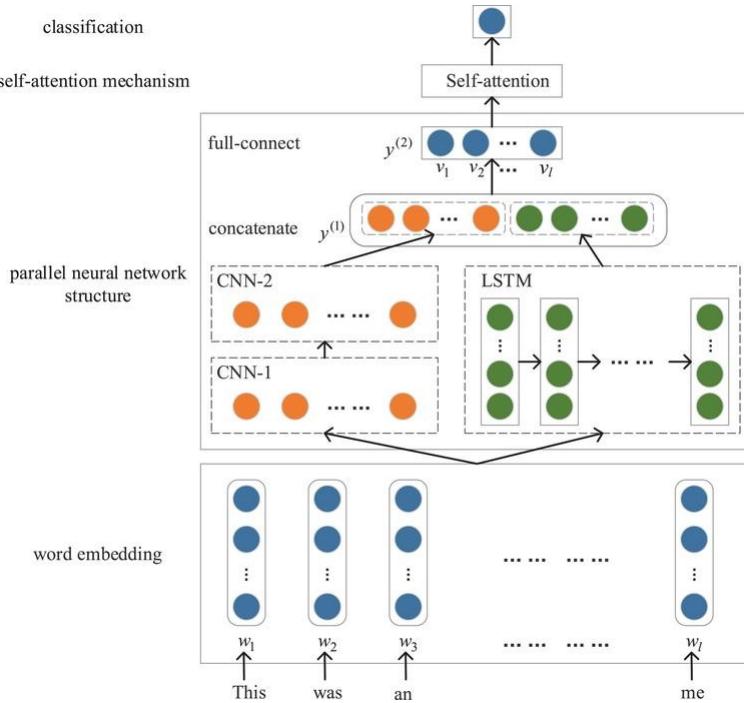
Μια σειριακή αρχιτεκτονική όπου το επίπεδο ΔMBM τοποθετείται πριν από το τμήμα ΣΝΔ. Σε αυτήν την περίπτωση, το ΣΝΔ εξάγει χαρακτηριστικά από τις εξόδους του ΔMBM και τα τροφοδοτεί στο επίπεδο εξόδου. Αυτός ο τύπος μοντέλου δεν χρησιμοποιείται συνήθως καθώς είναι λιγότερο αποτελεσματικός στην εκπαίδευση σε σύγκριση με την αρχιτεκτονική ΣΝΔ + ΔMBM. Αυτό συμβαίνει επειδή το ΣΝΔ χρησιμοποιείται για να μειώσει τη διάσταση της εισόδου με την παραγωγή των χαρτών χαρακτηριστικών. Επομένως, το ΔMBM θα πρέπει να επεξεργαστεί μικρότερες εισόδους σε σύγκριση με το ΔMBM + ΣΝΔ όπου το ΔMBM επεξεργάζεται τις πλήρεις αναπαραστάσεις εισόδου καθιστώντας το πιο αργό στην εκπαίδευση. Ένα παράδειγμα αυτής της αρχιτεκτονικής απεικονίζεται στο σχήμα 2.12.



Σχήμα 2.12: Παράδειγμα αρχιτεκτονικής ΔΜΒΜ + ΣΝΔ.

ΣΝΔ // ΔΜΒΜ

Ένας παράλληλος συνδυασμός αυτών των δύο τύπων νευρωνικών δικτύων. Οι ενσωματώσεις εισόδου τροφοδοτούνται και στο ΣΝΔ και στο ΔΜΒΜ και οι αντίστοιχες έξοδοι τους συνήθως συνδέονται. Στη συνέχεια, τα συνενωμένα χαρακτηριστικά τροφοδοτούνται στα ακόλουθα επίπεδα (εξόδου). Ένα παράδειγμα αυτής της αρχιτεκτονικής απεικονίζεται στο σχήμα 2.13.



Σχήμα 2.13: Παράδειγμα αρχιτεκτονικής ΣΝΔ // ΔΜΒΜ.

2.4 Αναπαραστάσεις Λέξεων

Για να κάνουμε τα δεδομένα κειμένου επεξεργάσιμα από οποιοδήποτε είδος αλγόριθμου, πρέπει να τα μετατρέψουμε σε αριθμούς. Αυτές οι αριθμητικές αναπαραστάσεις είναι σε διανυσματική μορφή και συνήθως αντιπροσωπεύουν λέξεις. Υπάρχουν δύο κατηγορίες αναπαραστάσεων/διανυσμάτων λέξεων.

- **Σταθερές Αναπαραστάσεις (Αραιές)** : Κάθε διάσταση/τιμή του διανύσματος χαρακτηριστικών αντιστοιχεί σε μια συγκεκριμένη λέξη από το λεξιλόγιο εισόδου.
- **Κατανεμημένες Αναπαραστάσεις (Πυκνές)** : Τα χαρακτηριστικά/διαστάσεις του διανύσματος δεν αντιστοιχούν σε λέξεις από το λεξιλόγιο αλλά σε κάποιο νόημα/οντότητα.

2.4.1 Σταθερές Αναπαραστάσεις

Αυτό το είδος αναπαραστάσεων ονομάζονται επίσης αραιές επειδή είναι συνήθως πολύ μεγάλες και περιέχουν λίγες πληροφορίες για τη σημασιολογία των λέξεων. Τρία είναι τα κύρια παραδείγματα σταθερών αναπαραστάσεων, που παρατίθενται παρακάτω.

- One-Hot Κωδικοποίηση

- Σύνολα Λέξεων
- ΣΟ-ΑΣΕ

2.4.1.1 One-hot Κωδικοποίηση

Δημιουργείται ένα λεξιλόγιο που περιέχει όλες τις μοναδικές λέξεις που βρίσκονται στο κείμενο εισόδου. Η κωδικοποίηση one-hot παράγει μια διανυσματική αναπαράσταση για κάθε λέξη όπου όλες οι τιμές είναι μηδέν εκτός από την αντίστοιχη θέση της λέξης στο λεξιλόγιο. Επομένως, τα διανύσματα λέξεων έχουν το ίδιο μέγεθος με το λεξιλόγιο [79]. Ένα παράδειγμα φαίνεται στο σχήμα 2.14.

Word	ID
Red	1
Green	2
Blue	3
Yellow	4
Orange	5
Purple	6

→

	Green	Yellow	Blue	Red	Purple	Orange
Green	0	1	0	0	0	0
Yellow	0	0	0	1	0	0
Orange	0	0	0	0	1	0

Σχήμα 2.14: Παράδειγμα One-hot κωδικοποίησης.

2.4.1.2 Σύνολα Λέξεων (ΣΛ)

Τα ΣΛ είναι μια μέθοδος που χρησιμοποιείται για την αναπαράσταση φράσεων, προτάσεων, παραγράφων, εγγράφων και όχι μεμονωμένων λέξεων. Τα διανύσματα που παράγονται έχουν το μέγεθος του λεξιλογίου, αλλά σε αντίθεση με την κωδικοποίηση one-hot περισσότερες από μία στήλες είναι μη μηδενικές, καθώς οι αναπαραστάσεις αντιπροσωπεύουν περισσότερες από μία λέξη [80]. Όταν οι μη μηδενικές τιμές είναι δυαδικές υποδεικνύοντας απλώς την ύπαρξη λέξεων στο αντιπροσωπευόμενο πλαίσιο, τα ονομάζουμε δυαδικά ΣΛ. Στα κλασικά ΣΛ, χρησιμοποιούνται οι συχνότητες εμφάνισης των λέξεων. Στο σχήμα 2.15 βλέπουμε τα αντίστοιχα παραδείγματα για κάθε περίπτωση.

BOW Representation

Representing the sentence, "it is the best of the best "

It is the best of a an

[1 ,1 ,1 ,1 ,1 ,0 ,0]

(only the words present in the document are activated)

(or)

[1 ,1 ,2 ,2 ,1 ,0 ,0]

(the word count is taken into consideration instead of activation)

Σχήμα 2.15: Παράδειγμα καδικοποίησης με ΣΛ.

2.4.1.3 Συχνότητα Όρου - Ανάστροφη Συχνότητα Εγγράφου (ΣΟ-ΑΣΕ)

Αυτή η μέθοδος υπολογίζει ένα στατιστικό μέτρο για να αξιολογήσει πόσο σημαντική είναι μια λέξη για ένα έγγραφο σε μια συλλογή ή ένα σώμα. Η σημασία αυξάνεται αναλογικά με τον αριθμό των φορών που εμφανίζεται μια λέξη στο έγγραφο, αλλά αντισταθμίζεται από τη συχνότητα της λέξης στο κείμενο. Συνήθως, το βάρος ΣΟ-ΑΣΕ αποτελείται από δύο όρους [81] :

- **Συχνότητα Όρου (ΣΟ)**, που μετρά πόσο συχνά εμφανίζεται ένας όρος σε ένα έγγραφο. Δεδομένου ότι κάθε έγγραφο έχει διαφορετικό μήκος, είναι πιθανό ένας όρος να εμφανίζεται πολύ περισσότερες φορές σε μεγάλα έγγραφα παρά σε μικρότερα. Έτσι, η συχνότητα του όρου συχνά διαιρείται με το μήκος του εγγράφου (γνωστός και ως ο συνολικός αριθμός όρων στο έγγραφο) ως τρόπος κανονικοποίησης:

$$\text{ΣΟ}(\tau) = (\text{Συχνότητα εμφάνισης του όρου } \tau \text{ σε ένα έγγραφο}) / (\text{Συνολικό πλήθος όρων στο έγγραφο}).$$

- **Ανάστροφη Συχνότητα Εγγράφου (ΑΣΕ)**, ο οποίος μετρά πόσο σημαντικός είναι ένας όρος. Κατά τον υπολογισμό του ΣΟ, όλοι οι όροι θεωρούνται εξίσου σημαντικοί. Ωστόσο, είναι γνωστό ότι ορισμένοι όροι, όπως "είναι", "του" και "αυτό", μπορεί να εμφανίζονται πολλές φορές αλλά έχουν μικρή σημασία. Επομένως, πρέπει να σταθμίσουμε τους συχνούς όρους ενώ κλιμακώνουμε τους σπάνιους, υπολογίζοντας τα ακόλουθα:

$$\text{ΑΣΕ}(\tau) = \lambda_{\text{ογ_ε}}(\text{Συνολικό πλήθος εγγράφων στο σύνολο δεδομένων} / \text{Αριθμός εγγράφων που περιέχουν τον όρο } \tau).$$

Το σχήμα 2.16 δείχνει ένα παράδειγμα όπου οι προτάσεις A, B αναπαρίστανται χρησιμοποιώντας τη μέθοδο ΣΟ-ΑΣΕ. Η στήλη Word είναι το λεξιλόγιο που περιέχει όλες τις μοναδικές λέξεις στο κείμενο.

Word	TF		IDF	TF*IDF	
	A	B		A	B
The	1/7	1/7	$\log(2/2) = 0$	0	0
Car	1/7	0	$\log(2/1) = 0.3$	0.043	0
Truck	0	1/7	$\log(2/1) = 0.3$	0	0.043
Is	1/7	1/7	$\log(2/2) = 0$	0	0
Driven	1/7	1/7	$\log(2/2) = 0$	0	0
On	1/7	1/7	$\log(2/2) = 0$	0	0
The	1/7	1/7	$\log(2/2) = 0$	0	0
Road	1/7	0	$\log(2/1) = 0.3$	0.043	0
Highway	0	1/7	$\log(2/1) = 0.3$	0	0.043

Σχήμα 2.16: Παράδειγμα κωδικοποίησης ΣΟ-ΑΣΕ.

2.4.1.4 Σταθερές Αναπαραστάσεις - Πλεονεκτήματα & Μειονεκτήματα

Μειονεκτήματα

- Απαιτούν μεγάλη μνήμη
- Δεν ενσωματώνουν τη σημασία/νόημα της λέξης

Πλεονεκτήματα

- Εύκολες στη χρήση

2.4.2 Κατανεμημένες Αναπαραστάσεις (Ενσωματώσεις)

Σε αντίθεση με τις σταθερές αναπαραστάσεις που είναι αραιές, καθώς έχουν το μέγεθος του λεξιλογίου, οι κατανεμημένες αναπαραστάσεις είναι πυκνές. Αυτό σημαίνει ότι αποτελούνται από έναν προκαθορισμένο αριθμό οντοτήτων/διαστάσεων που η καθεμία αποτυπώνει μια συγκεκριμένη σημασιολογική ιδιότητα. Επομένως, μια λέξη αντιπροσωπεύεται ως κατανομή σε αυτές τις οντότητες. Το πιο σημαντικό είναι ότι οι λέξεις με παρόμοια σημασία έχουν παρόμοιες αναπαραστάσεις/διανύσματα καθώς η σημασιολογία των λέξεων είναι ενσωματωμένη στο διάνυσμα (ενσωματώσεις λέξεων). Τέτοιες αναπαραστάσεις μαθαίνονται όπως τα βάρη ενός νευρωνικού στην εκπαίδευση. Τα μοντέλα που χρησιμοποιούνται για την εκπαίδευση/εκμάθηση ουσιαστικών ενσωματώσεων λέξεων ονομάζονται μοντέλα γλώσσας. Εκπαίδευονται με μεγάλα κείμενα, σε ένα πρόβλημα πρόβλεψης λέξεων. Υπάρχουν δύο κατηγορίες ενσωματώσεων λέξεων που αναφέρονται παρακάτω.

- *Κλασικές Ενσωματώσεις Λέξεων*
 - Word2Vec (T. Mikolov, 2013 [48])
 - GloVe (J. Pennington, 2014 [82])
 - FastText (P. Bojanowski, 2016 [83])
- *Πλαισιωμένες Ενσωματώσεις Λέξεων*
 - CoVe (B. McCann, 2017 [57])
 - ELMo (M. Peters, 2018 [58])
 - GPT (OpenAI, 2018 [59])
 - BERT (J. Devlin, 2018 [60])

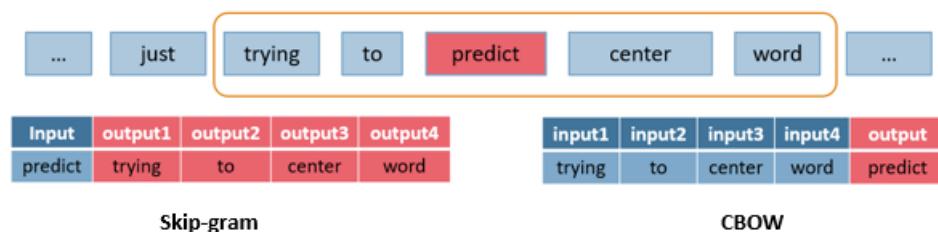
2.4.2.1 Κλασικές Ενσωματώσεις Λέξεων

Με τις κλασικές ενσωματώσεις λέξεων, κάθε λέξη έχει μια στάνταρ διανυσματική αναπαράσταση πραγματικών τιμών που μαθαίνεται από ένα γλωσσικό μοντέλο.

Word2Vec

To Word2Vec είναι ένα νευρωνικό δίκτυο δύο επιπέδων που προτάθηκε από τον T. Mikolov το 2013 [48]. Αυτή η εργασία σηματοδότησε την αρχή μιας νέας εποχής στην ΕΦΓ καθώς οι προτεινόμενες ενσωματώσεις επέτρεψαν την αποτελεσματική εκπαίδευση νευρωνικών δικτύων που δεν ήταν σε θέση να αποδώσουν καλά με τις προηγούμενες αραιές αναπαραστάσεις. To Word2Vec εκπαιδεύτηκε για την ανακατασκευή γλωσσικών πλαισίων λέξεων. Συγκεκριμένα, η εκπαίδευση περιελάμβανε ζεύγη λέξεων περιβάλλοντος-στόχου δίνοντάς μας δύο παραλλαγές της μεθόδου (σχήμα 2.17).

- *Skip-gram* : Οι ενσωματώσεις μαθαίνονται ενώ το μοντέλο προσπαθεί να προβλέψει τα συμφραζόμενα με δεδομένη τη λέξη-στόχο ως είσοδο.
- *CBOW* : Οι ενσωματώσεις μαθαίνονται ενώ το μοντέλο προσπαθεί να προβλέψει τη λέξη-στόχο δεδομένων των συμφραζόμενών της.



Σχήμα 2.17: Skip-gram vs CBOW.

Διαδικασία εκπαίδευσης Skip-gram

Σε αυτήν την ενότητα θα εξηγήσουμε την έκδοση skip-gram (πιο δημοφιλής) του Word2Vec και θα δούμε πώς εκπαιδεύεται [84]. Στην έκδοση Skip-gram, μια λέξη-στόχος δίνεται ως είσοδος και το μοντέλο προβλέψει τις γύρω λέξεις. Από τη φύση του, αυτό είναι ένα αρκετά δύσκολο έργο και έγινε πολύ πιο εύκολο με την ακόλουθη ιδέα. Αντί να ζητάμε από το μοντέλο να προβλέψει τη γειτονική λέξη, το ρωτάμε αν μια λέξη είναι γείτονας. Αυτό κάνει την επεξεργασία πολύ πιο γρήγορη, επιτρέπει την εκπαίδευση σε τεράστια δεδομένα και κατά συνέπεια μαθαίνει καλύτερες ενσωματώσεις. Το σχήμα 2.18 απεικονίζει την εκπαιδευτική διαδικασία.



Σχήμα 2.18: Η βασική ιδέα της προσέγγισης Skip-gram.

Αρνητική Δειγματοληψία

Για να εξισορροπηθεί το σύνολο δεδομένων εκπαίδευσης, χρησιμοποιείται αρνητική δειγματοληψία. Για κάθε δείγμα στο αρχικό σύνολο δεδομένων, τυχαίες λέξεις από το λεξιλόγιο - όχι γείτονες - τίθενται ως αρνητικά παραδείγματα. Με αυτόν τον τρόπο αποφεύγουμε να αποκτήσουμε ένα «τεμπέλικο» μοντέλο που δίνει κακές ενσωματώσεις. Ένα παράδειγμα αρνητικής δειγματοληψίας φαίνεται στο σχήμα 2.19.

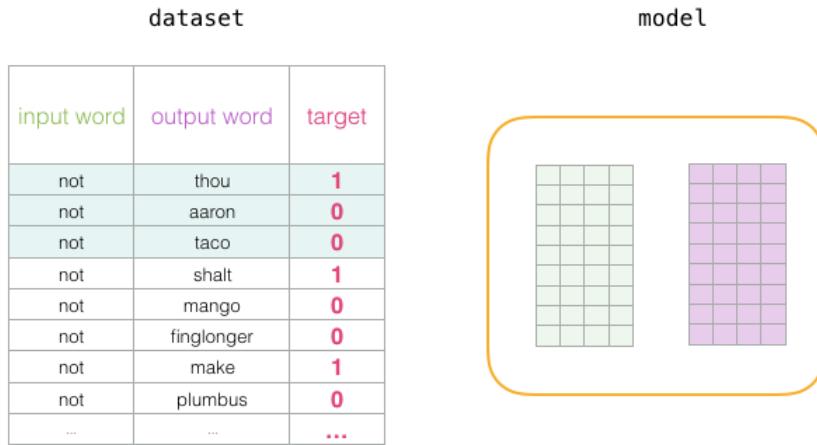
Skipgram					Negative Sampling		
shalt	not	make	a	machine	input word	output word	target
					make	shalt	1
					make	aaron	0
					make	taco	0

Σχήμα 2.19: Παράδειγμα αρνητικής δειγματοληψίας.

Διαδικασία Εκπαίδευσης

Αρχικά, δημιουργούνται πίνακες ενσωματώσεων λέξεων-στόχων (πράσινο) και λέξεων-περιβάλλοντος (ροζ) (και οι δύο μεγέθους λεξιλογίου ξ μεγέθους ενσωμάτωσης) όπως φαίνεται στο σχήμα 2.20. Αυτά τα δύο έχουν μια ενσωμάτωση για κάθε λέξη και ενημερώνονται κατά

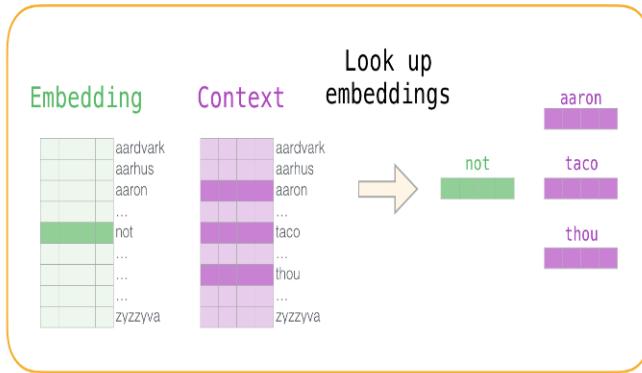
τη διάρκεια της εκπαίδευσης. Ο στόχος της εκπαίδευσης είναι να αποκτήσουμε έναν καλό πίνακα ενσωματώσεων με όλες τις αναπαραστάσεις λέξεων. Στη συνέχεια, μπορούμε να φορτώσουμε αυτόν τον προ-εκπαιδευμένο πίνακα για να αντιμετωπίσουμε οποιαδήποτε εργασία ΕΦΓ.



Σχήμα 2.20: Τα δεδομένα εκπαίδευσης του μοντέλου Skip-gram.

Η διαδικασία εκπαίδευσης έχει ως εξής:

1. Πάρε ένα θετικό παράδειγμα και τα σχετικά αρνητικά παραδείγματα (σχήμα 2.20).
2. Αντιστοίχισε τις λέξεις στις ενσωματώσεις τους με τη βοήθεια 2 πινάκων: Λέξεις Εισόδου -> Πίνακας Ενσωμάτωσης, Λέξεις Εξόδου/Συμφραζόμενα -> Πίνακας Συμφραζόμενων (σχήμα 2.21).
3. Βρες το δοτ-γινόμενο των αντίστοιχων ενσωματώσεων - αυτός ο αριθμός υποδεικνύει την ομοιότητα των ενσωματώσεων εισόδου και συμφραζόμενων (σχήμα 2.22).
4. Χρησιμοποίησε σιγμοειδή συνάρτηση (κατάλληλη για δυαδική ταξινόμηση) για να πάρεις την έξοδο του μοντέλου.
5. Υπολόγισε το σφάλμα πρόβλεψης και χρησιμοποίησε το για να ενημερώσεις τις παραμέτρους του μοντέλου (δηλαδή τους δύο πίνακες).
6. Επανάλαβε για ολόκληρο το σύνολο δεδομένων για να αποκτήσουμε έναν καλά εκπαιδευμένο πίνακα ενσωμάτωσης.



Σχήμα 2.21: Οι πίνακες ενσωματώσεων του μοντέλου Skip-gram.



Σχήμα 2.22: Η διαδικασία εκπαίδευσης του μοντέλου Skip-gram.

Καθολικά Διανύσματα (GloVe)

Το GloVe είναι ένα μοντέλο log-διγραμμικής παλινδρόμησης για μη εποπτευόμενη εκμάθηση αναπαραστάσεων λέξεων που προτάθηκε από το J. Pennington το 2014 [82]. Συνδυάζει τα χαρακτηριστικά δύο οικογενειών μοντέλων, δηλαδή την καθολική παραγοντοποίηση πίνακα και τις μεθόδους παραθύρου τοπικού περιβάλλοντος (Word2Vec). Το GloVe αξιοποιεί αποτελεσματικά τις στατιστικές πληροφορίες εκπαιδεύοντας μόνο στα μη μηδενικά στοιχεία σε έναν πίνακα συν-εμφάνισης λέξεων, αντί σε ολόκληρο τον αραιό πίνακα ή σε μεμονωμένα παράθυρα περιβάλλοντος σε ένα μεγάλο κείμενο. Ένας πίνακας συν-εμφάνισης όρων κατασκευάζεται από ολόκληρο το κείμενο. Συγκεκριμένα, για κάθε λέξη (π.χ. νερό), υπολογίζεται $\Pi(\kappa|water) = \text{πιθανότητα του } \kappa \text{ και του νερού να συνυπάρχουν}$, όπου $\kappa = \text{λέξη από το λεξιλόγιο}$. Αυτός ο πίνακας περιβάλλοντος υψηλών διαστάσεων στη συνέχεια μειώνεται με την κανονικοποίηση των μετρήσεων και την λογαριθμική εξομάλυνση.

FastText

Τα δημοφιλή μοντέλα που μαθαίνουν αναπαραστάσεις λέξεων συνήθως αγνοοούν τη μορφολογία των λέξεων, αναθέτοντας ένα ξεχωριστό διάνυσμα σε κάθε λέξη. Αυτός είναι ένας περιορισμός, ειδικά για γλώσσες με μεγάλο λεξιλόγιο και πολλές σπάνιες λέξεις. Το FastText, που προτάθηκε από τον P. Bojanowski το 2016 [83] από την ερευνητική ομάδα AI του Facebook, έλυσε αυτό το πρόβλημα. Η μέθοδος βασίζεται στο μοντέλο skipgram του Word2Vec. Η καινοτομία εδώ είναι ότι λαμβάνονται υπόψη οι μονάδες υπο-λέξεων και οι λέξεις αντιπροσωπεύονται από το άθροισμα των διανυσματικών αναπαραστάσεων των υποσυνόλων ν-χαρακτήρων της λέξης συν την ίδια τη λέξη. Όπως υποδηλώνει το όνομα, είναι μια γρήγορη μέθοδος που επιτρέπει την γρήγορη εκπαίδευση μοντέλων σε μεγάλα κείμενα και γι' αυτό το Facebook διαθέτει προεκπαιδευμένα μοντέλα για 294 γλώσσες. Το FastText μας δίνει επίσης τη δυνατότητα να υπολογίζουμε αναπαραστάσεις λέξεων για λέξεις που δεν εμφανίζονται στα δεδομένα εκπαίδευσης.

Εδώ βλέπουμε ένα παράδειγμα αναπαράστασης της λέξης 'where' χρησιμοποιώντας 3άδες χαρακτήρων:

where , n=3 -> <wh, whe, her, ere, re>, <where>

Η διανυσματική αναπαράσταση για τη λέξη 'where' θα προέλθει από το άθροισμα των αντίστοιχων διανυσματικών αναπαραστάσεων από τις οντότητες που φαίνονται στη δεξιά πλευρά. Είναι πολύ πιθανό μερικά ν-σύνολα να είναι κοινά μεταξύ λέξεων με παρόμοια μορφολογία και επομένως αυτές οι λέξεις θα έχουν παρόμοιες αναπαραστάσεις. Η κοινή χρήση των αναπαραστάσεων μεταξύ των λέξεων επιτρέπει την εκμάθηση αξιόπιστων αναπαραστάσεων για σπάνιες λέξεις, ενώ τα προηγούμενα μοντέλα είχαν κακή απόδοση από αυτήν την άποψη.

Πρόβλημα με τις Κλασικές Ενσωματώσεις Λέξεων

Ένα πρόβλημα που πρέπει να αντιμετωπίσουμε όταν χρησιμοποιούμε κλασικές ενσωματώσεις λέξεων είναι η πολυσημία. Η πολυσημία συναντάται όταν η ίδια λέξη έχει διαφορετική σημασία με βάση τα συμφραζόμενά της. Η λύση σε αυτό το πρόβλημα φέρεται από τις πλαισιωμένες ενσωματώσεις. Αυτές οι ενσωματώσεις είναι παρόμοιες με τις κλασικές, αλλά λαμβάνουν επιπλέον υπόψη το πλαίσιο/συμφραζόμενα της λέξης.

2.4.2.2 Πλαισιωμένες Ενσωματώσεις Λέξεων

Σε αντίθεση με τις κλασικές ενσωματώσεις λέξεων, οι πλαισιωμένες δεν είναι ίδιες για κάθε λέξη. Όπως υποδηλώνει το όνομα, αλλάζουν ανάλογα με το πλαίσιο στο οποίο βρίσκεται η λέξη. Αυτό σημαίνει ότι η ίδια λέξη μπορεί να έχει περισσότερες από μία αναπαραστάσεις. Για παράδειγμα, η λέξη «match» μπορεί να βρεθεί με λέξεις «football» ή «fire» όπου έχει εντελώς διαφορετική σημασία σε κάθε περίπτωση. Επομένως, οι πλαισιωμένες ενσωματώσεις

λέξεων είναι πολύ χρήσιμες και οδηγούν σε καλύτερα αποτελέσματα σε πολλές εργασίες ΕΦΓ. Θα συζητήσουμε για τέσσερα δημοφιλή γλωσσικά μοντέλα που μαθαίνουν πλαισιωμένες ενσωματώσεις λέξεων.

Διανύσματα Περιβάλλοντος (CoVe)

Το CoVe, που προτάθηκε από τον B. McCann το 2017 [57] ήταν η πρώτη προσέγγιση που προσπάθησε να λύσει το πρόβλημα της πολυσημίας λαμβάνοντας υπόψη το πλαίσιο μιας λέξης κατά την εκμάθηση διανυσματικών αναπαραστάσεων. Ένα μοντέλο κωδικοποιητή-αποκωδικοποιητή εκπαιδεύεται σε μια εργασία αυτόματης μετάφρασης (εποπτευόμενη μάθηση) για να μάθει πλαισιωμένες ενσωματώσεις. Για τον κωδικοποιητή, χρησιμοποιήθηκε ένα αμφίδρομο ΔΜΒΜ δύο επιπέδων ενώ ο αποκωδικοποιητής χρησιμοποίησε μονοκατευθυντικά ΔΜΒΜ Προσοχής. Κατά τη διάρκεια της εκπαίδευσης, ο κωδικοποιητής πρέπει να μάθει πώς να καταγράφει συντακτικές και σημασιολογικές έννοιες των λέξεων και να βγάζει πλαισιωμένες ενσωματώσεις. Στη συνέχεια, αυτός ο προεκπαιδευμένος κωδικοποιητής μπορεί να χρησιμοποιηθεί για την επίλυση ενός διαφορετικού προβλήματος. Φάνηκε στο [57] ότι η προσθήκη αυτών των διανυσμάτων περιβάλλοντος (CoVe) βελτιώνει την απόδοση σε σχέση με τη χρήση μόνο διανυσμάτων λέξεων και χαρακτήρων εξαγόμενα κατά την μη εποπτευόμενη εκπαίδευση (πρόβλημα πρόβλεψης επόμενης λέξης), σε μια μεγάλη ποικιλία εργασιών ΕΦΓ.

Περιορισμοί

- Εποπτευόμενη εκπαίδευση - Περιορισμένα επισημασμένα δεδομένα.
- Η αρχιτεκτονική για την επίλυση του τελικού προβλήματος ΕΦΓ παραμένει απροσδιόριστη (δεν αρκεί αυτό το μοντέλο για την επίλυση όλου του προβλήματος).

Στόχος

- Μη-εποπτευόμενη εκπαίδευση - Απεριόριστα δεδομένα.
- Εφαρμογή του γλωσσικού μοντέλου με μικρές τροποποιήσεις σε διάφορα προβλήματα ΕΦΓ.

ELMo

Το ELMo προτάθηκε από τον M. Peters το 2018 [58]. Το ELMo μοντελοποιεί και τα δύο πολύπλοκα χαρακτηριστικά της χρήσης λέξεων (σύνταξη και σημασιολογία), και πώς αυτές οι χρήσεις ποικίλουν μεταξύ των γλωσσικών πλαισίων (δηλαδή, για να μοντελοποιήσουν

την πολυσημία). Η διαφορά από τα ΔΠ είναι ότι οι ενσωματώσεις μαθαίνονται με την μη εποπτευόμενη εκπαίδευση του ELMo. Συγκεκριμένα, το μοντέλο μαθαίνει να προβλέπει την επόμενη/προηγούμενη λέξη δεδομένων των προηγούμενων/επόμενων (αμφίδρομη). Η αρχιτεκτονική του ELMo αποτελείται από στοιβαγμένα αμφίδρομα ΔMBM. Η μη εποπτευόμενη εκπαίδευση έλυσε τον περιορισμό των ανεπαρκών δεδομένων με ετικέτα και επέτρεψε την εκπαίδευση με τεράστια κείμενα. Το πρόβλημα που παραμένει άλυτο όμως, είναι ότι χρειαζόμαστε ακόμα ένα επιπλέον μοντέλο για την επίλυση προβλημάτων ΕΦΓ (το ELMo μας δίνει μόνο τις ενσωματώσεις).

Μετασχηματιστής

Σε αυτή την ενότητα θα συζητήσουμε για την αρχιτεκτονική του Μετασχηματιστή στην οποία βασίζονται τα ακόλουθα δύο μοντέλα γλώσσας: BERT και GPT. Ο Μετασχηματιστής είναι ένα μοντέλο βαθιάς μάθησης που προτάθηκε από τον Vaswani το 2017 [54]. Συνδυάζει την αρχιτεκτονική ακολουθία-σε-ακολουθία (κωδικοποιητής-αποκωδικοποιητής) [52] με τον μηχανισμό προσοχής [53]. Ο Μετασχηματιστής χρησιμοποιήθηκε αρχικά σε εργασίες ΕΦΓ όπως η αυτόματη μετάφραση και η σύνοψη κειμένου, αλλά βρίσκεται επίσης σε εργασίες υπολογιστικής όρασης.

Μετασχηματιστής vs ENΔ

Ομοιότητες

- Και τα δύο μπορούν να διαχειριστούν ακολουθιακά δεδομένα εισόδου.

Διαφορές

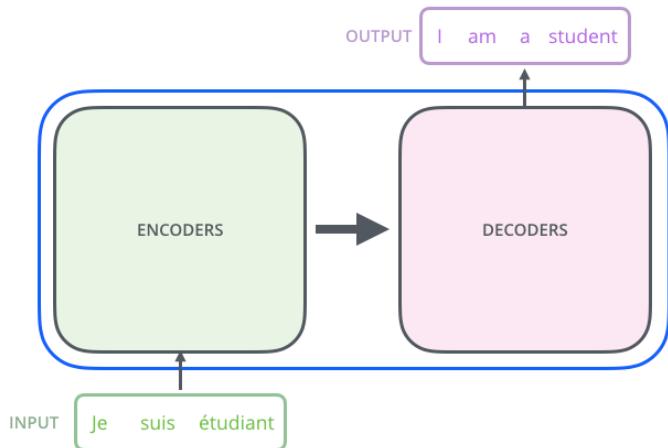
- Οι Μετασχηματιστές επεξεργάζονται όλη την ακολουθία εισόδου ταυτόχρονα (λόγω του μηχανισμού προσοχής), όχι διαδοχικά (λέξη προς λέξη) όπως τα ENΔ.
- Αυτό το χαρακτηριστικό επιτρέπει παραλληλοποίηση και επομένως μειώνει τους χρόνους εκπαίδευσης σε σχέση με τα ENΔ.
- Ο παραλληλισμός (μεγαλύτερες ταχύτητες) επιτρέπει την εκπαίδευση σε μεγαλύτερα σύνολα δεδομένων.
- Αυτό οδήγησε στην ανάπτυξη προεκπαίδευμένων γλωσσικών μοντέλων όπως το BERT και το GPT.

Οπτική περιγραφή Μετασχηματιστή

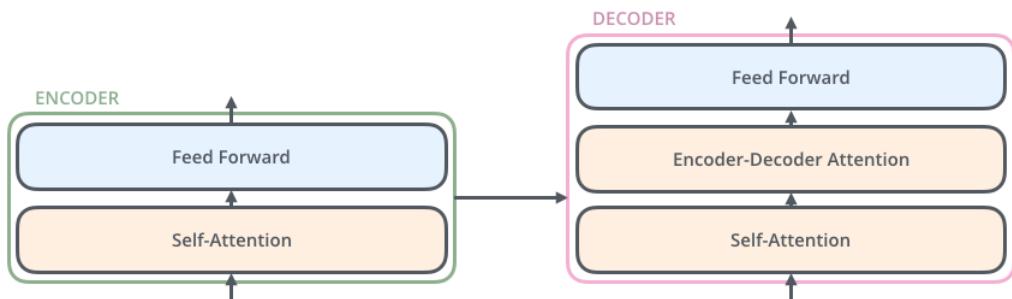
Παρακάτω παρουσιάζουμε την αρχιτεκτονική του Μετασχηματιστή και τον τρόπο λειτουργίας του σε μια εργασία ακολουθία-σε-ακολουθία (Αυτόματη Μετάφραση) [85].

Βασικά Στοιχεία

Τα βασικά δομικά στοιχεία της αρχιτεκτονικής του Μετασχηματιστή είναι οι στοίβες κωδικοποιητών και αποκωδικοποιητών όπως φαίνεται στο σχήμα 2.23. Στην αρχική αρχιτεκτονική, υπάρχουν έξι κωδικοποιητές και έξι αποκωδικοποιητές σε κάθε στοίβα.



Σχήμα 2.23: Οι στοίβες κωδικοποιητών και αποκωδικοποιητών του Μετασχηματιστή.

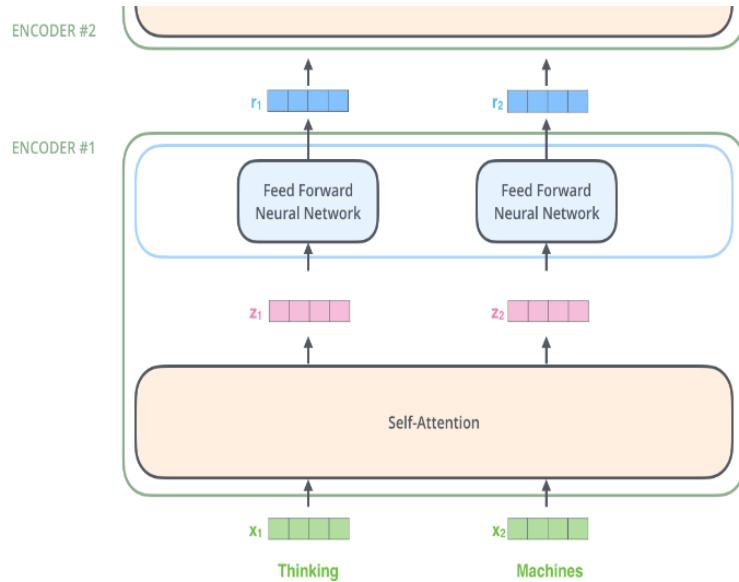


Σχήμα 2.24: Η εσωτερική δομή των μπλοκ Κωδικοποιητή-Αποκωδικοποιητή.

Κάθε κωδικοποιητής αποτελείται από ένα επίπεδο αυτο-προσοχής και ένα νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης. Το επίπεδο αυτο-προσοχής βοηθά τον κωδικοποιητή να κοιτάξει άλλες λέξεις στην πρόταση εισαγωγής καθώς κωδικοποιεί μια συγκεκριμένη λέξη. Ο αποκωδικοποιητής έχει ένα επιπλέον επίπεδο προσοχής που βοηθά στην εστίαση σε σχετικά μέρη της πρότασης εισαγωγής. Το σχήμα 2.24 δείχνει την εσωτερική δομή των μπλοκ κωδικοποιητή-αποκωδικοποιητή.

Ροή δεδομένων στον Κωδικοποιητή

Όλοι οι κωδικοποιητές λαμβάνουν μια λίστα με διανύσματα το καθένα με μέγεθος 512. Περνούν αυτά τα διανύσματα σε ένα στρώμα αυτο-προσοχής και μετά σε ένα νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης που στέλνει την έξοδο προς τα πάνω στον επόμενο κωδικοποιητή (φαίνεται στο σχήμα 2.25).



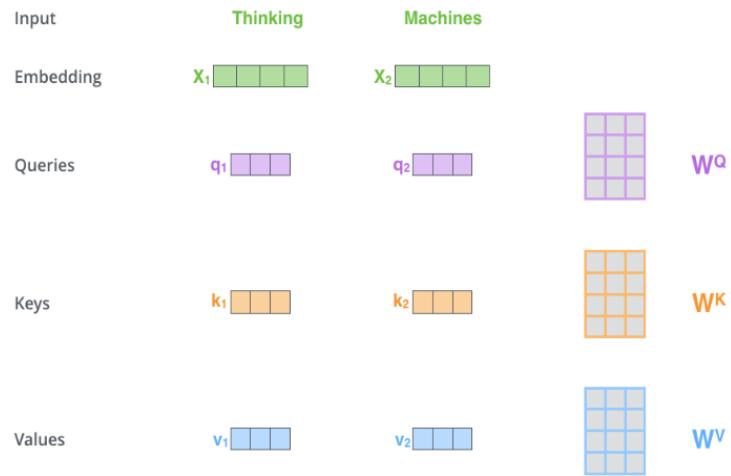
Σχήμα 2.25: Ροή δεδομένων στον Κωδικοποιητή.

Αυτο-προσοχή

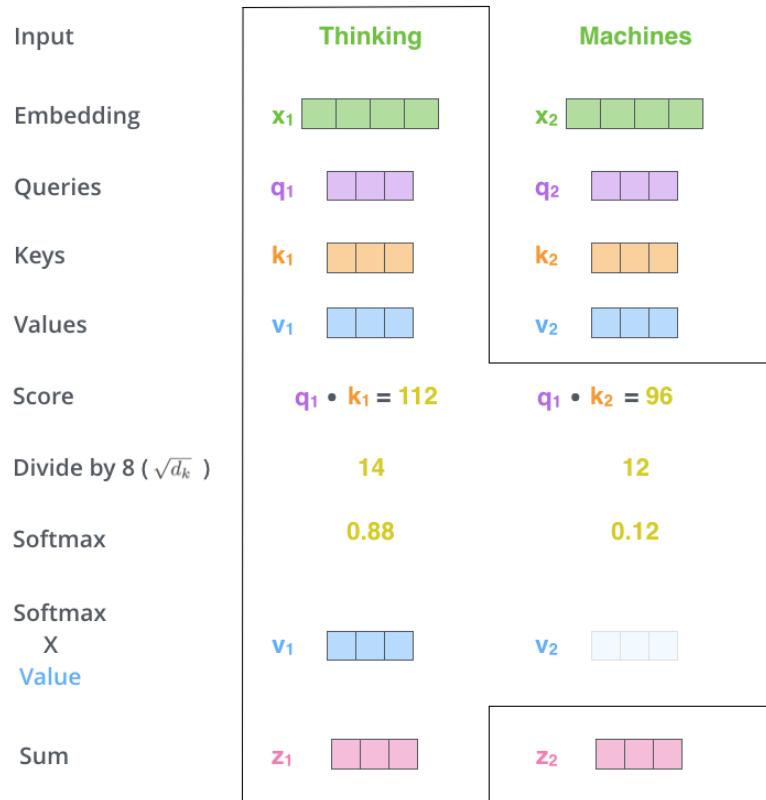
Η έξοδος του επιπέδου αυτο-προσοχής υπολογίζεται με τα ακόλουθα έξι βήματα (που απεικονίζονται στο σχήμα 2.27):

1. Δημιουργησε τα διανύσματα Query, Key και Value (μέγεθος=64) από κάθε ένα από τα διανύσματα εισόδου του κωδικοποιητή (μέγεθος=512). Αυτά τα διανύσματα δημιουργούνται πολλαπλασιάζοντας τα διανύσματα εισόδου με τους πίνακες βαρών W_q , W_k , W_v (σχήμα 2.26) που εκπαιδεύσαμε κατά τη διάρκεια της εκπαίδευτικής διαδικασίας ως τις παραμέτρους του μοντέλου. Για τον πρώτο μόνο κωδικοποιητή τα διανύσματα εισόδου είναι οι αρχικές ενσωματώσεις λέξεων.
2. Χρησιμοποιώντας τα q, k υπολόγισε μια τιμή (σκορε) για κάθε λέξη που καθορίζει πόση προσοχή θα δώσουμε σε κάθε όρο της ακολουθίας κατά την κωδικοποίηση μιας συγκεκριμένης λέξης.
3. Διαιρεσε τα σκορες με την τετραγωνική ρίζα της διάστασης των key διανυσμάτων.
4. Πέρνα το αποτέλεσμα από μια softmax συνάρτηση για την κανονικοποίηση των scores. Αυτή η softmax τιμή καθορίζει πόσο θα εκφραστεί κάθε λέξη στην τρέχουσα θέση.

5. Πολλαπλασίασε κάθε διάνυσμα αλυε με την softmax τιμή.
6. Άθροισε τα σταθμισμένα διανύσματα αλυε για να παραχθεί η έξοδος του επιπέδου αυτο-προσοχής σε αυτή τη θέση.



Σχήμα 2.26: Διανύσματα Query, Key και Value.



Σχήμα 2.27: Τα βήματα υπολογισμού της αυτο-προσοχής.

Αυτό ολοκληρώνει τον υπολογισμό της αυτο-προσοχής. Το διάνυσμα που προκύπτει είναι ένα διάνυσμα που μπορούμε να στείλουμε στο νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης.

Στην πραγματική υλοποίηση, ωστόσο, αυτός ο υπολογισμός γίνεται σε μορφή πινάκων για ταχύτερη επεξεργασία. Συγκεκριμένα, οι ενσωματώσεις εισόδου συσκευάζονται σε έναν πίνακα Ξ και στη συνέχεια πολλαπλασιάζονται με τους πίνακες βάρους W_q , W_k , W_v για να μας δώσουν τους πίνακες Query, Key και Value (αντί για διανύσματα) αντίστοιχα. Τέλος, τα βήματα δύο έως έξι μπορούν να εκτελεστούν με υπολογισμούς πινάκων όπως φαίνεται στο σχήμα 2.28 για να παραχθεί ο πίνακας εξόδου Z του επιπέδου αυτο-προσοχής.

$$\text{softmax} \left(\frac{\begin{array}{|c|c|} \hline \textcolor{violet}{Q} & \textcolor{orange}{K^T} \\ \hline \end{array} \times \begin{array}{|c|c|} \hline \textcolor{brown}{V} \\ \hline \end{array}}{\sqrt{d_k}} \right) = \textcolor{red}{Z}$$

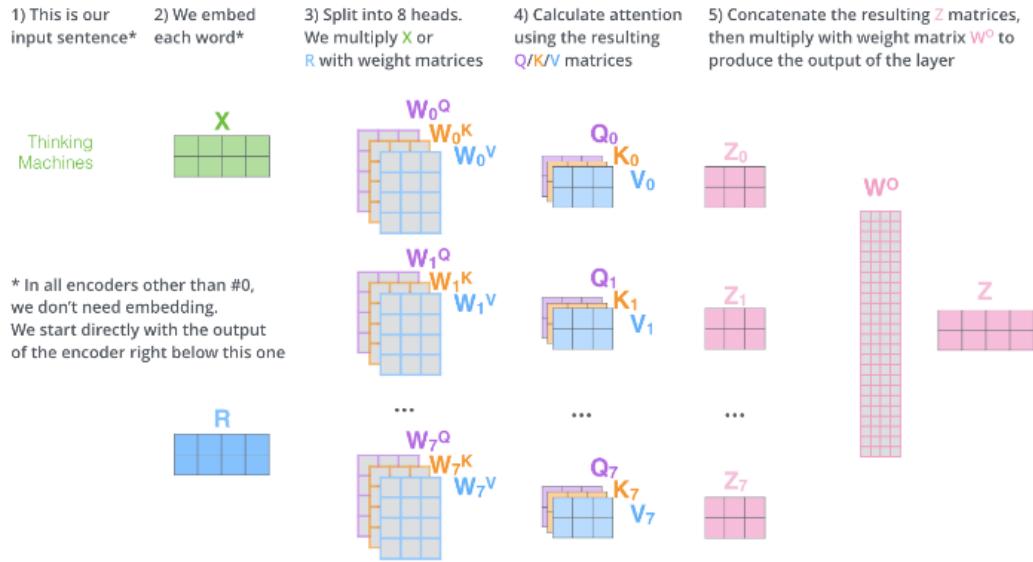
Σχήμα 2.28: Τα βήματα υπολογισμού της αυτο-προσοχής σε μορφή πινάκων.

Προσοχή πολλαπλών Κεφαλών

Το όρθρο [54] βελτίωσε περαιτέρω το επίπεδο αυτο-προσοχής προσθέτοντας έναν μηχανισμό που ονομάζεται «προσοχή πολλαπλών κεφαλών». Αυτό βελτιώνει την απόδοση του επιπέδου προσοχής με δύο τρόπους:

1. Επεκτείνει την ικανότητα του μοντέλου να εστιάζει σε διαφορετικές θέσεις. Στο παραπάνω παράδειγμα, το ζ1 περιέχει λίγο από κάθε άλλη κωδικοποίηση λέξεων, αλλά θα μπορούσε να κυριαρχείται από την ίδια τη λέξη.
2. Δίνει στο επίπεδο προσοχής πολλαπλούς «υποχώρους αναπαράστασης» αφού δεν υπάρχουν μόνο ένα, αλλά πολλαπλά σύνολα πινάκων βάρους Ερώτημα/Κλειδί/Τιμή.

Κάθε ένα από αυτά τα σύνολα αρχικοποιείται τυχαία. Στη συνέχεια, μετά την εκπαίδευση, κάθε σύνολο χρησιμοποιείται για την προβολή των ενσωματώσεων εισόδου (ή διανυσμάτων από χαμηλότερους κωδικοποιητές/αποκωδικοποιητές) σε διαφορετικό υποχώρο αναπαράστασης.

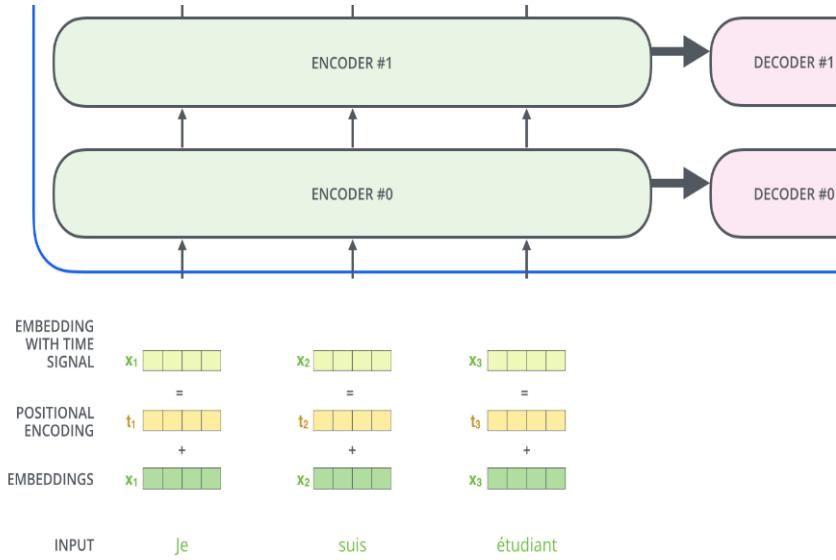


Σχήμα 2.29: Υπολογισμός Προσοχής πολλαπλών Κεφαλών.

Στην προσοχή πολλαπλών κεφαλών κάνουμε τον ίδιο υπολογισμό της αυτο-προσοχής οκτώ φορές με διαφορετικούς πίνακες βάρους και καταλήγουμε σε οκτώ διαφορετικούς πίνακες εξόδου Z . Το πρόβλημα είναι ότι το επίπεδο εμπρόσθιας τροφοδότησης αναμένει έναν ενιαίο πίνακα και όχι οκτώ. Αυτός είναι ο λόγος για τον οποίο αυτοί οι οκτώ πίνακες ενώνονται και στη συνέχεια πολλαπλασιάζονται με έναν επιπλέον πίνακα βαρών W^O που εκπαιδεύτηκε από κοινού με το μοντέλο για να ληφθεί ο τελικός πίνακας Z και να τον περάσει στο στρώμα εμπρόσθιας τροφοδότησης. Τα βήματα υπολογισμού της προσοχής πολλαπλών κεφαλών φαίνονται στο σχήμα 2.29.

Κωδικοποιήσεις θέσης

Για να ληφθεί υπόψη η σειρά των λέξεων στην ακολουθία εισόδου, ο Μετασχηματιστής προσέτει ένα διάνυσμα - κωδικοποίησης θέσης - σε κάθε ενσωμάτωση εισόδου όπως φαίνεται στο σχήμα 2.30.

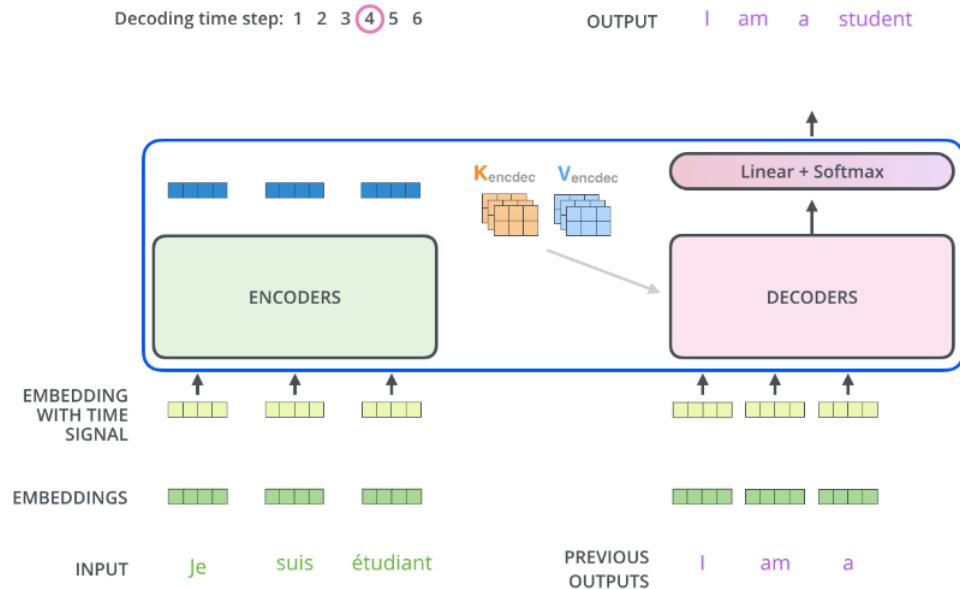


Σχήμα 2.30: Πρόσθεση κωδικοποιήσεων θέσης για να ληφθεί υπόψη η σειρά εμφάνισης των λέξεων στην ακολουθία.

Αποκωδικοποιητής

Οι κωδικοποιητές ξεκινούν με την επεξεργασία ολόκληρης της ακολουθίας εισόδου. Η έξοδος του επάνω κωδικοποιητή μετατρέπεται στη συνέχεια σε ένα σύνολο διανυσμάτων προσοχής Κ και V. Αυτά πρέπει να χρησιμοποιηθούν από κάθε αποκωδικοποιητή στο επίπεδο «προσοχή κωδικοποιητή-αποκωδικοποιητή» που βοηθά τον αποκωδικοποιητή να εστιάσει σε κατάλληλα σημεία στην ακολουθία εισόδου. Σε κάθε χρονικό βήμα οι αποκωδικοποιητές στοιβάζουν μια λέξη από τη νέα (μεταφρασμένη) ακολουθία. Η έξοδος κάθε βήματος τροφοδοτείται στον κάτω αποκωδικοποιητή στο επόμενο χρονικό βήμα, και οι αποκωδικοποιητές στέλνουν προς τα πάνω τα αποτελέσματα αποκωδικοποίησης όπως έκαναν οι κωδικοποιητές. Όπως και οι είσοδοι του κωδικοποιητή, οι κωδικοποιήσεις θέσης προστίθενται σε αυτές τις εισόδους του αποκωδικοποιητή για να υποδειξουν τη θέση κάθε λέξης. Αυτά τα βήματα επαναλαμβάνονται κατα τη διαδικασία μέχρι να φτάσουμε σε ένα ειδικό σύμβολο που υποδεικνύει ότι ο αποκωδικοποιητής του μετασχηματιστή έχει ολοκληρώσει την έξοδο του. Η διαδικασία απεικονίζεται στο σχήμα 2.31.

Τα επίπεδα αυτο-προσοχής στον αποκωδικοποιητή λειτουργούν με ελαφρώς διαφορετικό τρόπο από αυτά στον κωδικοποιητή. Στον αποκωδικοποιητή, το επίπεδο αυτο-προσοχής επιτρέπεται να παρακολουθεί μόνο προηγούμενες θέσεις στην ακολουθία εξόδου. Αυτό γίνεται καλύπτοντας μελλοντικές θέσεις πριν από το βήμα softmax στον υπολογισμό της αυτο-προσοχής. Το επίπεδο "Προσοχή Κωδικοποιητή-Αποκωδικοποιητή" λειτουργεί ακριβώς όπως η αυτο-προσοχή πολλών κεφαλών, με τη διαφορά ότι δημιουργεί τον πίνακα Ερωτημάτων του από το επίπεδο κάτω από αυτό και παίρνει τον πίνακα "Κλειδιά και Τιμές" από την έξοδο της στοίβας κωδικοποιητή.



Σχήμα 2.31: Αυτόματη μετάφραση από άκρο σε άκρο με τον Μετασχηματιστή.

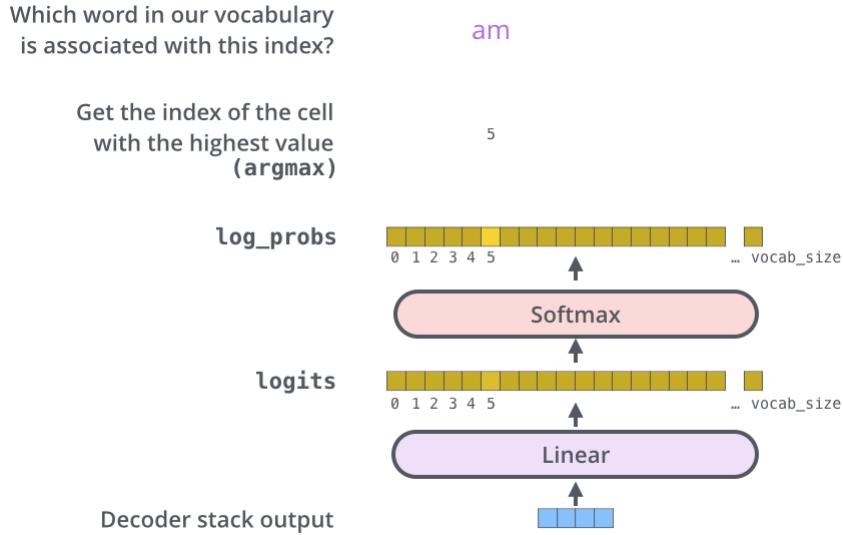
Τελική Έξοδος

Η στοίβα αποκωδικοποιητή εξάγει ένα διάνυσμα πραγματικών αριθμών. Ένα τελικό γραμμικό επίπεδο που ακολουθείται από ένα επίπεδο Softmax είναι υπεύθυνο για τη μετατροπή των διανυσμάτων αυτών σε λέξεις (όπως φαίνεται στο σχήμα 2.32).

Γραμμικό επίπεδο (Πλήρως συνδεδεμένο) : Προβάλλει το διάνυσμα που παράγεται από τη στοίβα των αποκωδικοποιητών, σε ένα πολύ μεγαλύτερο διάνυσμα που ονομάζεται διάνυσμα λογιτς.

Μέγεθος διανύσματος Logits = μέγεθος λεξιλογίου εξόδου (κάθε κελί αντιστοιχεί στο σκορ μιας μοναδικής λέξης)

Επίπεδο Softmax : Μετατρέπει αυτά τα σκορ σε πιθανότητες. Επιλέγεται το κελί με την υψηλότερη πιθανότητα και η λέξη που σχετίζεται με αυτό παράγεται ως έξοδος για αυτό το χρονικό βήμα.



Σχήμα 2.32: Το γραμμικό και το softmax επίπεδο μετατρέπουν την έξοδο της στοίβας του αποκωδικοποιητή σε λέξη.

Προεκπαιδευμένα γλωσσικά μοντέλα που βασίζονται στον Μετασχηματιστή

Γενετικός Προεκπαιδευμένος Μετασχηματιστής OpenAI (GPT)

Στις 11 Ιουνίου 2018, ερευνητές και μηχανικοί του ΟπενΑΙ δημοσίευσαν την αρχική τους εργασία σχετικά με τα γενετικά μοντέλα γλώσσας [59]. Ενώ τα μεγάλα σώματα κειμένου χωρίς ετικέτα είναι άφθονα, τα δεδομένα με ετικέτα για την εκμάθηση εποπτευόμενων εργασιών είναι σπάνια, γεγονός που καθιστά δύσκολη την επαρκή απόδοση των μοντέλων που έχουν εκπαιδευτεί με εποπτευόμενη μάθηση. Για να μετριάσουν αυτόν τον περιορισμό, οι συγγραφείς πρότειναν τη γενετική προεκπαίδευση ενός γλωσσικού μοντέλου σε ένα κείμενο χωρίς ετικέτα, ακολουθούμενη από μια έξτρα ρύθμιση προσαρμογής του μοντέλου σε κάθε συγκεκριμένη εργασία εφαρμογής. Η αρχιτεκτονική του βασίζεται στον αποκωδικοποιητή του Μετασχηματιστή. Σε αντίθεση με το ELMo, το GPT εκπαιδεύεται μόνο για να προβλέπει το μέλλον. Σε αντίθεση με προηγούμενες προσεγγίσεις, το GPT μπορεί να χρησιμοποιηθεί απευθείας για όλες τις τελικές εργασίες με μικρές μόνο τροποποιήσεις. Το GPT υπερτερεί των μοντέλων που έχουν δημιουργηθεί ειδικά για κάθε εργασία, όπως παρουσιάζεται στο [59].

vs ELMo

- Το ELMo δίνει πλουσιότερες ενσωματώσεις λόγω της αμφίδρομης φύσης του (κατανοεί καλύτερα τα συμφραζόμενα της λέξης).
- Το ELMo δεν μπορεί να χρησιμοποιηθεί απευθείας για όλες τις τελικές εργασίες ΕΦΓ (μόνο για καδικοποίηση-μπροστινό άκρο του μοντέλου).

Αναπαραστάσεις Αμφίδρομου Κωδικοποιητή από Μετασχηματιστή (BERT)

Το BERT προτάθηκε από τον J. Devlin (Ερευνητική ομάδα της Google) το 2018 [60]. Είναι στον πυρήνα του ένα μοντέλο γλώσσας Μετασχηματιστή με μεταβλητό αριθμό επιπέδων ικανοποιητή και κεφαλών αυτο-προσοχής. Η αρχιτεκτονική είναι “σχεδόν πανομοιότυπη” με την αρχική υλοποίηση του Μετασχηματιστή στο [54]. Το αρχικό αγγλόφωνο BERTBASE αποτελείται από 12 ικανοποιητές με 12 αμφίδρομες κεφαλές αυτο-προσοχής. Το BEPT είναι προεκπαιδευμένο από δεδομένα μη επισημασμένα που εξάγονται από το BooksCorpus με 800 εκατομμύρια λέξεις και την αγγλική Wikipedia με 2.5 εκατομμύρια λέξεις. Η προεκπαίδευσή του (μη εποπτευόμενη) αποτελείται από δύο προβλήματα. A) Μοντέλο γλώσσας μάσκας: Βρείτε τις καλυμμένες/κρυψές λέξεις κοιτάζοντας τα συμφραζόμενά τους. B) Πρόβλεψη επόμενης πρότασης: Με δύο προτάσεις ως είσοδο, A και B, προσδιορίστε αν η B ακολουθεί την A ή όχι. Σε αντίθεση με τα πρόσφατα μοντέλα αναπαράστασης γλώσσας όπως το OpenAI GPT, το BERT έχει σχεδιαστεί για να μαθαίνει βαθιές αναπαραστάσεις αμφίδρομης κατεύθυνσης από κείμενο μη επισημασμένο (χωρίς ετικέτα), με το να εκπαιδεύεται να προβλέπει τα συμφραζόμενα τόσο από αριστερά όσο και από δεξιά. Το προεκπαιδευμένο μοντέλο BEPT μπορεί να εκπαιδεύεται περαιτέρω με ένα μόνο πρόσθετο επίπεδο εξόδου για τη δημιουργία μοντέλων τελευταίας τεχνολογίας σε ένα ευρύ φάσμα εργασιών, όπως η απάντηση σε ερωτήσεις και η εξαγωγή συμπερασμάτων γλώσσας, χωρίς ουσιαστικές τροποποιήσεις αρχιτεκτονικής για την κάθε συγκεκριμένη εργασία. Το BERT έλαβε υπερσύγχρονα αποτελέσματα σε έντεκα εργασίες ΕΦΓ όπως φαίνεται στο [60].

vs GPT

- Το GPT εκπαιδεύεται μόνο για να προβλέπει το μέλλον (μονοκατευθυντικό), παρέχοντάς μας φτωχότερες ενσωματώσεις σε σύγκριση με το BERT (αμφίδρομο γλωσσικό μοντέλο).
- Τόσο το GPT όσο και το BERT μπορούν να χρησιμοποιηθούν απευθείας για όλες τις τελικές εργασίες ΕΦΓ με μικρές τροποποιήσεις στην αρχιτεκτονική.

Οπτική περιγραφή του BERT

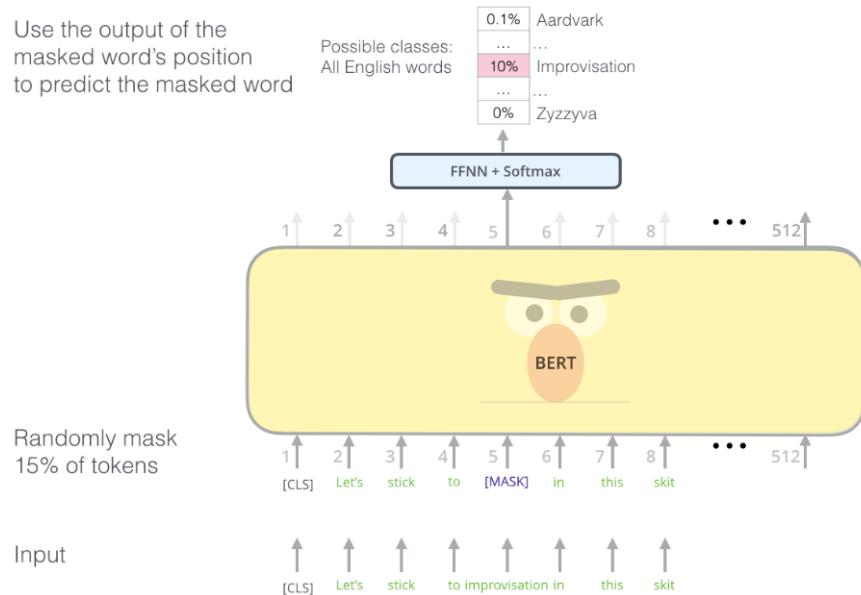
Παρακάτω παρουσιάζουμε πώς το BERT προ-εκπαιδεύεται σε δύο διαφορετικές εργασίες και πώς μπορούμε να χρησιμοποιήσουμε αυτό το μοντέλο για να αντιμετωπίσουμε διάφορες εργασίες ΕΦΓ [86].

Προ-εκπαίδευση

Γλωσσικό Μοντέλο Μάσκας

Το πρόβλημα με την αμφίδρομη ρύθμιση ήταν ότι θα επέτρεπε σε κάθε λέξη να δει έμμεσα τον εαυτό της σε ένα πολυεπίπεδο πλαίσιο. Γι' αυτό εισήχθη η χρήση μασκών για την πρόβλεψη λέξεων. Συγκεκριμένα, το 15% των λέξεων καλύφθηκαν και το BERT εκπαιδεύτηκε να τα

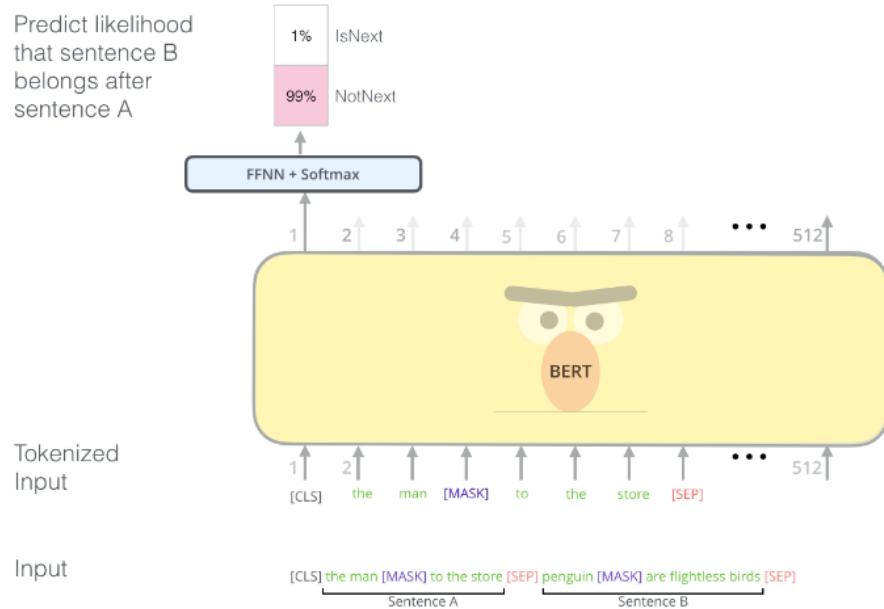
προβλέπει από τα συμφραζόμενα. Πέρα από την απόκρυψη του 15% της εισόδου, το BERT αναμιγνύει επίσης τα πράγματα λίγο για να βελτιώσει τον τρόπο με τον οποίο το μοντέλο προσαρμόζεται στο εκάστοτε πρόβλημα ΕΦΓ αργότερα. Μερικές φορές αντικαθιστά τυχαία μια λέξη με μια άλλη λέξη και ζητά από το μοντέλο να προβλέψει τη σωστή λέξη σε αυτή τη θέση. Μπορούμε να αποκτήσουμε μια διαίσθηση της εκπαιδευτικής διαδικασίας με τη βοήθεια του σχήματος 2.33.



Σχήμα 2.33: Γλωσσικό Μοντέλο Μάσκας.

Πρόβλεψη επόμενης πρότασης

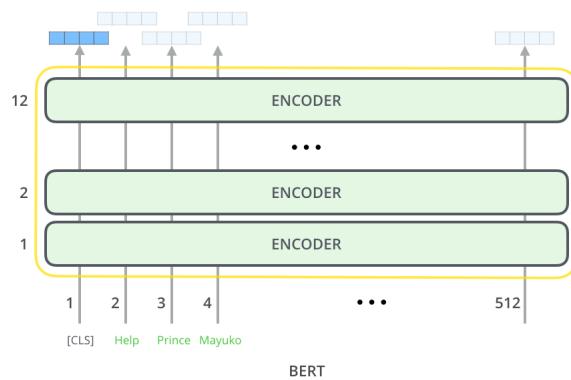
Για να γίνει ο BERT καλύτερος στο χειρισμό των σχέσεων μεταξύ πολλαπλών προτάσεων, η διαδικασία προεκπαίδευσης περιλαμβάνει μια πρόσθετη εργασία: Λαμβάνοντας υπόψη δύο προτάσεις A και B, είναι πιθανό η B να είναι η πρόταση που ακολουθεί την A ή όχι; Μπορούμε να αποκτήσουμε μια διαίσθηση της εκπαιδευτικής διαδικασίας με τη βοήθεια του σχήματος 2.34.



Σχήμα 2.34: Πρόβλεψη επόμενης πρότασης.

Ενσωματώσεις BERT

Η προσέγγιση της προσαρμογής στο εκάστοτε πρόβλημα ΕΦΓ δεν είναι ο μόνος τρόπος χρήσης του BERT. Ακριβώς όπως το ELMo, μπορούμε να χρησιμοποιήσουμε το προεκπαιδευμένο BERT για να δημιουργήσουμε πλαισιωμένες ενσωματώσεις λέξεων. Στη συνέχεια, μπορούμε να τροφοδοτήσουμε αυτές τις ενσωματώσεις στο υπάρχον μοντέλο μας για μια συγκεκριμένη εργασία ΕΦΓ (π.χ. ταξινόμηση κειμένου).



Σχήμα 2.35: Η στοίβα κωδικοποιητών του BERT παράγει πλαισιωμένες ενσωματώσεις λέξεων.

Η έξοδος κάθε επιπέδου καθορίζεται μέχρι της διαδρομής κάθε λέξης μπορεί να χρησιμοποιηθεί ως διάνυσμα χαρακτηριστικών που αναπαριστά αυτή τη λέξη. Μπορούμε να πάρουμε τα διανύσματα εξόδου οποιουδήποτε καθορίζεται θέλουμε, αλλά συνήθως προτιμάμε τον τελευταίο όπως φαίνεται στο σχήμα 2.35. Το μέγεθος ενσωμάτωσης θα είναι ίσο

με το μέγεθος του κρυφού επιπέδου του νευρωνικού δικτύου εμπρόσθιας τροφοδοσίας μέσα στον κωδικοποιητή. Οι ενσωματώσεις που παράγονται για κάθε λέξη περιέχουν πληροφορίες από το περιβάλλον (πλαισιωμένες ενσωματώσεις) λόγω του επιπέδου αυτο-προσοχής μέσα σε κάθε κωδικοποιητή.

2.5 Προβλήματα ΕΦΓ

Υπάρχουν πολλά προβλήματα ΕΦΓ με τα οποία ασχολούνται οι ερευνητές. Τα διαθέσιμα εργαλεία και μέθοδοι έχουν βελτιωθεί σημαντικά την τελευταία δεκαετία, οδηγώντας μας σε καλύτερα αποτελέσματα σε πολλά από αυτά τα προβλήματα. Ωστόσο, τα τεχνητά μας συστήματα απέχουν ακόμη πολύ από την προσομοίωση της ανθρώπινης κατανόησης της φυσικής γλώσσας. Μια λίστα με τα πιο δημοφιλή προβλήματα ΕΦΓ παρουσιάζεται παρακάτω [31].

2.5.1 Επεξεργασία κειμένου και ομιλίας

- *Αναγνώριση Ομιλίας* : Λαμβάνοντας ένα ηχητικό απόσπασμα ενός ατόμου ή ατόμων που μιλούν, καθορίστε τη κειμενική αναπαράσταση της ομιλίας.
- *Κείμενο-σε-Ομιλία* : Δεδομένου ενός κειμένου, μετασχηματίστε αυτές τις μονάδες (λέξεις, προτάσεις κλπ) και δημιουργήστε μια προφορική αναπαράσταση.
- *Διαχωρισμός σε λέξεις (Συμβολοποίηση)* : Διαχωρισμός των λέξεων του κειμένου.

2.5.2 Μορφολογική Ανάλυση

- *Λημματοποίηση* : Αφαίρεση μόνο καμπτικών καταλήξεων και επιστροφή της βασικής μορφής λεξικού μιας λέξης (λήμμα).
- *Στημονοποίηση* : Παρόμοιο με τη Λημματοποίηση απλά δεν επιστρέφει πάντα μια έγκυρη λέξη.
- *Ανίχνευση μέρους του λόγου* : Έχοντας μια πρόταση, καθορίστε τί μέρος του λόγου είναι η κάθε λέξη.

2.5.3 Συντακτική Ανάλυση

- *Γραμματική Επαγωγή* : Δημιουργήστε μια επίσημη γραμματική που περιγράφει τη σύνταξη μιας γλώσσας.

- *Εντοπισμός Προτάσεων* : Με δεδομένο ένα κομμάτι κειμένου, βρείτε τα όρια της κάθε πρότασης.
- *Γραμματική Ανάλυση* : Να προσδιορίσετε το δέντρο ανάλυσης (γραμματική) μιας δεδομένης πρότασης.

2.5.4 Σημασιολογία Λέξεων

- *Σημασιολογία Κατανομής* : Πώς μπορούμε να μάθουμε κατανεμημένες σημασιολογικές αναπαραστάσεις από δεδομένα.
- *Αναγνώριση Ονομαστικής Οντότητας* : Δεδομένης μιας ροής κειμένου, επισημάνετε τα στοιχεία στο κειμένου με τα κατάλληλα ονόματα (π.χ. άτομο, τοποθεσία, οργανισμός).
- *Ανάλυση Συναισθήματος* : Προσδιορίστε και κατηγοριοποιήστε τις απόψεις που εκφράζονται σε ένα κομμάτι κειμένου (π.χ. θετικές, αρνητικές ή ουδέτερες).

2.5.5 Σχεσιακή Σημασιολογία

- *Εξαγωγή Σχέσης* : Λαμβάνοντας ένα κομμάτι κειμένου, προσδιορίστε τις σχέσεις μεταξύ των ονομασμένων οντοτήτων (π.χ. ποιος είναι παντρεμένος με ποιον).
- *Σημασιολογική Επισήμανση Ρόλων* : Δεδομένης μιας πρότασης, προσδιορίστε και αποσαφηνίστε τα σημασιολογικά κατηγορήματα (π.χ. λεκτικά πλαίσια) και, στη συνέχεια, αναγνωρίστε και ταξινομήστε τα στοιχεία πλαισίου (σημασιολογικοί ρόλοι).

2.5.6 Συζήτηση

- *Ανάλυση Αναφοράς* : Με δεδομένη μια πρόταση ή μεγαλύτερο κομμάτι κειμένου, προσδιορίστε ποιες λέξεις („αναφορές“) αναφέρονται στα ίδια αντικείμενα („οντότητες“).
- *Αναγνώριση Θέματος Κειμένου* : Με δεδομένο ένα κομμάτι κειμένου, χωρίστε το σε τμήματα, καθένα από τα οποία είναι αφιερωμένο σε ένα θέμα και προσδιορίστε το θέμα του τμήματος.

2.5.7 Εφαρμογές ΕΦΓ υψηλού επιπέδου

- *Αυτόματη Σύνοψη Κειμένου* : Δημιουργήστε μια ευανάγνωστη περίληψη ενός κομματιού κειμένου.
- *Συγγραφή Βιβλίου* : Δημιουργία ολοκληρωμένων βιβλίων.

- *Απάντηση Ερωτήσεων* : Δεδομένης μιας ερώτησης ανθρώπινης γλώσσας, καθορίστε την απάντησή της.
- *Αυτόματη Μετάφραση* : Αυτόματη μετάφραση κειμένου από μια ανθρώπινη γλώσσα σε άλλη.

Κεφάλαιο 3

Ανάλυση Συναισθήματος

Σε αυτό το κεφάλαιο θα εξετάσουμε σε βάθος ένα συγκεκριμένο πρόβλημα ταξινόμησης κειμένου της ΕΦΓ, γνωστή ως Ανάλυση Συναισθήματος (ΑΣ). Θα ορίσουμε το πρόβλημα της ΑΣ και θα δούμε πώς μπορούμε να το λύσουμε. Συγκεκριμένα, αναφέρουμε συνοπτικά τα απαιτούμενα βήματα προεπεξεργασίας δεδομένων μαζί με ορισμένους βασικούς αλγόριθμους που χρησιμοποιούνται για την ΑΣ. Τέλος, μπαίνουμε σε περισσότερες λεπτομέρειες σχετικά με τα μοντέλα βαθιάς μάθησης που είναι αυτή τη στιγμή η τελευταία λέξη της τεχνολογίας για την ΑΣ. Αυτές οι νευρωνικές αρχιτεκτονικές χρησιμοποιήθηκαν για τα πειράματά μας που παρουσιάζονται στα δύο τελευταία κεφάλαια.

3.1 Σχετικά με την ΑΣ

3.1.1 Τί είναι η ΑΣ;

Η Ανάλυση Συναισθήματος ή *Εξόρυξη Απόψεων* είναι ένα πρόβλημα ταξινόμησης κειμένου που ανήκει στο ευρύ πεδίο της ΕΦΓ. Περιλαμβάνει την αυτόματη ερμηνεία και ταξινόμηση του συναισθήματος (θετικό, αρνητικό ή ουδέτερο) εντός δεδομένων κειμένου χρησιμοποιώντας τεχνικές ανάλυσης κειμένου [87].

ή,

Ανάλυση Συναισθήματος είναι το έργο της ανάθεσης μιας ετικέτας κλάσης σε ένα σώμα γραπτού κειμένου, όπου κάθε κλάση εκφράζει ένα πιθανό συναίσθημα του συγγραφέας σχετικά με το περιεχόμενο του κειμένου. Το συναίσθημα μπορεί απλώς να είναι *Πολικότητα* (που κυμαίνεται από πολύ αρνητική έως πολύ θετική στάση) ή πολυδιάστατο (που προσδιορίζει την παρουσία ή την απουσία διαφορετικών συναισθημάτων). Πρόσθετες ιδιότητες σημασιολογικού κειμένου που συσχετίζονται με τη γνώμη, εκτός από το συναίσθημα, μπορούν επίσης

να εντοπιστούν χρησιμοποιώντας σχεδόν πανομοιότυπους αλγόριθμους (για παράδειγμα, *Μεροληψία ή Σαρκασμός*).

3.1.2 Εφαρμογές της ΑΣ

Η ΑΣ έχει πολλούς τομείς εφαρμογής όπου μας ενδιαφέρει να παρακολουθούμε την κοινή γνώμη. Οι κριτικές (προϊόντων, ταινιών) είναι ένα κλασικό παράδειγμα. Ιστότοποι όπως το Amazon ή το IMDB εφαρμόζουν την ΑΣ για να παρέχουν στους επισκέπτες πληροφορίες σχετικά με το προϊόν/ταινία. Ένα άλλο δημοφιλές πεδίο εφαρμογής της ΑΣ είναι τα μέσα κοινωνικής δικτύωσης (Twitter, Facebook). Αυτές οι πλατφόρμες χρησιμοποιούνται ευρέως από τους ανθρώπους για να εκφράσουν τα συναισθήματα/απόψεις τους. Η ΑΣ βοηθά να εξορύξουμε αυτές τις απόψεις και να αποκτήσουμε μια εικόνα για το τι πιστεύουν οι άνθρωποι για ορισμένα θέματα (κοινωνικά, πολιτικά κ.λπ.). Η παρακολούθηση της κοινής γνώμης είναι επίσης χρήσιμη για τις μεγάλες εταιρείες και τις κυβερνήσεις για την προσαρμογή των προϊόντων/στρατηγικών τους προκειμένου να κερδίσουν την προτίμηση των πελατών/ψηφοφόρων. Τέλος, η ΑΣ μπορεί να χρησιμοποιηθεί για να κάνει προβλέψεις. Αυτές οι προβλέψεις θα μπορούσαν να υποδεικνύουν τον νικητή των εκλογών ή τις μελλοντικές τιμές των μετοχών, με βάση τις τρέχουσες προτιμήσεις του κοινού. Επομένως, γίνεται σαφές ότι η ΑΣ είναι ένα πολύ ισχυρό εργαλείο που παρέχει στον χρήστη πολύτιμες πληροφορίες.

3.2 Ένα πρόβλημα ταξινόμησης κειμένου

Όπως εξηγήσαμε νωρίτερα, η ΑΣ είναι στην πραγματικότητα η ταξινόμηση ενός δείγματος κειμένου σε τρεις συνήθως κατηγορίες (θετικό, αρνητικό, ουδέτερο). Ορισμένες παραλλαγές περιλαμβάνουν επιπλέον κατηγορίες (πολύ θετικό, πολύ αρνητικό) για τη λεπτομερέστερη ΑΣ ή λιγότερες (θετικό, αρνητικό) για την απλή ΑΣ (αλλιώς *Aníxneusη Πολικότητας*). Επιπλέον, *Aníxneusη Μεταφορικού Λόγου* μπορεί να πραγματοποιηθεί με κλάσεις μεταφορικό/κυριολεκτικό, *Aníxneusη Προσβλητικότητας* με κλάσεις προσβλητικό/μη προσβλητικό, *Aníxneusη Μεροληψίας* με κλάσεις μεροληπτικό/αντικειμενικό κ.ο.κ. Η διαδικασία υλοποίησης που ακολουθείται για όλες αυτές τις εργασίες είναι πανομοιότυπη.

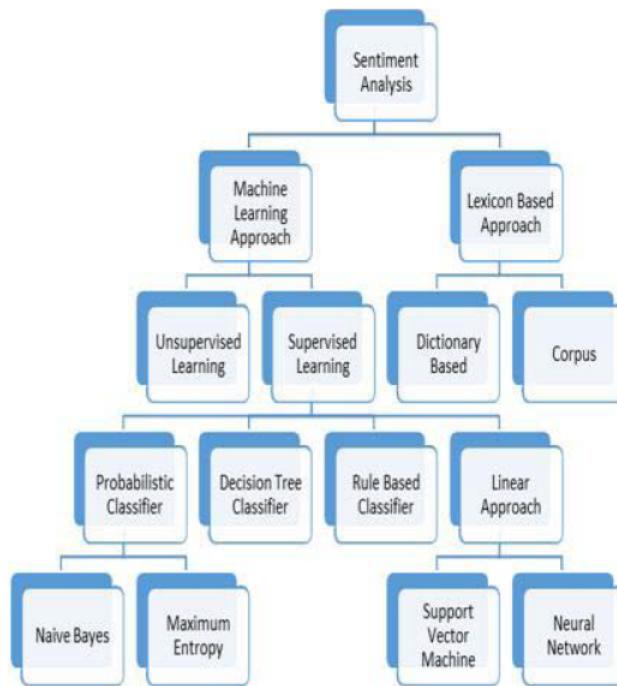
3.2.1 Επίπεδα ταξινόμησης

- *Επίπεδο Πρότασης* : Το κομμάτι κειμένου που πρέπει να ταξινομηθεί είναι μια πρόταση. Τα tweets ή οι τίτλοι άρθρων είναι μερικά παραδείγματα.
- *Επίπεδο Παραγράφου* : Το κομμάτι κειμένου που πρέπει να ταξινομηθεί είναι μια παράγραφος. Οι κριτικές ταινιών/προϊόντων είναι μερικά παραδείγματα.

- *Επίπεδο Εγγράφου* : Το κομμάτι κειμένου που πρέπει να ταξινομηθεί είναι ένα έγγραφο. Τα άρθρα ή τα εμαιλ είναι μερικά παραδείγματα.

3.2.2 Βασικοί Αλγόριθμοι

Οι αλγόριθμοι που χρησιμοποιούνται για την ΑΣ χωρίζονται σε δύο βασικές κατηγορίες: μέθοδοι *Μηχανικής Μάθησης* και μέθοδοι *Βάσει Κανόνων* (χρήση λεξικού). Το σχήμα 3.1 δείχνει μια δενδρική απεικόνιση αλγορίθμων ΑΣ.



Σχήμα 3.1: Αλγόριθμοι που χρησιμοποιούνται στην ΑΣ.

3.2.2.1 Μέθοδοι βάσει Κανόνων (λεξικού)

Αυτή είναι μια πρακτική προσέγγιση για την ανάλυση κειμένου χωρίς εκπαίδευση ή χρήση μοντέλων μηχανικής μάθησης. Το αποτέλεσμα αυτής της προσέγγισης είναι ένα σύνολο χειρόγραφων κανόνων βάσει των οποίων το κείμενο χαρακτηρίζεται ως θετικό/αρνητικό/ουδέτερο. Αυτοί οι κανόνες είναι επίσης γνωστοί ως λεξικά. Ως εκ τούτου, η προσέγγιση βάσει κανόνων ονομάζεται επίσης προσέγγιση βάσει λεξικού [88]. Οι ευρέως χρησιμοποιούμενες προσεγγίσεις βάσει λεξικού είναι το TextBlob, το VADER και το SentiWordNet.

TextBlob

Το TextBlob [89] βαθμολογεί κάθε λέξη στο λεξικό με μια αριθμητική τιμή μεταξύ -1 (αρνητικό) και +1 (θετικό) για να δείξει την πολικότητα της λέξης. Η βαθμολογία 0 ερμηνεύεται ως ουδέτερη. Κατά τον υπολογισμό ενός συναισθήματος για μια μεμονωμένη λέξη, το ΤεξτΒλοβ χρησιμοποιεί την τεχνική "μέσου όρου" που εφαρμόζεται στις τιμές της πολικότητας για τον υπολογισμό της βαθμολογίας πολικότητας για μια μεμονωμένη λέξη και επομένως παρόμοια διαδικασία εφαρμόζεται για κάθε λέξη και παίρνουμε μια συνδυασμένη πολικότητα για μεγαλύτερα κείμενα.

VADER

Ομοίως, η ανάλυση συναισθήματος VADER [90] βασίζεται σε ένα λεξικό που αντιστοιχίζει λεξικά χαρακτηριστικά σε βαθμολογίες συναισθήματος. Η βαθμολογία συναισθήματος ενός κειμένου μπορεί να ληφθεί συνοψίζοντας τις βαθμολογίες κάθε λέξης του κειμένου. Η βαθμολογία συναισθήματος μετράται σε μια κλίμακα από -4 (πιο αρνητικό) έως +4 (πιο θετικό). Το 0 αντιπροσωπεύει και πάλι ένα ουδέτερο συναίσθημα. Το VADER επιστρέφει μια βαθμολογία συναισθήματος στην περιοχή -1 έως 1, από το πιο αρνητικό στο πιο θετικό. Αν και οι μεμονωμένες λέξεις έχουν βαθμολογία συναισθήματος μεταξύ -4 και 4, η επιστρεφόμενη βαθμολογία συναισθήματος μιας πρότασης είναι μεταξύ -1 έως 1. Αυτό επιτυγχάνεται με την εφαρμογή μιας τεχνικής κανονικοποίησης στη συνολική βαθμολογία, προκειμένου η έξοδος να είναι από -1 έως 1.

SentiWordNet

Το SentiWordNet [91] λειτουργεί στη βάση δεδομένων που παρέχεται από το WordNet. Το WordNet είναι μια λεξιλογική βάση δεδομένων που συνθέτει αγγλικές λέξεις, ομαδοποιημένες ως συνώνυμες σε ομάδες που είναι γνωστές ως synsets. Όλες οι λέξεις συνδέονται μεταξύ τους με τη σχέση ISA. Για παράδειγμα, ένα αυτοκίνητο είναι ένας τύπος οχήματος, όπως και ένα φορτηγό. Η πρόσθετη λειτουργικότητα που παρέχει είναι το μέτρο της θετικότητας, της αρνητικότητας και της ουδετερότητας όπως απαιτείται για την ΑΣ.

Έτσι, κάθε synset s συνδέεται με μια τιμή $\text{Pos}(s)$: μια βαθμολογία θετικότητας, $\text{Neg}(s)$: μια βαθμολογία αρνητικότητας και $\text{Obj}(s)$: μια βαθμολογία αντικειμενικότητας (ουδετερότητας). Και οι τρεις βαθμολογίες κυμαίνονται εντός των τιμών [0,1] και αθροίζονται σε 1.

$$\text{Pos}(s) + \text{Neg}(s) + \text{Obj}(s) = 1$$

Παράδειγμα ανάλυσης μιας κριτικής ταινίας με το SentiWordNet : "I disliked the movie."

Η βαθμολογία αρνητικότητας για τη λέξη "disliked" είναι 0,5. Τα υπόλοιπα διακριτικά, όπως το "I" και το "the" στην πρόταση θα φιλτραριστούν κατά την προεπεξεργασία. Εν τω μεταξύ, η βαθμολογία θετικότητας και αρνητικότητας της λέξης "movie" είναι μηδενική,

κάνοντας έτσι τη βαθμολογία αντικειμενικότητάς της 1,0. Έτσι, το συνολικό συναισθήμα της πρότασης θα είναι αρνητικό, αφού μόνο ουδέτεροι και αρνητικοί όροι χρησιμοποιούνται για τον υπολογισμό του συναισθήματος.

3.2.2.2 Μέθοδοι Εποπτευόμενης Μάθησης

Σε αυτήν την προσέγγιση, η μηχανική μάθηση χρησιμοποιείται για την εκπαίδευση μοντέλων με επισημασμένα δεδομένα (εποπτευόμενη μάθηση). Αυτό σημαίνει ότι χρειαζόμαστε ένα σύνολο δεδομένων με δείγματα κειμένου που σχολιάζονται χειροκίνητα με μια ετικέτα συναισθήματος, προκειμένου να εκπαιδεύσουμε τα μοντέλα σχετικά με τον τρόπο παραγωγής του επιθυμητού αποτελέσματος. Έχοντας τα δεδομένα, το μόνο που απομένει είναι να επιλέξουμε το κατάλληλο μοντέλο και να καθορίσουμε τις υπερ-παραμέτρους του για να λάβουμε υψηλές ακρίβειες πρόβλεψης αποφεύγοντας την υπερβολική προσαρμογή στα δεδομένα εκπαίδευσης (καλή απόδοση στα δεδομένα εκπαίδευσης αλλά αδυναμία στη γενίκευση σε νέα δεδομένα). Οι πιο δημοφιλείς ταξινομητές μηχανικής μάθησης παρατίθενται παρακάτω (δείτε 2.3.2.2).

- Αφελής Ταξινομητής Bayes
- Μέγιστη Εντροπία
- Μηχανές Διανυσμάτων Υποστήριξης
- Δέντρα Απόφασης
- Τυχαία Δάση
- K-Κοντινότεροι Γείτονες
- Νευρωνικά Δίκτυα

Δεν θα μπούμε σε περισσότερες λεπτομέρειες σχετικά με τις ιδιαιτερότητες υλοποίησης του καθενός για ΑΣ. Αντίθετα, θα επικεντρωθούμε στα νευρωνικά δίκτυα (τελευταία λέξη της τεχνολογίας) και στο πώς χρησιμοποιούνται για την ανάλυση του συναισθήματος που βρίσκεται στο κείμενο.

3.2.3 Προεπεξεργασία Κειμένου

Τα δεδομένα ακατέργαστου κειμένου ενδέχεται να περιέχουν ανεπιθύμητο ή ασήμαντο κείμενο, λόγω του οποίου τα μοντέλα μας ενδέχεται να μην παρέχουν αποτελεσματική ακρίβεια και να δυσκολεύουν την κατανόηση και την ανάλυση. Επομένως, η σωστή προεπεξεργασία πρέπει να γίνει στα ανεπεξέργαστα δεδομένα πριν από τις περισσότερες εργασίες ΕΦΓ.

Ο απώτερος στόχος του καθαρισμού και της προετοιμασίας δεδομένων κειμένου είναι να περιοριστεί το κείμενο μόνο στις λέξεις που χρειαζόμαστε για την εργασία μας [92].

3.2.3.1 Κοινά βήματα προεπεξεργασίας για ΑΣ

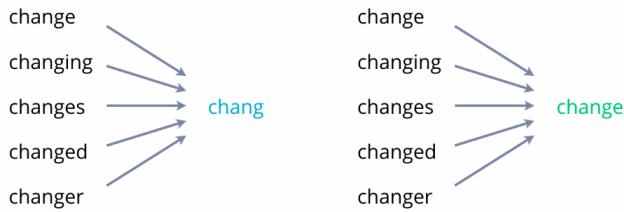
Τα πιο συχνά χρησιμοποιούμενα βήματα προεπεξεργασίας για την ανάλυση συναισθήματος [93] είναι τα ακόλουθα:

- *Καθαρισμός Κειμένου*: Αφαιρέστε όλες τις περιττές λέξεις/χαρακτήρες (ΥΡΛ, ηασηταγς, συνδετικές λέξεις, άρθρα) καθώς είναι συναισθηματικά ουδέτερα.
- *Μετατροπή σε πεζά*: Για να μην κάνει τον συμβολοποιητή να μετράει την ίδια λέξη δύο φορές (μία με πεζά και μία με κεφαλαία).
- *Στημονοποίηση/Λημματοποίηση*: Και πάλι θέλουμε να αποφύγουμε τον αλγόριθμό μας να χειρίζεται χωριστά πολλές λέξεις με το ίδιο νόημα.
- *Συμβολοποίηση*: Μετατρέψτε το κείμενο σε ακολουθίες ακεραίων για να το επεξεργαστεί ευκολότερα το μοντέλο μας.
- *Διανυσματική Αναπαράσταση*: Αντιπροσωπεύστε κάθε λέξη με ένα περιεκτικό, αριθμητικό διάνυσμα χαρακτηριστικών, ώστε το μοντέλο να μπορεί να μάθει ποια χαρακτηριστικά καθορίζουν το συνολικό συναίσθημα.

Στημονοποίηση/Λημματοποίηση

Ο στόχος τόσο της Στημονοποίηση όσο και της Λημματοποίηση [94] είναι η μείωση των μορφών διαφορετικών ιλίσεων και μερικές φορές των παραγώγων σχημάτων μιας λέξης σε μια κοινή βασική μορφή. Η Στημονοποίηση κόβει τα άκρα των λέξεων με την ελπίδα να επιτευχθεί σωστά αυτός ο στόχος τις περισσότερες φορές και συχνά περιλαμβάνει την αφαίρεση των παραγώγων επιθεμάτων. Η Λημματοποίηση χρησιμοποιεί λεξιλόγιο και μορφολογική ανάλυση λέξεων, με στόχο να αφαιρέσει μόνο τις καμπτικές καταλήξεις και να επιστρέψει τη μορφή βάσης ή λεξικού μιας λέξης, η οποία είναι γνωστή ως λήμμα. Το σχήμα 3.2 απεικονίζει ένα παράδειγμα για κάθε μέθοδο.

Stemming vs Lemmatization



Σχήμα 3.2: Στημονοποίηση vs Λημματοποίηση.

Συμβολοποίηση

Η Συμβολοποίηση [95] είναι ένας τρόπος διαχωρισμού ενός τμήματος κειμένου σε μικρότερες μονάδες που ονομάζονται σύμβολα. Τα σύμβολα μπορούν να αναφέρονται είτε σε λέξεις, χαρακτήρες ή υπολέξεις. Ως εκ τούτου, η Συμβολοποίηση μπορεί να ταξινομηθεί ευρέως σε 3 τύπους - Συμβολοποίηση λέξεων, χαρακτήρων και υπολέξεων. Ο πιο δημοφιλής και πρακτικός τύπος είναι ο χαρακτηρισμός λέξεων, καθώς οι λέξεις είναι οι πιο σημαντικές οντότητες και είναι εύκολο να εξαχθούν χρησιμοποιώντας το κενό ως οριοθέτη.

Η Συμβολοποίηση είναι το πιο σημαντικό βήμα κατά τη μοντελοποίηση δεδομένων κειμένου. Εκτελείται με την ανάθεση μιας ακέραιας τιμής (σύμβολο) σε κάθε λέξη του λεξιλογίου. Ο όρος λεξιλόγιο αναφέρεται σε όλες τις μοναδικές λέξεις του κειμένου. Τα σύμβολα που αποκτήθηκαν χρησιμοποιούνται στη συνέχεια για την προετοιμασία ενός λεξικού που περιέχει όλες τις λέξεις λεξιλογίου σε συνδυασμό με το αντίστοιχο σύμβολό τους. Το λεξικό μπορεί να κατασκευαστεί λαμβάνοντας υπόψη κάθε μοναδικό σύμβολο στο κείμενο ή λαμβάνοντας υπόψη τις Κ συχνότερα εμφανιζόμενες λέξεις.

Το σχήμα 3.3 δείχνει ένα παράδειγμα Συμβολοποίησης. Συνήθως, οι πιο συχνές λέξεις εμφανίζονται ψηλότερα στο συμβολικό λεξικό («της», «οφ», «σο»). Οι ακολουθίες από σύμβολα (προτάσεις) τροφοδοτούνται ως είσοδοι στα μοντέλα βαθιάς μάθησης όπου κάθε σύμβολο (λέξη) στη συνέχεια αντιστοιχίζεται στη διανυσματική της αναπαράσταση (ενσωμάτωση λέξης) και επεξεργάζεται από τα νευρωνικά επίπεδα.

- vocabulary** - all unique words in a source of text
token - an integer value assigned to each word in the vocabulary

token dictionary

{'the': 0, 'of': 1, 'so': 2, 'then': 3, 'you': 4, ... 'learn': 3191, ... 'artificial': 30297... }

sample text	tokenized text
"the pettiness of the whole situation"	————— [0, 121241, 1, 0, 988, 25910]

Σχήμα 3.3: Παράδειγμα Συμβολοποίησης.

Διανυσματική Αναπαράσταση Κειμένου

Όπως εξηγήσαμε στο 2.4, δεν μπορούμε να εργαστούμε απευθείας με το αυθεντικό κείμενο ως είσοδο, επομένως χρησιμοποιούμε αναπαραστάσεις λέξεων. Αυτές οι αναπαραστάσεις είναι αριθμητικά διανύσματα και υπάρχουν πολλοί τρόποι για να τα δημιουργήσουμε. Υπάρχουν δύο κύριες κατηγορίες με βάση τα παραγόμενα διανύσματα. Σταθερές (αραιές) και Κατανεμημένες (πυκνές) αναπαραστάσεις. Ωστόσο, οι σταθερές αναπαραστάσεις δεν μπορούν να εξηγήσουν την ομοιότητα μεταξύ των λέξεων, καθώς δεν περιέχουν πληροφορίες σχετικά με τη σημασία των λέξεων. Από την άλλη πλευρά, οι κατανεμημένες αναπαραστάσεις ενσωματώνουν (ενσωματώσεις λέξεων) τη σημασιολογία των λέξεων σε διανύσματα πραγματικών τιμών. Ως εκ τούτου, οι λέξεις με παρόμοια σημασία έχουν παρόμοιες διανυσματικές αναπαραστάσεις. Επιπλέον, η πυκνότητα των κατανεμημένων αναπαραστάσεων κατέστησε υπολογιστικά δυνατό το βαρύ έργο της εκπαίδευσης ενός νευρωνικού δικτύου με τέτοια δεδομένα. Το σχήμα 3.4 δείχνει ένα παράδειγμα ενσωματώσεων τριών λέξεων.

~300 columns

the →	0.2	0.4	-0.1
good →	0.7	-0.5	0.3
movie →	0.1	0.2	0.6

Σχήμα 3.4: Τρεις ενσωματώσεις λέξεων.

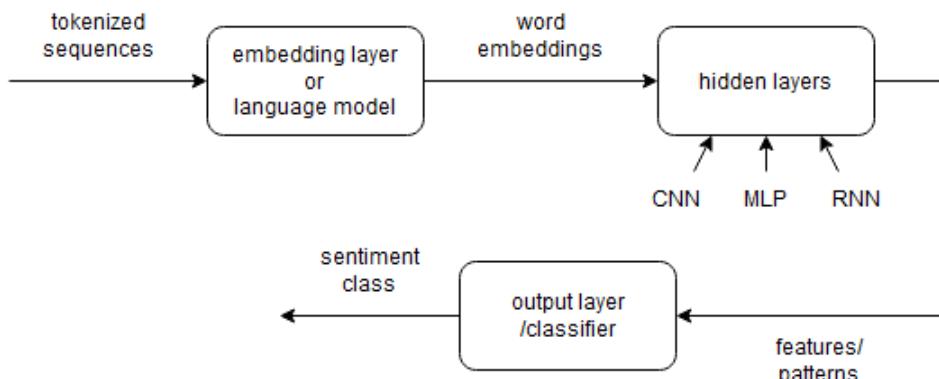
3.2.4 Ανάλυση Συναισθήματος με Νευρωνικά Δίκτυα (ΕΝΔ, ΣΝΔ, BERT)

3.2.4.1 Βασική Ιδέα

- Είσοδος: Διανύσματα πραγματικών τιμών (ενσωματώσεις λέξεων) που αντιπροσωπεύουν ένα κομμάτι κειμένου. Για παράδειγμα, μια πρόταση μπορεί να αναπαρασταθεί

ως 2Δ πίνακας όπου κάθε σειρά είναι μια ενσωμάτωση λέξης.

- Έξοδος: Μια κατανομή πιθανοτήτων μεταξύ κλάσεων συναισθήματος.
- Σε μια νευρωνική αρχιτεκτονική, οι ενσωματώσεις εξάγονται από το επίπεδο ενσωμάτωσης ή ένα μοντέλο γλώσσας (Word2Vec, GloVe, BERT).
- Μετά τη λήψη των ενσωματώσεων, η ταξινόμηση κειμένου και εικόνας είναι σχεδόν η ίδια. Και οι δύο λαμβάνουν έναν πίνακα αναπαράστασης (εικόνα ή πρόταση) ως είσοδο και εξάγουν χαρακτηριστικά υψηλού επιπέδου μέσω των κρυφών επιπέδων. Αυτά τα χαρακτηριστικά τροφοδοτούνται στο επίπεδο εξόδου/ταξινομητή (πλήρως συνδεδεμένο επίπεδο) το οποίο καλείται να ταξινομήσει τελικά το δείγμα εισόδου σε μία από τις κλάσεις συναισθήματος. Το σχήμα ταξινόμησης φαίνεται στην εικόνα 3.5.



Σχήμα 3.5: Γενικό σχήμα Ανάλυσης Συναισθήματος (ταξινόμηση κειμένου).

Δυαδική Ταξινόμηση (Θετικό/αρνητικό) : Μια σιγμοειδής συνάρτηση ενεργοποίησης θα χρησιμοποιηθεί στο επίπεδο εξόδου για την έξοδο μιας ενιαίας τιμής μεταξύ 0 και 1. Αυτοί οι δύο αριθμοί αντιστοιχούν στις δύο κλάσεις του προβλήματος ταξινόμησης και στρογγυλοποιώντας τη σιγμοειδή έξοδο παίρνουμε την τελική κλάση εξόδου που περιγράφει το συναίσθημα που βρίσκεται στην πρόταση εισόδου.

Ταξινόμηση Πολλαπλών Κλάσεων (Θετικό/αρνητικό/ουδέτερο) : Εδώ θα χρησιμοποιηθεί μια συνάρτηση ενεργοποίησης softmax. Σε αυτήν την περίπτωση, ο ταξινομητής θα επέστρεφε μια κατανομή πιθανότητας (3 θετικοί αριθμοί που αθροίζονται σε 1) για τις αντίστοιχες κλάσεις. Για να λάβουμε την τελική κλάση εξόδου, απλώς επιλέγουμε τη μέγιστη τιμή (υψηλότερη πιθανότητα) από τις εξόδους softmax.

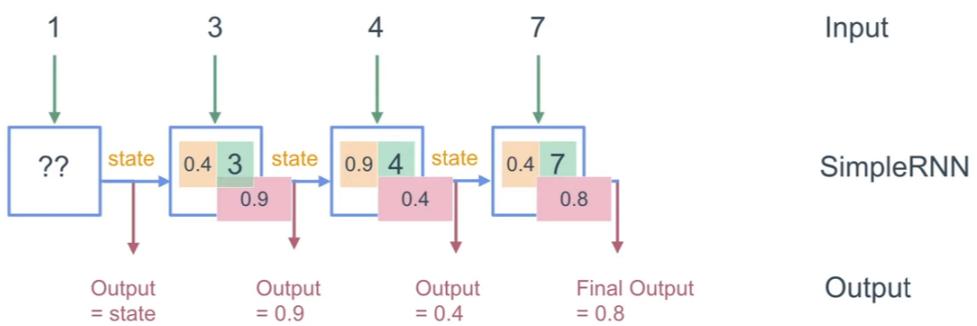
Όπως απεικονίζεται στο σχήμα ταξινόμησης 3.5, τα κρυφά επίπεδα που χρησιμοποιούνται συνήθως είναι συνελικτικά (ΣΝΔ), επαναλαμβανόμενα (ΕΝΔ) ή πλήρως συνδεδεμένα (MLP). Στις επόμενες ενότητες θα αναφερθούμε στις λεπτομέρειες του τρόπου με τον οποίο λειτουργούν τα "NN και τα PNN σε ένα πρόβλημα ΑΣ (ταξινόμηση κειμένου).

3.2.4.2 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (ΕΝΔ)

Τα ΕΝΔ (βλ. 2.3.3.2) έχουν σχεδιαστεί για να χειρίζονται ακολουθιακά δεδομένα όπου η σειρά των στοιχείων της ακολουθίας είναι σημαντικό να μοντελοποιηθεί. Στην ΑΣ, η σειρά με την οποία εμφανίζονται οι λέξεις στο κείμενο δεν παίζει σημαντικό ρόλο. Μας ενδιαφέρει περισσότερο να εντοπίσουμε ειδικές λέξεις που δείχνουν συναίσθημα. Ωστόσο, δεδομένου ότι το κείμενο είναι μια ακολουθία λέξεων, τα ΕΝΔ είναι μια πολύ καλή επιλογή που ταιριάζει στη φύση των δεδομένων. Παρακάτω εξετάζουμε το απλό ΕΝΔ και την παραλλαγή ΔΜΒΜ.

Απλό ΕΝΔ

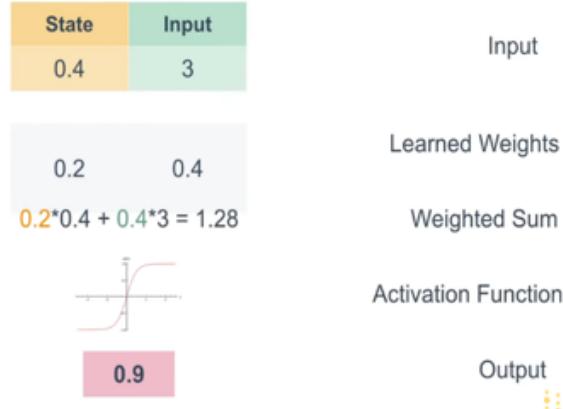
Το σχήμα 3.6 δείχνει πώς μια ακολουθία εισόδου επεξεργάζεται από ένα απλό μοντέλο ΕΝΔ. Η ακολουθία εισόδου εδώ είναι [1, 3, 4, 7]. Η επαναλαμβανόμενη μονάδα/κελί (μπλε πλαίσιο) ξεδιπλώνεται για κάθε χρονικό βήμα καθώς πηγαίνουμε προς τα δεξιά. Στο πρώτο χρονικό βήμα, ο αριθμός 1 είναι η είσοδος του ΕΝΔ και το 0,4 είναι η κατάσταση εξόδου που μεταβιβάζεται στο επόμενο χρονικό βήμα ως είσοδος. Στο δεύτερο χρονικό βήμα, ο αριθμός 3 είναι η είσοδος στο ΕΝΔ μαζί με την προηγούμενη έξοδο/κατάσταση 0,4. Η έξοδος του δεύτερου χρονικού βήματος είναι 0,9 και μεταβιβάζεται ομοίως στην μονάδα επόμενου χρονικού βήματος. Η διαδικασία συνεχίζεται μέχρι να φτάσουμε στον τελικό αριθμό της ακολουθίας (7) όπου λαμβάνουμε την τελική έξοδο του ΕΝΔ (0,8) για ολόκληρη την ακολουθία εισόδου. Η τελική έξοδος αντιστοιχεί σε ολόκληρη την ακολουθία επειδή το ΕΝΔ έχει επεξεργαστεί όλους τους αριθμούς και έχει περάσει τις πληροφορίες από τον καθένα μέσω της εξόδου κάθε χρονικού βήματος ως καταστάσεις. Όλες αυτές οι πληροφορίες σχετικά με προηγούμενα δεδομένα που λαμβάνει η τελική μονάδα για να παράγει την τελική έξοδο μπορούν να ερμηνευθούν ως μνήμη.



Σχήμα 3.6: Επεξεργασία ακολουθίας με ένα απλό ΕΝΔ.

Το σχήμα 3.7 δείχνει πώς υπολογίζεται η έξοδος κάθε επαναλαμβανόμενης μονάδας/κελιού. Οι δύο είσοδοι (τρέχουσα είσοδος + προηγούμενη έξοδος/κατάσταση) πολλαπλασιάζονται

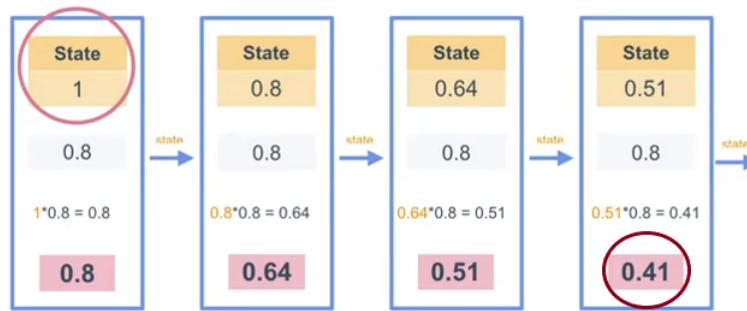
με ορισμένα βάρη. Το σταθμισμένο άθροισμα περνά στη συνέχεια από μια συνάρτηση ενεργοποίησης (\tanh) για να μας δώσει την έξοδο σε εκείνο το χρονικό βήμα (0,9).



Σχήμα 3.7: Υπολογισμοί μέσα στη μονάδα απλού ENΔ.

Πρόβλημα με τα απλά ENΔ – Βραχυπρόθεσμη Μνήμη

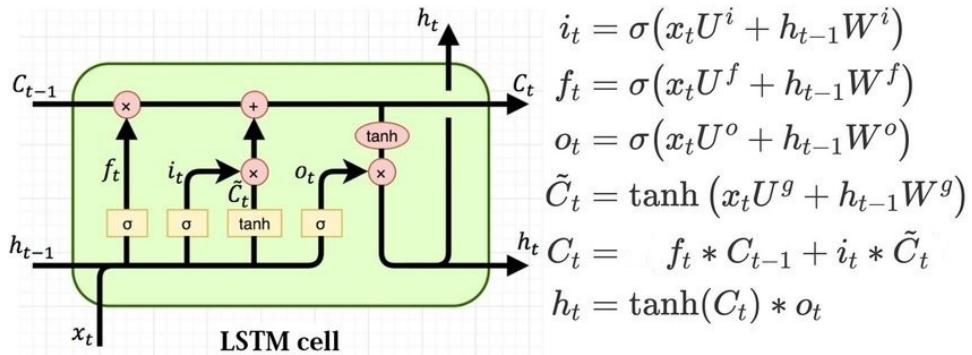
Το πρόβλημα με τα απλά ENΔ είναι ότι είναι πραγματικά δύσκολο για αυτά να κωδικοποιησουν εξαρτήσεις μεγάλης εμβέλειας (μακρυπρόθεσμη μνήμη). Αυτό έχει να κάνει με τον τρόπο κατασκευής του δικτύου. Η κατάσταση εισόδου σε κάθε χρονικό βήμα πολλαπλασιάζεται με ένα βάρος (<1) προκειμένου να υπολογιστεί η έξοδος που είναι ίση με την κατάσταση του επόμενου κελιού. Ωστόσο, είναι πολύ δύσκολο για το δίκτυο να διατηρήσει την ίδια κατάσταση για μεγάλο χρονικό διάστημα. Για παράδειγμα, αν προχωρήσουμε 20 βήματα μπροστά στο χρόνο, αυτοί οι 20 διαδοχικοί πολλαπλασιασμοί κάνουν την αρχική κατάσταση πολύ μικρότερη. Αυτό σημαίνει ότι το δίκτυο στο χρονικό βήμα 20 ξέχασε την κατάσταση στο χρονικό βήμα 1. Το σχήμα 3.8 δείχνει την επίδραση των διαδοχικών πολλαπλασιασμών σε μόλις 4 χρονικά βήματα.



Σχήμα 3.8: Πρόβλημα βραχυπρόθεσμης μνήμης των απλών ENΔ.

Δίκτυα Μακρυπρόθεσμης-Βραχυπρόθεσμης Μνήμης (ΔMBM)

Τα ΔMBM [76] εφευρέθηκαν για να λύσουν το πρόβλημα της μακρυπρόθεσμης μνήμης και να μπορούν να κωδικοποιήσουν τόσο τη βραχυπρόθεσμη όσο και τη μακροπρόθεσμη μνήμη. Στα τυπικά ENΔ, η επαναλαμβανόμενη μονάδα αποτελείται από ένα ενιαίο tanh πλήρως συνδεδεμένο επίπεδο. Ωστόσο, στα ΔMBM η επαναλαμβανόμενη μονάδα έχει τέσσερα πλήρως συνδεδεμένα επίπεδα (3 σιγμοειδή + 1 tanh). Το σχήμα 3.9 μας δίνει μια εικόνα της μονάδας του ΔMBM μαζί με τη μαθηματική του διατύπωση [77].



Σχήμα 3.9: Η μονάδα ΔMBM.

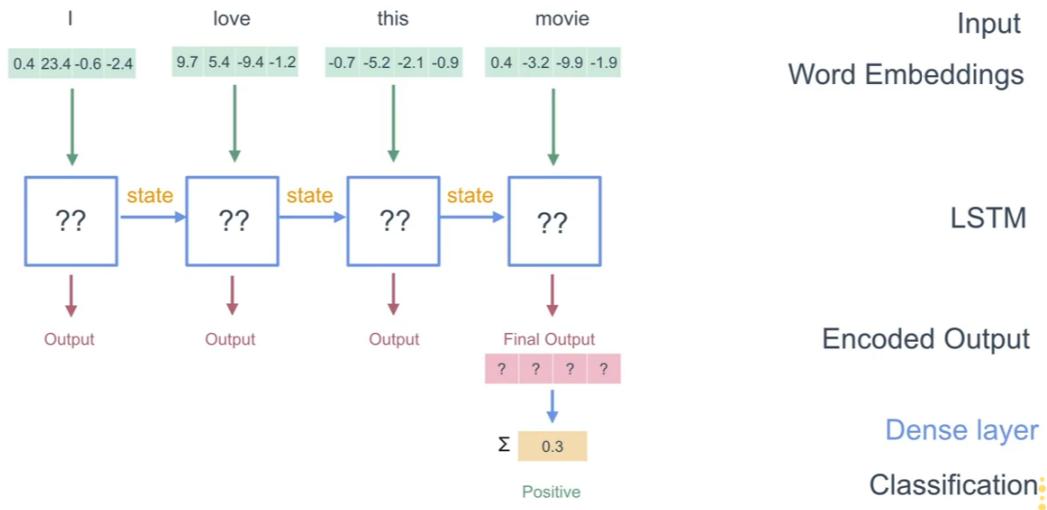
Και τα τέσσερα επίπεδα (U , W) βλέπουν τα h_{t-1} (προηγούμενη έξοδος) και x_t (τρέχουσα είσοδος) και εξάγουν ένα διάνυσμα $(i_t, f_t, o_t, \tilde{C}_t)$ με το ίδιο μέγεθος με την κατάσταση κελιού C_{t-1} . Όλες οι λειτουργίες κόκκινου χρώματος στο σχήμα 3.9 είναι στοιχείο προς στοιχείο.

- Το επίπεδο πύλης Αφαίρεσης αποφασίζει ποιες πληροφορίες θα πετάξουμε από την κατάσταση κελιού ($f_t * C_{t-1}$).
- Το επίπεδο tanh δημιουργεί ένα διάνυσμα νέων υποψήφιων τιμών \tilde{C}_t , που θα μπορούσαν να προστεθούν στην κατάσταση κελιού.
- Το επίπεδο πύλης Εισόδου αποφασίζει ποιες υποψήφιες τιμές θα προστεθούν στην κατάσταση κελιού ($i_t * \tilde{C}_t$).
- Το επίπεδο πύλης Εξόδου αποφασίζει ποια μέρη της ελαφρώς τροποποιημένης νέας κατάστασης κελιού θα αποτελέσουν την έξοδο της μονάδας ΔMBM στο συγκεκριμένο χρονικό σημείο ($\tanh(C_t) * o_t$).

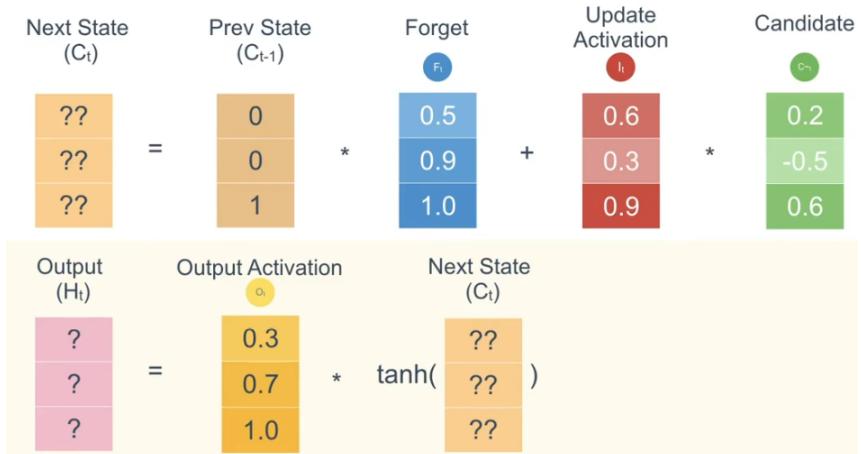
Έτσι, οι πύλες Αφαίρεσης και Εισόδου είναι υπεύθυνες για την ενημέρωση της παλιάς κατάστασης κελιού C_{t-1} , στη νέα κατάσταση κελιού C_t . Και οι τρεις πύλες (Αφαίρεσης, Εισόδου, Εξόδου) έχουν σιγμοειδείς ενεργοποιήσεις προκειμένου να εξάγουν τιμές μεταξύ 0 και 1 και να ελέγχουν τη ροή δεδομένων.

Παράδειγμα ταξινόμησης κειμένου με ΔMBM

Για μια εφαρμογή ΑΣ του ΔMBM, θεωρούμε ως διανύσματα εισόδου x_t τις ενσωματώσεις λέξεων μιας κριτικής ταινίας που θέλουμε να ταξινομήσουμε ως θετικές/αρνητικές (σχήμα 3.10). Οι προαναφερθέντες υπολογισμοί (σχήμα 3.11) θα εκτελεστούν τόσες φορές όσες ο αριθμός των λέξεων στην πρόταση. Η τελική έξοδος του ΔMBM παράγεται στο τελικό χρονικό βήμα (τέλος της ακολουθίας εισόδου) και στη συνέχεια τροφοδοτείται σε ένα σιγμοειδές πλήρως συνδεδεμένο επίπεδο εξόδου που μας δίνει το τελικό αποτέλεσμα ταξινόμησης.



Σχήμα 3.10: Ανάλυση συναισθήματος κριτικών ταινιών με χρήση ΔMBM.



Σχήμα 3.11: Υπολογισμοί που εκτελούνται από τη μονάδα ΔMBM σε κάθε χρονικό βήμα.

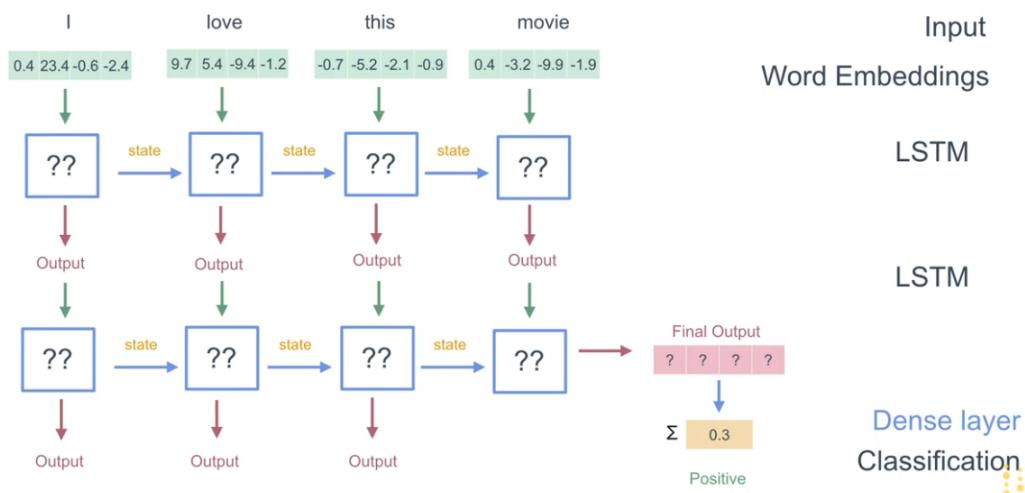
Φυσικά, οι πρώτες προβλέψεις θα είναι ανακριβείς. Αυτός είναι ο λόγος για τον οποίο χρειαζόμαστε τα επισημασμένα δεδομένα με ετικέτα για την εποπτευόμενη μάθηση. Κατά την εκπαίδευση του μοντέλου, κάθε έξοδος συγχρίνεται με την ετικέτα (επιθυμητή έξοδος) και το σφάλμα χρησιμοποιείται για την ενημέρωση των παραμέτρων/βαρών του μοντέλου

προκειμένου να ελαχιστοποιηθεί το σφάλμα. Όταν ολοκληρωθεί η εκπαίδευση, θα έχουμε πιθανώς ένα μοντέλο ικανό να ταξινομήσει σωστά τις προτάσεις.

Άλλες περιπτώσεις χρήσης του μοντέλου ΔMBM για ανάλυση συναισθήματος είναι [96], [97], [98], [99], [100], [101].

Στοιβαγμένο ΔMBM

Τα ΔMBM μπορούν επίσης να στοιβάζονται σε μία ενιαία αρχιτεκτονική [102]. Σε αυτήν την περίπτωση (σχήμα 3.12), τα υψηλότερα ΔMBM περνούν τις εξόδους τους στα χαμηλότερα ως εισόδους. Μπορούμε να χρησιμοποιήσουμε όσα επίπεδα θέλουμε στη στοίβα. Το χαμηλότερο επίπεδο ΔMBM στην αρχιτεκτονική παράγει στο τελικό χρονικό βήμα (τέλος της ακολουθίας εισόδου) την τελική έξοδο. Τα υπόλοιπα βήματα ταξινόμησης είναι τα ίδια όπως με ένα μόνο ΔMBM.



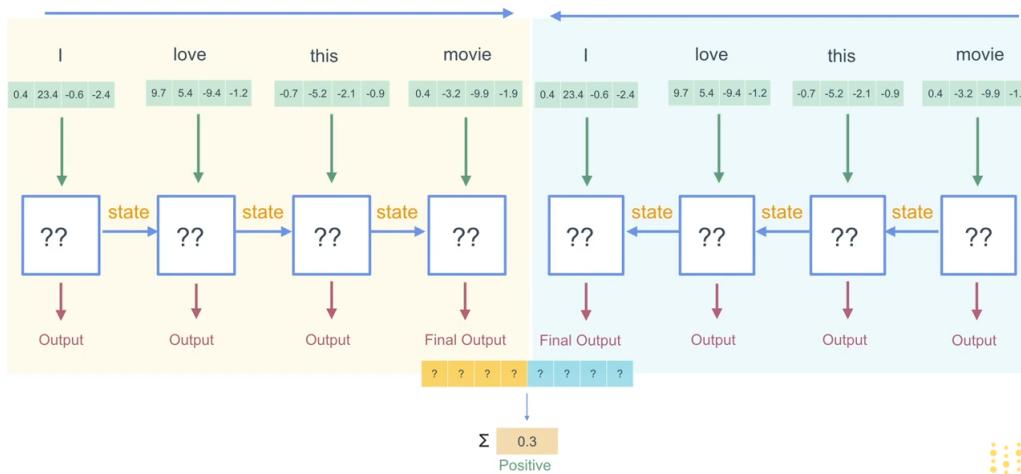
Σχήμα 3.12: Παράδειγμα ανάλυσης συναισθήματος με Στοιβαγμένο ΔMBM.

Γιατί να ανξήσουμε το βάθος;

Η επιτυχία των βαθιών νευρωνικών δικτύων αποδίδεται συνήθως στην ιεραρχία που εισάγεται λόγω των πολλών επιπέδων [103]. Κάθε επίπεδο επεξεργάζεται ένα μέρος του προβλήματος που θέλουμε να λύσουμε και το μεταβιβάζει στο επόμενο. Με αυτή την έννοια, το δίκτυο μπορεί να θεωρηθεί ως ένας αγωγός επεξεργασίας, στον οποίο κάθε επίπεδο επιλύει ένα μέρος του προβλήματος (πιο εύκολο από το συνολικό) πριν το μεταβιβάσει στο επόμενο επίπεδο, έως ότου τελικά το τελευταίο επίπεδο παρέχει την τελική έξοδο. Από μια άλλη οπτική γωνία, τα πρόσθετα επίπεδα εννοείται ότι ανασυνδυάζουν τη μαθημένη αναπαράσταση από προηγούμενα επίπεδα και δημιουργούν νέες αναπαραστάσεις σε επίπεδα υψηλής αφαιρετικότητας. Για παράδειγμα, από γραμμές σε σχήματα και αντικείμενα (υπολογιστική όραση).

Αμφίδρομο ΔMBM

Το αμφίδρομο ΔMBM (AM-ΔMBM) [104] χρησιμοποιεί ένα ΔMBM εμπρόσθιας προσπέλασης και ένα οπίσθιας προσπέλασης (σχήμα 3.13). Το ΔMBM εμπρόσθιας προσπέλασης «βλέπει» τις λέξεις με την κανονική τους σειρά ανάγνωσης (από τα αριστερά προς τα δεξιά). Αντίθετα αυτό της οπίσθιας προσπέλασης τα «βλέπει» ανάποδα. Οι τελικές έξοδοι από αυτά τα δύο ΔMBM ενώνονται για να σχηματίσουν την τελική έξοδο ολόκληρου του AM-ΔMBM. Αυτό το διάνυσμα τροφοδοτείται στη συνέχεια στο επίπεδο εξόδου που μας δίνει την τιμή συναισθήματος.



Σχήμα 3.13: Παράδειγμα ανάλυσης συναισθήματος με το AM-ΔMBM.

Γιατί να χρησιμοποιήσουμε AM-ΔMBM;

Το AM-ΔMBM διατηρεί πληροφορίες τόσο από το παρελθόν όσο και από το μέλλον επειδή έχει δει εισόδους τόσο από το παρελθόν (εμπρόσθια προσπέλαση) όσο και από το μέλλον (οπίσθια προσπέλαση). Ωστόσο, η στήριξη στη γνώση του μέλλοντος φαίνεται εκ πρώτης όψεως να παραβιάζει την αιτιότητα. Πώς μπορούμε να βασίσουμε την κατανόησή μας για όσα ακούσαμε σε κάτι που δεν έχει ειπωθεί ακόμα· Η χρήση της παροχής της αλληλουχίας αμφίδρομα δικαιολογήθηκε αρχικά στον τομέα της Αναγνώρισης Ομιλίας, επειδή υπάρχουν ενδείξεις ότι το πλαίσιο ολόκληρης της εκφοράς χρησιμοποιείται από ανθρώπους για να ερμηνεύσουν αυτό που λέγεται. Με άλλα λόγια, ήχοι, λέξεις, ακόμη και ολόκληρες προτάσεις που στην αρχή δεν σημαίνουν τίποτα, βρίσκουν νόημα υπό το πρίσμα του μελλοντικού πλαισίου. Αυτό που πρέπει να θυμόμαστε είναι η διάκριση μεταξύ εργασιών που είναι πραγματικά ονλινε – που απαιτούν έξοδο μετά από κάθε είσοδο – και εκείνων όπου οι έξοδοι χρειάζονται μόνο στο τέλος κάποιου τμήματος εισόδου. Γι' αυτό χρησιμοποιούμε AM-ΔMBM σε προβλήματα ταξινόμησης ακολουθιών. Σε τέτοιες εργασίες (ανάλυση συναισθήματος) συνήθως αποδίδουν καλύτερα από τα μονοκατευθυντικά ΔMBM καθώς μπορούν να κατανοήσουν καλύτερα το κείμενο [105].

3.2.4.3 Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ)

Τα ΣΝΔ χρησιμοποιούνται παραδοσιακά στην επεξεργασία εικόνας, καθώς τα συνελικτικά φίλτρα έχουν σχεδιαστεί για να βρίσκουν χωρικά μοτίβα. Ωστόσο, την τελευταία δεκαετία έχουν νιοθετηθεί για εργασίες ταξινόμησης κειμένων με εξαιρετικά αποτελέσματα. Επιπλέον, η υλοποίηση ΣΝΔ επιτρέπει τη χρήση ΓΠΥ (παραλληλισμός) που οδηγεί σε γρήγορους χρόνους εκπαίδευσης. Θα δούμε πώς τα ΣΝΔ μπορούν να χρησιμοποιηθούν για την εύρεση γενικών μοτίβων στο κείμενο και την εκτέλεση ταξινόμησης κειμένου με τον παρακάτω διαφωτιστικό οδηγό της Cezanne Camacho [106].

Συνελικτικοί Πυρήνες

Τα συνελικτικά επίπεδα έχουν σχεδιαστεί για να βρίσκουν χωρικά μοτίβα σε μια εικόνα σύροντας ένα μικρό παράθυρο (πυρήνας) πάνω από μια εικόνα. Αυτά τα παράθυρα είναι συχνά μικρά, ίσως 3x3 πίξελς σε μέγεθος, και κάθε κελί του πυρήνα έχει ένα σχετικό βάρος. Καθώς ένας πυρήνας ολισθαίνει πάνω από μια εικόνα, πίξελ προς πίξελ, τα βάρη του πυρήνα πολλαπλασιάζονται με την τιμή του πίξελ στην εικόνα από κάτω και, στη συνέχεια, όλες οι πολλαπλασιασμένες τιμές αθροίζονται για να ληφθεί μια έξοδος, φιλτραρισμένη τιμή πίξελ.

Στην περίπτωση της ταξινόμησης κειμένου, ένας συνελικτικός πυρήνας θα εξακολουθεί να είναι ένα συρόμενο παράθυρο, μόνο που η δουλειά του είναι να εξετάζει τις ενσωματώσεις για πολλές λέξεις, αντί για μικρές περιοχές πίξελ σε μια εικόνα. Οι διαστάσεις του συνελικτικού πυρήνα θα πρέπει επίσης να αλλάξουν, σύμφωνα με αυτήν την εργασία. Για να δούμε σειρές ενσωματώσεων λέξεων, θέλουμε ένα παράθυρο για να δούμε πολλές ενσωματώσεις λέξεων σε μια ακολουθία. Οι πυρήνες δεν θα είναι πλέον τετράγωνοι, αντίθετα, θα είναι ένα φαρδύ ορθογώνιο με διαστάσεις όπως 3x300 ή 5x300 (υποθέτοντας μήκος ενσωμάτωσης 300). Το σχήμα 3.14 δείχνει το σχήμα του συνελικτικού πυρήνα.

	width =	length of embedding						
height =	numbers of words to look at in sequence							
		<table border="1"> <tbody> <tr> <td>0.5</td><td>0.4</td><td>0.7</td></tr> <tr> <td>0.2</td><td>-0.1</td><td>0.3</td></tr> </tbody> </table>	0.5	0.4	0.7	0.2	-0.1	0.3
0.5	0.4	0.7						
0.2	-0.1	0.3						

Σχήμα 3.14: Ένας συνελικτικός πυρήνας.

Συνέλιξη σε ακολουθίες λέξεων

Το σχήμα 3.15 δείχνει ένα παράδειγμα για το πώς μοιάζει ένα ζευγάρι φιλτραρισμένων ενσωματώσεων λέξεων. Για ευκολία οπτικοποίησης, το μήκος ενσωμάτωσης είναι ίσο με 3. Για να

δούμε δύο λέξεις σε αυτήν την ακολουθία παραδείγματος, μπορούμε να χρησιμοποιήσουμε έναν συνελικτικό πυρήνα 2x3. Τα βάρη του πυρήνα τοποθετούνται πάνω από δύο ενσωματώσεις λέξεων. Σε αυτήν την περίπτωση, η κατεύθυνση προς τα κάτω αντιπροσωπεύει το χρόνο, επομένως, η λέξη "movie" έρχεται αμέσως μετά το "good" σε αυτή τη σύντομη πρόταση. Τα βάρη του πυρήνα και οι τιμές ενσωμάτωσης πολλαπλασιάζονται σε ζεύγη και στη συνέχεια αθροίζονται για να ληφθεί μια ενιαία τιμή εξόδου (0,54).

the →	0.2	0.4	-0.1
good →	0.7	-0.5	0.3
movie →	0.1	0.2	0.6

0.5	0.4	0.7
0.2	-0.1	0.3

convolutional kernel

$$\begin{aligned}
 & 0.5 * 0.7 + 0.4 * -0.5 + 0.7 * 0.3 \\
 & + 0.2 * 0.1 + -0.1 * 0.2 + 0.3 * 0.6 \\
 & = 0.54
 \end{aligned}$$

Σχήμα 3.15: Λειτουργία συνέλιξης που εφαρμόζεται σε κείμενο.

Ένα ΣΝΔ θα περιλαμβάνει πολλούς από αυτούς τους πυρήνες και, καθώς το δίκτυο εκπαιδεύεται, αυτά τα βάρη του πυρήνα βελτιώνονται. Κάθε πυρήνας έχει σχεδιαστεί για να κοιτάζει μια λέξη και τις περιβάλλουσες σε ένα παράθυρο και να βγάζει μια τιμή που καταγράφει κάτι σχετικά με αυτήν τη φράση. Με αυτόν τον τρόπο, η λειτουργία συνέλιξης μπορεί να θεωρηθεί ως εξαγωγή χαρακτηριστικών που βασίζεται σε παράθυρο, όπου τα χαρακτηριστικά είναι μοτίβα σε διαδοχικές ομαδοποιήσεις λέξεων που υποδεικνύουν χαρακτηριστικά όπως το συναίσθημα ενός κειμένου, τη γραμματική λειτουργία διαφορετικών λέξεων κ.λπ.

Αναγνώριση γενικών προτύπων

Όπως εξηγήσαμε στο 2.4.2, παρόμοιες λέξεις θα έχουν παρόμοιες ενσωματώσεις (αναπαραστάσεις). Επομένως, δεδομένου ότι η λειτουργία συνέλιξης είναι απλώς μια γραμμική πράξη σε αυτά τα διανύσματα, όταν ένας συνελικτικός πυρήνας εφαρμόζεται σε διαφορετικά σύνολα παρόμοιων λέξεων, θα παράγει παρόμοια τιμή εξόδου. Αυτό φαίνεται στο σχήμα 3.16.

the →	0.2	0.4	-0.1
good →	0.7	-0.5	0.3
movie →	0.1	0.2	0.6

a →	0.1	0.3	-0.2
fantastic →	0.8	-0.6	0.3
song →	0.2	0.3	0.5

$$\begin{aligned} & 0.5 * 0.7 + 0.4 * -0.5 + 0.7 * 0.3 \\ & + 0.2 * 0.1 + -0.1 * 0.2 + 0.3 * 0.6 \\ & = \textcolor{red}{0.54} \end{aligned}$$

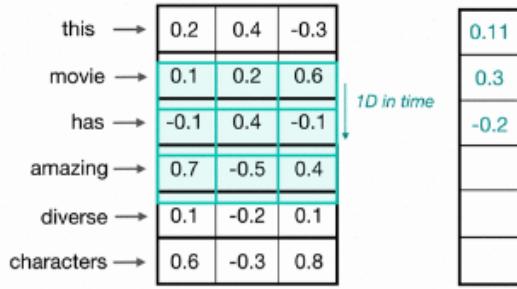
$$\begin{aligned} & 0.5 * 0.8 + 0.4 * -0.6 + 0.7 * 0.3 \\ & + 0.2 * 0.2 + -0.1 * 0.3 + 0.3 * 0.5 \\ & = \textcolor{violet}{0.53} \end{aligned}$$

Σχήμα 3.16: Παράδειγμα συνελικτικού πυρήνα που εντοπίζει θετικά χαρακτηριστικά.

Σε αυτό το παράδειγμα, ο συνελικτικός πυρήνας έχει μάθει να καταγράφει ένα πιο γενικό χαρακτηριστικό. Όχι απλώς μια καλή ταινία ή τραγούδι, αλλά ένα θετικό πράγμα, γενικά. Η αναγνώριση αυτών των χαρακτηριστικών υψηλού επιπέδου είναι το κλειδί για την αντιμετώπιση εργασιών ταξινόμησης κειμένου, οι οποίες συχνά βασίζονται σε γενικές ομαδοποιήσεις. Για παράδειγμα, στην ΑΣ, ένα μοντέλο θα επωφεληθεί από το να μπορεί να αναπαραστήσει αρνητικές, ουδέτερες και θετικές ομάδες λέξεων.

1Δ Συνελίξεις

Για την επεξεργασία μιας ολόκληρης ακολουθίας λέξεων, αυτοί οι πυρήνες θα όλισθησουν' προς τα κάτω προσπελαύνοντας μια λίστα με ενσωματώσεις λέξεων, με τη σειρά. Αυτό ονομάζεται συνέλιξη 1Δ επειδή ο πυρήνας κινείται σε μία μόνο διάσταση: τον χρόνο (ενώ στις εικόνες ο πυρήνας κινείται σε δύο διαστάσεις -2Δ- καθώς ολισθαίνει κατά μήκος της εικόνας). Ένας μεμονωμένος πυρήνας θα κοιτάξει την πρώτη ενσωμάτωση λέξης (και ένα μικρό παράθυρο επόμενων ενσωματώσεων) μετά την επόμενη ενσωμάτωση λέξης και την επόμενη και ούτω καθεξής. Η προκύπτουσα έξοδος θα είναι ένα διάνυσμα χαρακτηριστικών που περιέχει περίπου τόσες τιμές όσες και οι ενσωματώσεις εισόδου. Το σχήμα 3.17 δείχνει την έξοδο μιας συνέλιξης 1Δ που εφαρμόζεται σε μια σύντομη ακολουθία ενσωματώσεων λέξεων (κριτική ταινίας).



Σχήμα 3.17: Έξοδος 1Δ συνέλιξης.

Πολλαπλοί πυρήνες

Σε ένα τυπικό ΣΝΔ, ένας συνελικτικός πυρήνας δεν είναι αρκετός για να ανιχνεύσει όλα τα διαφορετικά είδη χαρακτηριστικών που θα είναι χρήσιμα για μια εργασία ταξινόμησης κειμένου. Για να δημιουργήσουμε ένα δίκτυο έτσι ώστε να είναι σε θέση να μάθει μια ποικιλία διαφορετικών σχέσεων μεταξύ των λέξεων, θα χρειαστούμε πολλούς πυρήνες διαφορετικού ύψους. Έτσι, καθώς ένας συνελικτικός πυρήνας ολισθαίνει πάνω από ενσωματώσεις λέξεων (μία κάθε φορά) καταγράφει τοπικά χαρακτηριστικά ή χαρακτηριστικά σε ένα κοντινό παράθυρο διαδοχικών λέξεων. Τα στοιβαγμένα διανύσματα χαρακτηριστικών εξόδου που προκύπτουν από αρκετές από αυτές τις συνελικτικές λειτουργίες αποτελούν ένα συνελικτικό επίπεδο.

Max-pooling

Αν προσπαθούμε να ταξινομήσουμε κριτικές ταινιών και βλέπουμε τη φράση, "ωραία πλοκή", δεν έχει σημασία πού εμφανίζεται αυτό σε μια κριτική. Είναι μια καλή ένδειξη ότι αυτή είναι μια θετική κριτική, ανεξάρτητα από τη θέση της στην ακολουθία. Για να υποδείξουμε την παρουσία αυτών των χαρακτηριστικών υψηλού επιπέδου, ανεξάρτητα από τη θέση εντός της μεγαλύτερης ακολουθίας εισόδου, χρησιμοποιούμε μια λειτουργία max-pooling. Με το max-pooling επιλέγουμε τη μέγιστη τιμή από ένα οριοθετημένο σύνολο τιμών (ποολ). Αυτή η λειτουργία αναγκάζει το δίκτυο να απορρίψει λιγότερο σχετικές πληροφορίες διατηρώντας μόνο τη μέγιστη τιμή σε ένα διάνυσμα χαρακτηριστικών, το οποίο θα πρέπει να είναι το πιο χρήσιμο. Οι συνελικτικοί πυρήνες παράγουν την ισχυρότερη απόκριση σε ένα τοπικό μοτίβο / χαρακτηριστικό που εμφανίζεται στην είσοδο και γι' αυτό η μέγιστη τιμή του διανύσματος χαρακτηριστικών είναι η πιο χρήσιμη. Ένα παράδειγμα max-pooling απεικονίζεται στο σχήμα 3.18.

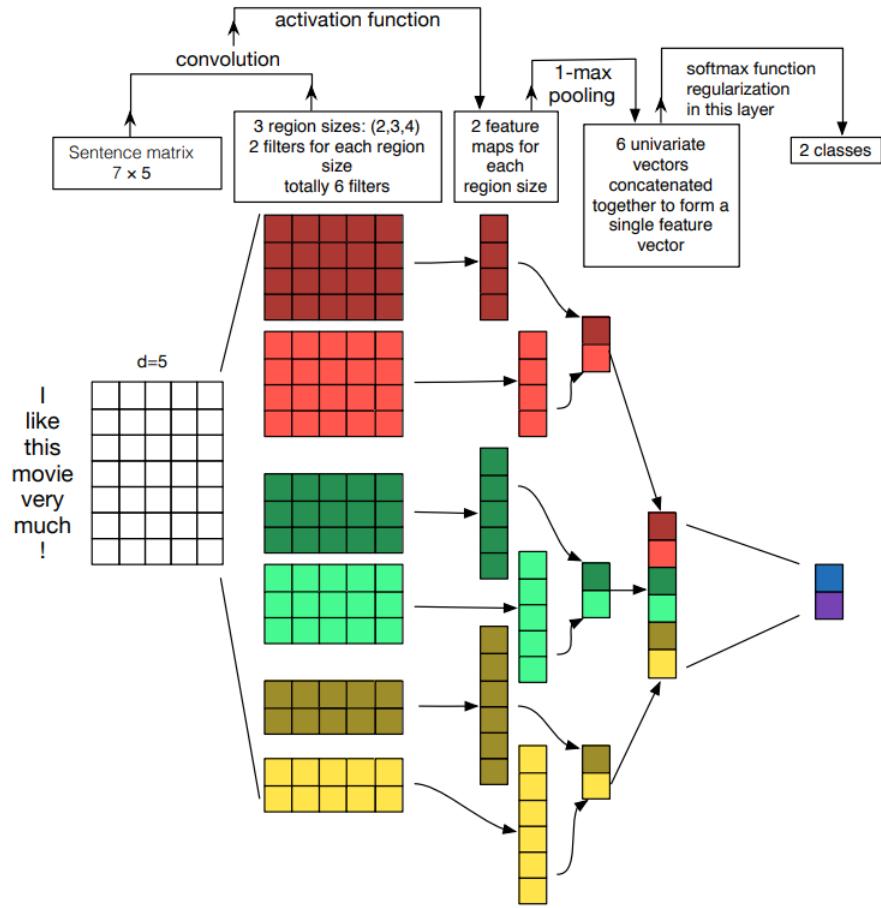


Σχήμα 3.18: Εφαρμογή max-pooling στην έξοδο της 1Δ συνέλιξης.

Οι μέγιστες τιμές που παράγονται με την επεξεργασία καθενός από τα συνελικτικά διανύσματα χαρακτηριστικών θα ενωθούν και θα περάσουν σε ένα τελικό, πλήρως συνδεδεμένο επίπεδο. Αυτό το επίπεδο θα παράγει την τελική τιμή συναισθήματος, ταξινομώντας την κριτική ταινίας ως θετική ή αρνητική.

Παράδειγμα ταξινόμησης κειμένου με το ΣΝΔ

Το σχήμα 3.19 απεικονίζει μια πλήρη αρχιτεκτονική ΣΝΔ που χρησιμοποιείται για την ΑΣ κριτικών ταινιών [107]. Το πλήρες δίκτυο θα δει μια παρτίδα κριτικών ταινιών ως είσοδο. Αυτά περνούν από ένα προ-εκπαιδευμένο επίπεδο ενσωμάτωσης και, στη συνέχεια, οι ακολουθίες ενσωματώσεων λέξεων περνούν από διάφορες συνελικτικές λειτουργίες. Κάθε μεμονωμένη συνέλιξη πυρήνα εξάγει ένα διάνυσμα χαρακτηριστικών. Ένα ολόκληρο συνελικτικό επίπεδο εξάγει τα διανύσματα χαρακτηριστικών που παράγονται από κάθε πυρήνα (ίδιο ύψος) του επιπέδου. Οι μέγιστες τιμές (max-pooling) από τις εξόδους των διαφορετικών συνελικτικών επιπέδων (διαφορετικά ύψη πυρήνα) συνενώνονται και περνούν στο επίπεδο εξόδου που εξάγει την ιλάση συναισθήματος στην οποία ανήκει η εκάστοτε κριτική ταινίας.



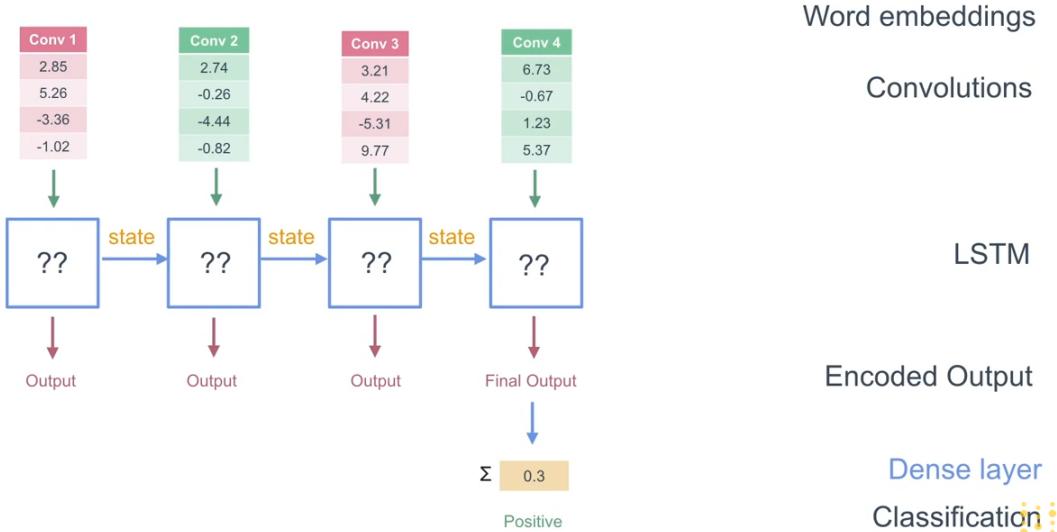
Σχήμα 3.19: Παράδειγμα ανάλυσης συναισθήματος (ταξινόμηση κειμένου) με χρήση ΣΝΔ.

Άλλες περιπτώσεις χρήσης της αρχιτεκτονικής του ΣΝΔ για ανάλυση συναισθήματος είναι [108], [109].

3.2.4.4 Υβριδικό ΣΝΔ+ΔΜΒΜ

Όπως εξηγήσαμε στο 2.3.3.2, το ΣΝΔ και το ΔΜΒΜ μπορούν να συνδυαστούν σε ένα ενιαίο μοντέλο που εκμεταλλεύεται τις καλές ιδιότητές τους ενώ αντισταθμίζει τις ελλείψεις του καθενός [110]. Αυτό σημαίνει ότι το ΣΝΔ θα ανιχνεύσει τοπικά χαρακτηριστικά (ομαδοποιήσεις λέξεων) και το ΔΜΒΜ θα μοντελοποιήσει χρονικές σχέσεις μεγάλων αποστάσεων στην ακολουθία εισόδου των ενσωματώσεων. Το σχήμα 3.20 δείχνει την πιο κοινή αρχιτεκτονική αυτού του τύπου μοντέλων όπου το ΣΝΔ προηγείται του ΔΜΒΜ. Το ΣΝΔ χρησιμοποιείται πρώτα για να μειώσει τη διάσταση της εισόδου παράγοντας τα διανύσματα χαρακτηριστικών. Αυτά τα διανύσματα εισάγονται στο ΔΜΒΜ. Επομένως, το ΔΜΒΜ θα πρέπει να επεξεργαστεί μικρότερες εισόδους από τις πλήρεις αναπαραστάσεις εισόδου, καθιστώντας

την εκπαίδευσή του πολύ πιο γρήγορη και αποτελεσματική. Όπως πάντα, απαιτείται ένα πλήρως συνδεδεμένο επίπεδο εξόδου στο τέλος του μοντέλου για το αποτέλεσμα ταξινόμησης.

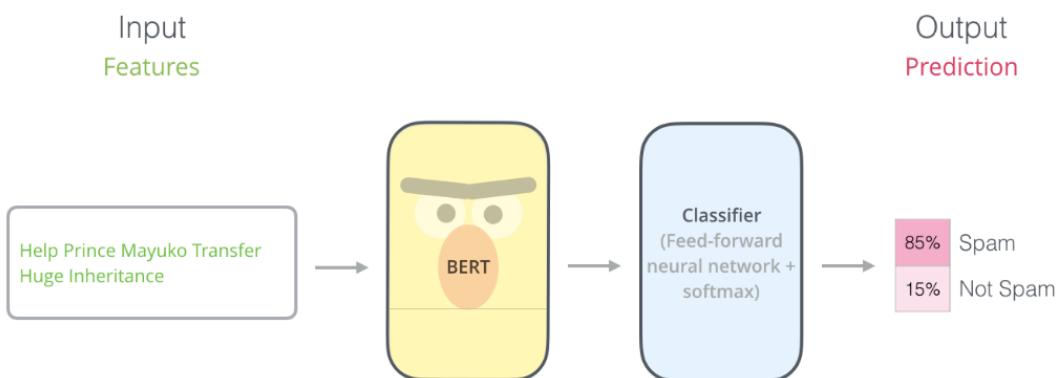


Σχήμα 3.20: Παράδειγμα ανάλυσης συναισθήματος χρησιμοποιώντας αρχιτεκτονική ΣΝΔ+ΔΜΒΜ.

Άλλες περιπτώσεις χρήσης της υβριδικής αρχιτεκτονικής για ανάλυση συναισθήματος είναι [111], [112], [113], [114], [115].

3.2.4.5 Ταξινόμηση κειμένου με BERT

Εξηγήσαμε νωρίτερα (σχήμα ταξινόμησης 3.5) ότι μπορεί να χρησιμοποιηθεί ένα μοντέλο γλώσσας αντί για το επίπεδο ενσωμάτωσης για την παραγωγή των ενσωματώσεων λέξεων από τις αρχικές συμβολοποιημένες ακολουθίες. Το BERT (δείτε 2.4.2.2) είναι ένα μοντέλο γλώσσας αιχμής που οδηγεί σε κορυφαία αποτελέσματα λόγω των υψηλής ποιότητας πλαισιομένων ενσωματώσεων που παράγει (δείτε 2.4.2.2).



Σχήμα 3.21: Παράδειγμα ταξινόμησης κειμένου χρησιμοποιώντας το προεκπαιδευμένο BERT.

Για εργασίες ταξινόμησης φορτώνουμε το προεκπαιδευμένο BERT και προσθέτουμε έναν ταξινομητή από πάνω του όπως φαίνεται στο σχήμα 3.21. Αυτό το συγκεκριμένο παράδειγμα [86] αφορά τον Εντοπισμό Ανεπιδύμητων Μηνυμάτων και μπορεί να χρησιμοποιηθεί για την ΑΣ αλλάζοντας απλώς τις κλάσεις σε θετικό/αρνητικό. Ο ταξινομητής μπορεί να είναι ένα ΣΝΔ, ΔΜΒΜ, MLP ή μόνο ένα πλήρως συνδεδεμένο επίπεδο. Το BERT τροφοδοτεί τον ταξινομητή με τις ενσωματώσεις λέξεων για να παράξει αυτός το τελικό αποτέλεσμα. Δεδομένου ότι το BERT είναι προεκπαιδευμένο, χρειάζεται απλώς να εκπαιδεύσουμε τον ταξινομητή για την εργασία ταξινόμησης χρησιμοποιώντας το επισημασμένο σύνολο δεδομένων μας. Μια άλλη επιλογή θα ήταν να ρυθμίσουμε το BERT μαζί με τον ταξινομητή για τη συγκεκριμένη εργασία για τη βελτίωση της συνολικής απόδοσης. Ωστόσο, αυτό είναι υπολογιστικά κοστοβόρο και συνήθως αποφεύγεται, καθώς το όφελος είναι συχνά πολύ μικρό ή μηδενικό.

Κεφάλαιο 4

Παρακολούθηση Κοινής Γνώμης μέσω Συλλογικής Σημασιολογικής Ανάλυσης Tweet

4.1 Περίληψη

Η υψηλή δημοτικότητα του twitter το καθιστά εξαιρετικό εργαλείο για πολιτική έρευνα, ενώ η εξόρυξη γνώμης μέσω σημασιολογικής ανάλυσης μεμονωμένων tweets έχει αποδειχθεί πολύτιμη. Ωστόσο, η αξιοποίηση των σχετικών επιστημονικών προόδων για συλλογική ανάλυση των μηνυμάτων twitter προκειμένου να ποσοτικοποιηθεί η κοινή γνώμη δεν έχει διερευνηθεί. Αυτή η μελέτη παρουσιάζει έναν τέτοιο νέο, αυτοματοποιημένο μηχανισμό παρακολούθησης της κοινής γνώμης, που αποτελείται από έναν σημασιολογικό περιγραφέα που βασίζεται σε αλγόριθμους ΕΦΓ. Ένας τετραδιάστατος περιγραφέας εξάγεται πρώτα για κάθε tweet ανεξάρτητα, ποσοτικοποιώντας την πολικότητα, την προσβλητικότητα, την μεροληψία και την μεταφορικότητα του λόγου. Στη συνέχεια, συναθροίζεται για πολλά tweets, σύμφωνα με μια επιδυμητή στρατηγική συνάθροισης και στόχο συνάθροισης. Αυτό μπορεί στη συνέχεια να αξιοποιηθεί με διάφορους τρόπους, όπως η εκπαίδευση μοντέλων μηχανικής μάθησης για την καθημερινή πρόβλεψη της κοινής γνώμης. Ο προτεινόμενος μηχανισμός εφαρμόζεται στα σύνολα δεδομένων tweet των προεδρικών εκλογών των ΗΠΑ του 2016 και 2020 και οι συνοπτικές περιγραφές της κοινής γνώμης που προκύπτουν διερευνώνται σε βάθος.

4.2 Εισαγωγή

Τα μέσα κοινωνικής δικτύωσης έχουν αναδειχθεί σταδιακά σε κεντρικά στοιχεία της σύγχρονης ζωής στον δυτικό κόσμο. Η εύκολη πρόσβαση σε διαδικτυακές πλατφόρμες και τα οφέλη της συνεχούς κοινωνικής αλληλεπίδρασης διατηρούν τον αριθμό των χρηστών σταθερά αυξανόμενο. Επιτρέπουν στους ανθρώπους να εκφράσουν άμεσα τις απόψεις και τα συναισθήματά τους χρησιμοποιώντας μια ποικιλία μέσων, όπως κείμενο, εικόνες, βίντεο κ.λπ.

Κατά συνέπεια, αυτές οι πλατφόρμες μπορούν να χρησιμοποιηθούν για την παρακολούθηση της κοινής γνώμης που σχετίζεται με ένα θέμα ιδιαίτερου ενδιαφέροντος. Η κοινή γνώμη “αντιπροσωπεύει τις απόψεις και τις επιθυμίες της πλειοψηφίας ενός πληθυσμού σχετικά με ένα συγκεκριμένο θέμα, είτε πολιτικό, εμπορικό, κοινωνικό ή άλλο” [1]. Το twitter φαίνεται να είναι το μέσο επιλογής για τη διατύπωση απόψεων σχετικά με κοινωνικοπολιτικά ζητήματα όπως ο COVID-19, οι εκλογές, ο ρατσισμός κ.λπ. Ο τεράστιος αριθμός ανθρώπων που χρησιμοποιούν το twitter για να ενημερώνονται και να εκφράζουν τις απόψεις τους έχει προσφέρει στους πολιτικούς την ευκαιρία να μεταδώσουν το μήνυμά τους γρήγορα και φθηνά, χωρίς να περάσουν στις παραδοσιακές τεχνικές μέσων ενημέρωσης και συνεντεύξεις τύπου [2].

Οι δυνατότητες του twitter σχετικά με πολιτικά γεγονότα επισημάνθηκαν για πρώτη φορά κατά τις προεδρικές εκλογές των ΗΠΑ του 2008, όπου ο Μπαράκ Ομπάμα χρησιμοποίησε την πλατφόρμα αποτελεσματικά για την εκστρατεία του [116]. Μετά από αυτή την επιτυχημένη καμπάνια στο twitter, όλοι οι μεγάλοι υποψήφιοι και τα πολιτικά κόμματα έκαναν γρήγορα την παρουσία τους αισθητή στα μέσα κοινωνικής δικτύωσης. Επιπλέον, η δημοτικότητα του twitter παρέχει μια μοναδική ευκαιρία για αποτελεσματική ηλεκτρονική διαχυτήση, ιδίως όσον αφορά στην απλοποιημένη επικοινωνία μεταξύ κυβερνητικών θεσμών και πολιτών [117]. Αυτό μπορεί να επιτρέψει μεγαλύτερη διαφάνεια και αυξημένη εμπιστοσύνη των πολιτών στους τοπικούς θεσμούς.

Δεδομένης αυτής της πολύ μεγάλης προοπτικής του twitter, η εξόρυξη γνώμης μέσω tweets μπορεί να μας δώσει πολύτιμες πληροφορίες. Τα αυτοματοποιημένα εργαλεία σημασιολογικής ανάλυσης κειμένου, που βασίζονται σε σύγχρονους αλγόριθμους ΕΦΓ που βασίζονται στην Τεχνητή Νοημοσύνη, μπορούν να προσδιορίσουν τη διάθεση ενός tweet (π.χ. πολικότητα/συναίσθημα [3]) με αξιοσημείωτη ακρίβεια. Τα βαθιά νευρωνικά δίκτυα (BNΔ) έχουν εξελίξει πολύ τις σύγχρονες μεθόδους αντιμετώπισης τέτοιων προβλημάτων, ειδικά στην ΑΣ.

Έτσι, για παράδειγμα, η σύγχρονη τεχνητή νοημοσύνη καθιστά δυνατό να αποκτήσουμε μια αίσθηση για το ποιος υποψήφιος είναι πιο πιθανό να κερδίσει τις επόμενες εκλογές, αναλύοντας το συναίσθημα των σχετικών tweets [118]. Επιπλέον, η εξόρυξη γνώμης μπορεί να χρησιμοποιηθεί για τον προσδιορισμό των απόψεων του κοινού σχετικά με ένα κρίσιμο

θέμα, π.χ. ένα δημοψήφισμα [119], και να βοηθήσει την κυβέρνηση να λάβει τις σωστές αποφάσεις.

Ωστόσο, η συλλογική και πολυδιάστατη σημασιολογική ανάλυση των tweets, βασισμένη σε BNΔ τελευταίας τεχνολογίας, προκειμένου να ποσοτικοποιηθεί και να παρακολουθείται η κοινή γνώμη δεν έχει διερευνηθεί σημαντικά, παρά τις προφανείς δυνατότητες. Οι περισσότερες σχετικές μέθοδοι εξάγουν μόνο περιορισμένες ποσότητες σημασιολογικού περιεχομένου (συνήθως πολικότητα), αντί για πολλαπλές σημασιολογικές διαστάσεις, ενώ τα tweets επεξεργάζονται μεμονωμένα. Τα αποτελέσματα απλώς συνοψίζονται για χειροκίνητη ανθρώπινη επισκόπηση. Η αυτοματοποιημένη συλλογική ανάλυση περιεχομένου που εξαγεται από τα μέσα κοινωνικής δικτύωσης από πολλαπλές σημασιολογικές πτυχές, προκειμένου να εντοπιστούν τάσεις στη συνολική κοινή γνώμη στο σύνολό της, είναι μάλλον σπάνια.

Αυτή η μελέτη επιχειρεί να διερευνήσει τις σχετικές ανεξερεύνητες δυνατότητες πολυδιάστατης συλλογικής ανάλυσης tweet μέσω υπερσύγχρονων BNΔ. Έτσι, προτείνεται ένας νέος, αυτοματοποιημένος μηχανισμός παρακολούθησης της κοινής γνώμης, που αποτελείται από έναν σύνθετο, ποσοτικό, σημασιολογικό περιγραφέα που βασίζεται σε αλγόριθμους ΕΦΓ με δυνατότητα BNΔ. Ένα τετραδιάστατο διάνυσμα, δηλαδή, ένα παράδειγμα του προτεινόμενου περιγραφέα, εξαγεται πρώτα για κάθε tweet ανεξάρτητα, ποσοτικοποιώντας την πολικότητα κειμένου, προσβλητικότητα, μεροληψία και μεταφορικότητα. Στη συνέχεια, οι υπολογισμένοι περιγραφείς πολλών tweets συναθροίζονται σύμφωνα με μια επιθυμητή στρατηγική συνάθροισης (π.χ. αριθμητικός μέσος όρος) και στόχο συνάθροισης (π.χ. μια συγκεκριμένη χρονική περίοδο). Αυτό μπορεί να αξιοποιηθεί με διάφορους τρόπους. για παράδειγμα, η συνάθροιση των tweet κάθε ημέρας ξεχωριστά μας επιτρέπει να κατασκευάσουμε μια πολυμεταβλητή χρονοσειρά που μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός αλγόριθμου τεχνητής νοημοσύνης, για καθημερινές προβλέψεις της κοινής γνώμης [120]. Προκειμένου να αξιολογηθεί η χρησιμότητα του προτεινόμενου μηχανισμού, έχει εφαρμοστεί στο σύνολο tweet για τις προεδρικές εκλογές των ΗΠΑ το 2016 και 2020. Οι περιγραφείς της κοινής γνώμης που προκύπτουν διερευνώνται σε βάθος.

Το υπόλοιπο αυτής της μελέτης οργανώνεται με τον ακόλουθο τρόπο. Η ενότητα 4.3 συζητά σχετική προηγούμενη βιβλιογραφία, εστιάζοντας όχι σε μεθόδους πρόβλεψης ΕΦΓ ή χρονοσειρών (τις οποίες εκμεταλλευόμαστε με τρόπο μαύρου κουτιού) αλλά σε διάφορους υπάρχοντες μηχανισμούς εξαγωγής και παρακολούθησης της κοινής γνώμης με εφαρμογή ΕΦΓ και/ ή πρόβλεψη χρονοσειρών σε αναρτήσεις στα μέσα κοινωνικής δικτύωσης. Η ενότητα 4.4 παρουσιάζει την προτεινόμενη σημασιολογική περιγραφή της κοινής γνώμης. Η ενότητα 4.5 εξετάζει την πειραματική αξιολόγηση των συνόλων tweet των Προεδρικών Εκλογών των ΗΠΑ 2016/2020 και τις συνοπτικές περιγραφές της κοινής γνώμης που προέκυψαν χρησιμοποιώντας τον προτεινόμενο μηχανισμό. Στη συνέχεια, η ενότητα 4.6 συζητά αξιοσημείωτα ευρήματα σχετικά με τα χρησιμοποιούμενα σύνολα δεδομένων που εξήχθησαν

χρησιμοποιώντας τον προτεινόμενο μηχανισμό, καθώς και τις κύριες νέες συνεισφορές του μηχανισμού. Τέλος, η ενότητα 4.7 εξάγει συμπεράσματα από την προηγούμενη παρουσίαση.

4.3 Σχετική Βιβλιογραφία

Αυτή η ενότητα παρουσιάζει μια σύντομη επισκόπηση των υφιστάμενων προσεγγίσεων που βασίζονται σε μοντέλα τεχνητής νοημοσύνης (ΕΦΓ και/ή BNΔ) για τη σημασιολογική ανάλυση του κειμένου των μέσων κοινωνικής δικτύωσης για: α) ποσοτική περιγραφή της κοινής γνώμης και β) πρόβλεψη χρονοσειρών.

4.3.1 Περιγραφή Κοινής Γνώμης

Υπάρχει αυξανόμενο επιστημονικό ενδιαφέρον για την ανάλυση των αναρτήσεων στα μέσα κοινωνικής δικτύωσης από τις αρχές της δεκαετίας του 2010, με το twitter να κυριαρχεί στη σχετική έρευνα. Για παράδειγμα, το [121] αξιολόγησε την κοινή γνώμη σχετικά με την πολιτική διαφήμιση στο Facebook, εξάγοντας αυτόματα θέματα συζήτησης από σχετικά tweets που δημοσιεύτηκαν τον Οκτώβριο του 2019. Πραγματοποιήθηκε μη αυτόματη επιθεώρηση αυτών των θεμάτων, οδηγώντας στο συμπέρασμα ότι η άποψη των χρηστών για τη διαφήμιση στο Facebook χειροτερεύει σταδιακά, ως αποτέλεσμα ανησυχιών για το απόρρητο που σχετίζονται με την εμπιστοσύνη στην πλατφόρμα. Ομοίως, έχοντας ως στόχο την ανάλυση των tweets από την οπτική γωνία των υποψηφίων, οι [122] και [123] επεξεργάστηκαν το περιεχόμενο των tweets του Ντόναλντ Τράμπ και της Χίλαρι Κλίντον κατά τις προεδρικές εκλογές των ΗΠΑ το 2016. Στόχος ήταν να εντοπιστεί ποιοτικά σε ποιο θέμα έδωσε έμφαση ο κάθε υποψήφιος και ποιες επικοινωνιακές στρατηγικές χρησιμοποίησε.

Ωστόσο, αυτή η εργασία δεν αφορά τη μοντελοποίηση θεμάτων και τη χειροκίνητη επιθεώρηση προσδιοριζόμενων θεμάτων. Αντίθετα, σχετίζεται με μεθόδους που εκμεταλλεύονται χαρακτηριστικά σημασιολογικού περιεχομένου των tweets για να ποσοτικοποιήσουν την κοινή γνώμη με αυτόματο τρόπο. Αυτές οι μέθοδοι μπορούν να κατηγοριοποιηθούν ευρέως σε: α) μη σημασιολογικές, οι οποίες δεν εκτελούν σημασιολογική ανάλυση σε tweet βασισμένη σε τεχνητή νοημοσύνη και β) σημασιολογικές, που συνήθως εκτελούν έναν τύπο ανάλυσης συναισθήματος/εξόρυξης απόψεων με δυνατότητα τεχνητής νοημοσύνης σε tweet.

4.3.1.1 Μη σημασιολογικές μέθοδοι

Οι μη σημασιολογικές μέθοδοι λαμβάνουν υπόψη μόνο τις συχνότητες των λέξεων-κλειδιών ή/και τον όγκο των tweet, με αποτέλεσμα μάλλον ανακριβείς ή/και καθαρά ποιοτικές πληροφορίες. Για παράδειγμα, το [124] πραγματοποίησε μια στατιστική ανάλυση σε tweets

από την ισπανική προεδρική εκστρατεία του 2019, τα οποία επιλέχθηκαν με βάση τις λέξεις-κλειδιά και ποσοτικοποιήθηκαν μέσω του όγκου τους με την πάροδο του χρόνου. Στόχος ήταν να αποκαλυφθεί ο πολιτικός λόγος των εμπλεκόμενων κομμάτων και να επισημανθούν τα κύρια μηνύματα που μεταφέρθηκαν και ο αντίκτυπός τους στο μερίδιο της φωνής των υποψηφίων. Οι ταξινομητές μηχανικής μάθησης χρησιμοποιήθηκαν μόνο για τον εντοπισμό ανεπιθύμητων μηνυμάτων και τα συμπεράσματα ήταν καθαρά ποιοτικά. Ομοίως, το [125] ερεύνησε τη δραστηριότητα στα μέσα κοινωνικής δικτύωσης κατά τη διάρκεια μιας πολιτικής εκδήλωσης, αναλύοντας τη συζήτηση του GOP των ΗΠΑ το 2016 και παρατηρώντας τον όγκο των ειδικών λέξεων-κλειδιών σε αναρτήσεις στο twitter.

Προφανώς, ο μηχανισμός που προτείνεται σε αυτή την εργασία είναι εντελώς διαφορετικός στη φύση από αυτές τις μη σημασιολογικές προσεγγίσεις: εκχωρεί σε κάθε tweet ένα 4Δ σημασιολογικό αριθμητικό διάνυσμα που αποκτάται με ταξινομητές κειμένου που βασίζονται σε BNΔ/ΕΦΓ, προχωρώντας έτσι ένα βήμα παραπέρα από τις συμβατικές στατιστικές προσεγγίσεις και στη συνέχεια συγκεντρώνει αυτά τα αποτελέσματα για όλα τα tweets προκειμένου να δημιουργήσει μια συνολική, ποσοτική περιγραφή της κοινής γνώμης.

4.3.1.2 Σημασιολογικές μέθοδοι χωρίς Συνάθροιση

Οι σημασιολογικές μέθοδοι είναι πιο προηγμένες και παρέχουν πιο ακριβή αποτελέσματα. Η πλειοψηφία τέτοιων μεθόδων λειτουργούν στο twitter, αλλά δεν αντιμετωπίζουν τα σχετικά tweets συλλογικά και συνήθως λαμβάνουν υπόψη τον όγκο των tweets ανά κατηγορία συναισθήματος. Για παράδειγμα, το [1] παρουσίασε ένα πλαίσιο για την παρακολούθηση της εξέλιξης (16 μήνες) της κοινής γνώμης σχετικά με το θέμα της κλιματικής αλλαγής. Ο προτεινόμενος μηχανισμός είναι σε θέση να εντοπίζει και να παρακολουθεί on-the-fly το συναίσθημα σε ένα επιθυμητό σύνολο μεμονωμένων tweets, πιθανώς κατά τη δημοσίευσή τους, αλλά πραγματοποιείται μόνο στοιχειώδης ανάλυση σε αυτά τα αποτελέσματα ως συλλογή. Έτσι, η κοινή γνώμη στο σύνολό της ελάχιστα εξετάζεται. Έχοντας κατά νου έναν διαφορετικό στόχο, η [7] διεξήγαγε ΑΣ των tweets για την πρόβλεψη του εκλογικού αποτελέσματος. Χρησιμοποιήθηκε ένα απλό μοντέλο ΣΝΔ για αυτήν την εργασία, ενώ ο συνολικός όγκος των tweets (μη σημασιολογική ιδιότητα) συγκρίθηκε με την πολικότητα του συναισθήματος (όγκος θετικών tweets) για να διαπιστωθεί ποιος είναι καλύτερος προγνωστικός δείκτης των εκλογικών αποτελεσμάτων. Ο σημασιολογικός περιγραφέας βρέθηκε να είναι πιο ακριβής. Προχωρώντας σε αυτήν τη γραμμή έρευνας, το [8] είχε ως στόχο να προβλέψει όχι μόνο τον νικητή αλλά και το ποσοστό ψήφου κάθε υποψηφίου στις ισπανικές προεδρικές εκλογές του 2019, λαμβάνοντας υπόψη τον όγκο των θετικών tweets ανά υποψήφιο. Ωστόσο, τα ίδια τα σημασιολογικά χαρακτηριστικά δεν συναθροίστηκαν στο σύνολο όλων των tweets.

Με παρόμοιο τρόπο, το [10] πραγματοποίησε εξόρυξη γνώμης για τα ιταλικά tweets σχετικά με τον εμβολιασμό για μια περίοδο 12 μηνών και απλώς μέτρησε τον αριθμό των tweets ανά κατηγορία πολικότητας (άρνητικό, 'θετικό' ή όυδέτερο') για κάθε μήνα. Διαπιστώθηκε ότι τα συμβάντα που σχετίζονται με το εμβόλιο επηρέασαν την κατανομή των κατηγοριών πολικότητας, αλλά δεν πραγματοποιήθηκε συνάθροιση του σημασιολογικού περιεχομένου. Σύμφωνα με μια σχεδόν πανομοιότυπη γενική ιδέα, το [13] είχε ως στόχο να προσδιορίσει το ακρίσιμο χρονικό παράθυρο της κοινής γνώμης σχετικά με ένα συμβάν, εφαρμόζοντας ταξινόμηση πολυσυναισθημάτων σε αναρτήσεις στο Sina Weibo (δημοσιεύτηκαν σε σύντομο χρονικό διάστημα περίπου 10 ημερών μετά από ορισμένα γεγονότα). Ο όγκος των tweets ανά κλάση (μεταξύ των 7 κλάσεων συναισθημάτων που χρησιμοποιήθηκαν) εξετάσθηκε για να διαπιστωθεί ότι η παρακολούθηση της τάσης των αρνητικών συναισθημάτων είναι ζωτικής σημασίας για την πρόβλεψη της έκβασης των γεγονότων.

4.3.1.3 Σημασιολογικές μέθοδοι με Συνάθροιση

Σε αντίθεση με αυτές τις προσεγγίσεις, ένα σύνολο πιο προηγμένων σημασιολογικών μεθόδων πραγματοποιεί συνάθροιση του σημασιολογικού περιεχομένου και συνεπώς αντιμετωπίζει τα tweets συλλογικά, κατασκευάζοντας έναν σημασιολογικό περιγραφέα της κοινής γνώμης με τη μορφή μιας χρονοσειράς χαμηλής διάστασης. Αυτή ακριβώς είναι η οικογένεια μεθόδων στην οποία ανήκει ο μηχανισμός που προτείνεται σε αυτή τη μελέτη. Για παράδειγμα, το [12] διερεύνησε την αλλαγή του κοινού συναισθήματος στην Κίνα μετά τη "σύγκρουση τρένου Wenzhou", πραγματοποιώντας ανάλυση συναισθήματος σε αναρτήσεις από την πλατφόρμα Sina Weibo, συγκεντρώνοντας οκτώ αναγνωρισμένα συναισθήματα ανά tweet με την πάροδο του χρόνου για να παράγουν ένα 8Δ διάνυσμα για κάθε ημέρα, που παρακολουθείται ως 8 ξεχωριστές χρονοσειρές για ένα διάστημα 10 ημερών. Ωστόσο, η ακρίβεια της ανάλυσης συναισθήματος εκείνη τη στιγμή δεν ήταν υψηλή και τα αποτελέσματα δεν ήταν ιδιαίτερα χρήσιμα. Ομοίως, στο [14], tweets σχετικά με τις πολιτειακές εκλογές του κυβερνήτη στη Νιγηρία το 2017 αναλύονται σημασιολογικά και τα αποτελέσματα συναθροίζονται για όλες τις αναρτήσεις ανά διάστημα δύο ωρών. Οι χρονοσειρές που παράγονται καλύπτουν ένα χρονικό πλαίσιο 18 ωρών την ημέρα των εκλογών.

Λειτουργώντας επίσης προς αυτήν την κατεύθυνση, ο [9] χρησιμοποίησε σημασιολογική ανάλυση και συνάθροιση tweet για να δημιουργήσει μια χρονοσειρά μέσων ημερήσιων συναισθημάτων για κάθε κόμμα, καλύπτοντας μια προεκλογική περίοδο 21 ημερών κατά τις προεδρικές εκλογές των ΗΠΑ του 2016. Τέλος, στο [11], παρόμοιες ιδέες εφαρμόστηκαν στην πρόβλεψη των τιμών κρυπτονομισμάτων μέσω συλλογικής σημασιολογικής ανάλυσης των tweets. Οι σχετικές χρονοσειρές κατασκευάστηκαν μέσω της ημερήσιας συνάθροισης μεμονωμένων σημασιολογικών αποτελεσμάτων tweet επανέχημένης με οικονομικά δεδομένα, που καλύπτουν περίοδο 2 μηνών και εκπαιδεύτηκε ένα μοντέλο μηχανικής μάθησης για την

πρόβλεψη χρονοσειρών. Το δημόσιο συναίσθημα που προέρχεται από το twitter βρέθηκε ότι έχει πράγματι προγνωστική δύναμη, αλλά δεν ήταν αρκετό από μόνο του για ακριβείς προβλέψεις. Συνολικά, οι μέθοδοι αυτού του τύπου είναι περισσότερο παρόμοιες με τη δική μας, αλλά η κλίμακα της πειραματικής αξιολόγησης (π.χ., η χρονική διάρκεια των κατασκευασμένων χρονοσειρών) είναι σημαντικά περιορισμένη σε σύγκριση με αυτήν τη μελέτη. Άλλες σημαντικές διαφορές συζητούνται στην επόμενη υποενότητα.

4.3.1.4 Διαστάσεις και Αλγόριθμοι Σημασιολογικής Ανάλυσης

Η συντριπτική πλειονότητα των σημασιολογιών μεθόδων που παρουσιάζονται παραπάνω χρησιμοποιεί μόνο το συναίσθημα του tweet, επομένως μπορούν να θεωρηθούν ότι εκμεταλλεύονται μονοδιάστατη σημασιολογία κειμένου. Αυτό είναι πιο προφανές σε περιπτώσεις όπου η σημασιολογική ανάλυση εξάγει μια πολικότητα (π.χ. δυαδική ή τριαδική ταξινόμηση σε θετικά/αρνητικά tweets ή σε θετικά/ουδέτερα/αρνητικά). Αυτή η περιορισμένη προσέγγιση είναι η πιο κυρίαρχη [7] [8] [9] [10] [11]. Ωστόσο, η μονοδιάστατη ανάλυση συναισθημάτων μπορεί να θεωρηθεί ότι ισχύει ακόμη και όταν χρησιμοποιούνται πολυσυναισθηματικοί ταξινομητές αντί της απλής πολικότητας. Π.χ., στο [12] (προσδοκία, χαρά, αγάπη, έκπληξη, άγχος, λύπη, θυμός και μίσος), [13] (ευτυχία, αρέσκεια, λύπη, αηδία, έκπληξη, θυμός και φόβος) και [1] (χαρά, έμπνευση, θυμός, ρατσισμός, υποστήριξη). Αν και είναι μια πιο διαφοροποιημένη προσέγγιση, αυτά τα συναισθήματα εξακολουθούν να εμπίπτουν στη γενική ομπρέλα του συναισθήματος, επομένως αυτές οι μέθοδοι συνεχίζουν να αγνοούν άλλα χαρακτηριστικά σημασιολογικού κειμένου. Η μόνη περίπτωση όπου λαμβάνεται υπόψη η πολυδιάστατη σημασιολογία είναι η [14], όπου λαμβάνονται υπόψη 2 διαφορετικές διαστάσεις γνώμης (πολικότητα και προκατάληψη). Αντίθετα, αυτή η μελέτη προτείνει έναν 4-διάστατο περιγραφέα που εξετάζει από κοινού την πολικότητα, την μεροληψία, την μεταφορικότητα και την προσβλητικότητα, τα οποία είναι όλα διαφορετικά χαρακτηριστικά κειμένου και επαληθεύει πειραματικά τη χρησιμότητά τους.

Μια άλλη σχετική πτυχή είναι ο τρόπος με τον οποίο εκτελείται η σημασιολογική ανάλυση tweet σε τέτοιες δημοσιευμένες μεθόδους. Η συντριπτική πλειονότητά τους χρησιμοποιεί απαρχαιωμένους αλγόριθμους για την περιγραφή κειμένου, βασιζόμενοι σε λεξικά ή παλαιότερα είδη αναπαραστάσεων. Π.χ., το [12] χρησιμοποιεί το λεξικό HowNet [126], το [13] χρησιμοποιεί ένα συναισθηματικό λεξικό [127] και ένα λεξικό αρνητικών λέξεων, το [9] χρησιμοποιεί το SentiStrength [128], Το [10] χρησιμοποιεί μια προσέγγιση Bag-of-Words, κλπ. Μόνο λίγοι βασίζονται σε σύγχρονα σχήματα αναπαράστασης λέξεων που βασίζονται σε BNΔ, όπως τα [13] και [1] που εκμεταλλεύονται το Word2Vec [129]. Επιπλέον, η πλειοψηφία των μεθόδων μηχανικής μάθησης που χρησιμοποιούνται για τη σημασιολογική ανάλυση αποτελούν ξεπερασμένες προσεγγίσεις, όπως [13], [10] ή [14], οι οποίες χρησιμοποιούν

έναν απλό ταξινομητή K-Κοντινότερων Γειτόνων (KKT), Μηχανές Διανυσμάτων Υποστήριξης (ΜΔΥ) ή τον Αφελή ταξινομητή Bayes¹ του Textblob, αντίστοιχα. Το [8] αξιολογεί μια ποικιλία από παραδοσιακούς (μη νευρωνικούς) αλγόριθμους μηχανικής μάθησης, ενώ ο [11] εκμεταλλεύεται ένα σύνολο κανόνων ανάλυσης συναισθήματος [130]. Πρόσφατα μοντέλα μηχανικής μάθησης που βασίζονται σε BNΔ αξιοποιήθηκαν μόνο στο [1] (ΑΜ-ΔΜΒΜ) και στο [7] (ΣΝΔ).

Σε αντίθεση με όλες τις παραπάνω προσεγγίσεις, ο μηχανισμός που προτείνεται σε αυτή τη μελέτη βασίζεται από άκρο σε άκρο σε λύσεις BNΔ τελευταίας τεχνολογίας, τόσο για αναπαράσταση λέξεων όσο και για σημασιολογική ανάλυση.

4.3.2 Πρόβλεψη Χρονοσειρών

Η πρόβλεψη χρονοσειρών που προέκυψε από την ανάλυση συναισθήματος των tweets είχε χρησιμοποιηθεί κυρίως στο παρελθόν για την πρόβλεψη μελλοντικών χρηματοοικονομικών δεικτών. Έτσι, το [131] διερεύνησε την επίδραση διαφορετικών σημαντικών γεγονότων που συνέβησαν κατά την περίοδο 2012–2016 στις χρηματιστηριακές αγορές. Παρόμοια προσέγγιση ακολούθηκε στο [132]. Το [133] εξέτασε τη χρήση τιμών πολικότητας, που προέρχονται από tweets σχετικά με την εξωτερική πολιτική των Ηνωμένων Πολιτειών και τις εταιρείες πετρελαίου, προκειμένου να προβλέψει την κατεύθυνση των εβδομαδιαίων τιμών του αργού πετρελαίου WTI. Τέλος, το [134] διερεύνησε εάν ένας δείκτης δημόσιου συναισθήματος που εξάγεται από καθημερινά tweets στο χρηματιστήριο ή στο box office ταινιών μπορεί πράγματι να βελτιώσει την πρόβλεψη μιας σχετικής χρονοσειράς.

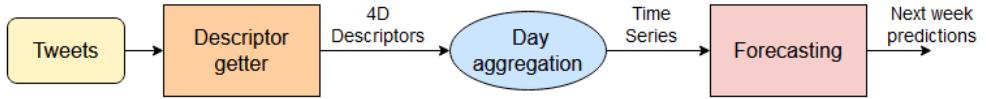
Σε όλες αυτές τις περιπτώσεις, η μόνη αξιοποιημένη σημασιολογική διάσταση ήταν η πολικότητα. Αποδείχθηκε ότι η ακρίβεια της πρόβλεψης βελτιώνεται χρησιμοποιώντας την πολικότητα, αλλά η κοινή γνώμη που εξάγεται μέσω tweets χρησιμοποιήθηκε μόνο ως βοηθητική πηγή πληροφοριών για μια εργασία χρηματοοικονομικού τομέα, συμπληρώνοντας μια κύρια πηγή (π.χ. δεδομένα χρηματιστηρίου στο [131]). Επιπλέον, χρησιμοποιήθηκαν απαρχαιωμένοι αλγόριθμοι αναπαράστασης λέξεων και εξόρυξης γνώμης σε όλες αυτές τις μελέτες: [131] χρησιμοποίησε λεξιά SentiWordNet², [133] εκμεταλλεύτηκε το SentiStrength και το Stanford NLP Sentiment Analyzer³, ενώ το [134] χρησιμοποίησε έναν Αφελή ταξινομητή Bayes.

Ομοίως, παλαιότεροι αλγόριθμοι χρησιμοποιήθηκαν κυρίως για την ίδια την εργασία πρόβλεψης χρονοσειρών. Το [131] εκμεταλλεύτηκε τη Γραμμική Παλινδρόμηση και την Παλινδρόμηση Διανυσμάτων Υποστήριξης (ΠΔΥ), το [133] χρησιμοποίησε μοντέλα εκμάθησης ΜΔΥ,

¹<https://textblob.readthedocs.io/en/dev/>

²<https://github.com/aesuli/SentiWordNet>

³<https://nlp.stanford.edu/sentiment/>



Σχήμα 4.1: Ποσοτική πρόβλεψη κοινής γνώμης με χρήση του προτεινόμενου μηχανισμού/σημασιολογικού περιγραφέα.

Αφελή Bayes και Perceptron Multi-Layer (MLP), ενώ το [134] βασίστηκε σε Γραμμική Παλινδρόμηση και MLP, ΜΔΥ για την πρόβλεψη της άμεσης τροχιάς της χρονοσειράς-στόχου, λαμβάνοντας υπόψη την πολικότητα ή τον όγκο των tweets ως χαρακτηριστικά εισόδου.

Σε αντίθεση με τις παραπάνω μεθόδους χρηματοοικονομικής πρόβλεψης αξιοποιώντας εκτιμήσεις πολικότητας που προέρχονται από tweet, αυτή η μελέτη εστιάζει στον ίδιο τον προτεινόμενο σημασιολογικό περιγραφέα της κοινής γνώμης και τις πιθανές χρήσεις του για πολιτική ανάλυση (συμπεριλαμβανομένης της πρόβλεψης της κοινής γνώμης). Αντός ο περιγραφέας αποτυπώνει συμπαγώς πολλαπλές απόψεις/σημασιολογικές διαστάσεις, αντί για απλή πολικότητα. Η πρόβλεψη χρονοσειρών χρησιμοποιείται ως παράδειγμα εφαρμογής, μεταξύ άλλων, και τα μοντέλα BNΔ τελευταίας τεχνολογίας αξιοποιούνται σε όλα τα στάδια, σε αντίθεση με προηγούμενες μεθόδους στη βιβλιογραφία.

4.4 Προτεινόμενος Μηχανισμός

Ο προτεινόμενος νέος αυτοματοποιημένος μηχανισμός παρακολούθησης της κοινής γνώμης αποτελείται από έναν σύνθετο, ποσοτικό, σημασιολογικό περιγραφέα που βασίζεται σε ταξινομητές που βασίζονται σε ΕΦΓ/BNΔ. Με τη χρησιμοποίησή τους, ένα τετραδιάστατο διάνυσμα, δηλαδή, ένα δείγμα του προτεινόμενου περιγραφέα, εξάγεται πρώτα για κάθε tweet ανεξάρτητα, ποσοτικοποιώντας έτσι την πολικότητα, προσβλητικότητα, μεροληψία και μεταφορικότητα. Στη συνέχεια, οι υπολογισμένοι περιγραφείς συναθροίζονται για πολλά tweets, σύμφωνα με μια επιθυμητή στρατηγική συνάθροισης (π.χ. αριθμητικός μέσος όρος) και στόχο συνάθροισης (π.χ. μια συγκεκριμένη χρονική περίοδο). Οι συνοπτικοί περιγραφείς μπορούν να αξιοποιηθούν με διάφορους τρόπους: για παράδειγμα, με τη συνάθροισή τους σε καθημερινή βάση μας επιτρέπει να κατασκευάσουμε μια πολυμεταβλητή χρονοσειρά που μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός BNΔ για την πρόβλεψη μελλοντικών περιγραφέων κοινής γνώμης. Για παράδειγμα, ο αγωγός μιας τέτοιας εφαρμογής πρόβλεψης κοινής γνώμης που χρησιμοποιεί τον προτεινόμενο μηχανισμό μας απεικονίζεται στο Σχήμα 4.1.

Παρακάτω, αναλύονται τα βήματα υπολογισμού του προτεινόμενου περιγραφέα κοινής γνώμης, μαζί με τον αλγορίθμικό μηχανισμό υλοποίησης της διαδικασίας.

4.4.1 Βήμα 1: Επιλέγοντας το επιθυμητό σύνολο tweets

To Twitter API επιτρέπει την εύκολη και αυτοματοποιημένη εξαγωγή tweets με βάση τα κριτήρια που ορίζονται με μη αυτόματο τρόπο σχετικά με το θέμα και την ημερομηνία τους. Για παράδειγμα, η παρουσία συγκεκριμένων λέξεων-κλειδιών ή/και hashtags, η χρονική σήμανση του tweet, το γεγονός ότι έχει δημοσιευτεί εντός ενός επιθυμητού εύρους ημερομηνιών, κ.λπ. Η παρακολούθηση της κοινής γνώμης σχετικά με ένα ζήτημα (π.χ. στάση απέναντι στους υποψηφίους κατά τη διάρκεια μιας εκτεταμένης προεκλογικής περιόδου) προφανώς απαιτεί έξυπνη προσαρμογή αυτών των κριτηρίων, έτσι ώστε να μπορεί να ληφθεί ένα πραγματικά αντιπροσωπευτικό σύνολο μηνυμάτων χρήστη. Ωστόσο, σε γενικές γραμμές, αυτό το στάδιο είναι αρκετά απλό και ξεκάθαρο, επομένως δεν θα επεκταθούμε περαιτέρω.

4.4.2 Βήμα 2: Εξαγωγή μεμονωμένου περιγραφέα ανά tweet

Το δεύτερο βήμα του προτεινόμενου μηχανισμού είναι να περιγράψει σημασιολογικά κάθε tweet από την επιλεγμένη ομάδα μηνυμάτων ως ένα 4-διάστατο (4Δ) διάνυσμα γνώμης με πραγματικές τιμές. Αυτό το διάνυσμα περιγραφής εξάγεται ξεχωριστά για κάθε tweet. Για το σκοπό αυτό χρησιμοποιείται ένα σύνολο 4 προεκπαιδευμένων μοντέλων BNΔ. Με βάση τα πιο σύγχρονα μοντέλα της ΕΦΓ, χρησιμοποιήθηκαν δύο διαφορετικές νευρωνικές αρχιτεκτονικές. Το υβριδικό ΣΝΔ-ΔΜΒΜ από το [135] εκπαιδεύτηκε χωριστά τρεις φορές από την αρχή, χρησιμοποιώντας τρία διαφορετικά δημόσια επισημασμένα σύνολα δεδομένων για την αναγνώριση προσβλητικότητας, μεροληψίας και μεταφορικότητας (σαρκασμός, ειρωνεία και/ή μεταφορά) σε tweets. Επιπλέον, ένα υπερσύγχρονο προεκπαιδευμένο, δημόσια διαθέσιμο νευρωνικό μοντέλο⁴ χρησιμοποιήθηκε για την αναγνώριση πολικότητας. Και στις τέσσερις περιπτώσεις, η επιθυμητή εργασία τέθηκε ως δυαδική ταξινόμηση κειμένου, με αντίστοιχες ετικέτες tweet (προσβλητικό/μη προσβλητικό, μεροληπτικό/αντικειμενικό, μεταφορικό/κυριολεκτικό, αρνητικό/θετικό). Είναι σημαντικό ότι η επιβολή της ταξινόμησης να είναι όσο το δυνατόν πιο απλή (η διάκριση μεταξύ δύο κατηγοριών είναι συνήθως ευκολότερη από τη διάκριση μεταξύ πολλαπλών αλάσεων) καθιστά τα χρησιμοποιούμενα BNΔ πιο ισχυρά, ακριβή και αξιόπιστα για το πραγματικό πρόβλημα που αντιμετωπίζεται από τον προτεινόμενο μηχανισμό, την παρακολούθηση της κοινής γνώμης. Όλοι οι εκπαιδευμένοι ταξινομητές εξάγουν μια πραγματική τιμή εντός του εύρους $[0, 1]$ για κάθε tweet, με την τιμή 0/1 να υποδηλώνει τέλεια και αδιαφορίσβήτητη εκχώρηση μιας από τις δύο αντίθετες ετικέτες (π.χ. 0/1 σημαίνει σίγουρα και πλήρως αρνητικό/θετικό συναίσθημα, αντίστοιχα, στην περίπτωση της πολικότητας).

Πρακτικές λεπτομέρειες σχετικά με την εκπαίδευση των τεσσάρων BNΔ που αποτελούν τη ραχοκοκαλιά του προτεινόμενου μηχανισμού ακολουθούν παρακάτω.

⁴<https://github.com/DheerajKumar97/US-2020-Election-Campaign-Youtube-Comments-Sentiment-Analysis>

4.4.2.1 Δεδομένα Εκπαίδευσης

SemEval-2019 Task 6 sub-task A (S19-T6) [136]: Αυτό το σύνολο δεδομένων περιέχει 14.100 tweets σχολιασμένα ως προσβλητικά/μη προσβλητικά που χρησιμοποιήθηκαν για την εκπαίδευση του BNΔ αναγνώρισης προσβλητικότητας.

Political Social Media Posts (PSMP)⁵: Αυτό το σύνολο δεδομένων περιέχει 5.000 μηνύματα από το twitter και το Facebook σχολιασμένα ως αντικειμενικά/μεροληπτικά που χρησιμοποιήθηκαν για τη δημιουργία ενός BNΔ αναγνώρισης μεροληψίας. Η παρουσία μηνυμάτων Facebook στο σύνολο δεδομένων δεν δημιουργούσε πρόβλημα, καθώς βρίσκονται επίσης στη γενική κατηγορία των σύντομων κειμένων, όπως και τα tweets.

Tweets with Sarcasm and Irony (TSI) [137]: Αυτό το σύνολο δεδομένων περιέχει περίπου 76.000 tweet που σχολιάζονται ως ειρωνικά/σαρκαστικά/μεταφορικά/κυριολεκτικά. Στο πλαίσιο αυτής της μελέτης, οι τρεις πρώτες τάξεις ομαδοποιήθηκαν σε μια ενιαία τάξη που ονομάζεται «μεταφορική», έτσι ώστε να μπορεί να εκπαιδευτεί ένα δυαδικό BNΔ αναγνώρισης μεταφορικότητας.

YouTube Comments (YTC): Προχωρώντας στο προεκπαιδευμένο BNΔ αναγνώρισης πολιτότητας, αρχικά εκπαιδεύτηκε σε ένα σύνολο δεδομένων με 12.559 σχόλια στο YouTube. Τα σχόλια αφαιρέθηκαν από 8 διαφορετικά βίντεο του YouTube που σχετίζονται με τις προεδρικές εκλογές του 2020 στις ΗΠΑ. Η επισήμανση έγινε αυτόματα μέσω TextBlob⁶ κι έτσι μια ετικέτα θετικής/αρνητικής γνώμης προστέθηκε σε κάθε σχόλιο.

4.4.2.2 Προεπεξεργασία Κειμένου

Εφαρμόστηκε πανομοιότυπη προεπεξεργασία στα τρία σύνολα δεδομένων εκπαίδευσης, δηλαδή αφαιρέθηκαν οι συνδετικές λέξεις, τα hashtags, οι αναφορές και οι διευθύνσεις URL. Αυτές οι οντότητες είτε δεν μας παρέχουν σημασιολογικές πληροφορίες, είτε κωδικοποιούν μόνο πληροφορίες για το συζητούμενο θέμα. Ωστόσο, ο προτεινόμενος μηχανισμός προϋποθέτει ότι το θέμα έχει επιλεγεί χειροκίνητα από τον χρήστη (στο Βήμα 1). Ως εκ τούτου, ο παρουσιαζόμενος αυτοματοποιημένος μεμονωμένος περιγραφέας tweet που βασίζεται στα τέσσερα προεκπαιδευμένα BNΔ καταγράφει μόνο συμπληρωματικές σημασιολογικές πληροφορίες, όπως πολιτότητα, μεροληψία, κ.λπ. Επιπλέον, εφαρμόστηκε λημματοποίηση για να αποφευχθεί η ύπαρξη πολλών λέξεων με το ίδιο νόημα. Τέλος, όλες οι λέξεις μετατράπηκαν σε πεζά. Αυτές είναι τυπικές επιλογές προεπεξεργασίας κειμένου στην ΕΦΓ.

⁵<https://www.kaggle.com/crowdflower/political-social-media-posts>

⁶<https://textblob.readthedocs.io/en/dev/>

4.4.2.3 Νευρωνικά Μοντέλα

Η αρχιτεκτονική BNΔ που εκπαιδεύτηκε ξεχωριστά για την αναγνώριση προσβλητικότητας, μεροληψίας και μεταφορικότητας [135] αποτελείται από ένα υβριδικό, παράλληλο ΑΜ-ΔΜΒΜ-ΣΝΔ. Οι αναπαραστάσεις κειμένου εισόδου (υπολογισμένες μετά την προεπεξεργασία του συνόλου μηνυμάτων), που τροφοδοτούνται σε αυτήν τη νευρωνική αρχιτεκτονική τόσο κατά τα στάδια εκπαίδευσης όσο και κατά τη διάρκεια του ελέγχου, προέρχονται από τη χρήση διανυσμάτων ενσωμάτωσης 200 διαστάσεων από ένα προεκπαιδευμένο μοντέλο GloVe [138]. Το ΣΝΔ εφαρμόζει συνέλιξη μεγεθών πυρήνα 3, 4 και 5, μαθαίνοντας έτσι χαρακτηριστικά σταθερού μήκους 3, 4 και 5 λέξεων, αντίστοιχα. Η συνέλιξη ακολουθείται από μια συνάρτηση ενεργοποίησης ReLU. Αυτά τα συνελικτικά χαρακτηριστικά στη συνέχεια υποβάλλονται σε δειγματοληψία χρησιμοποιώντας 1D max-pooling. Οι έξοδοι του ΣΝΔ συνενώνονται και στη συνέχεια συνδυάζονται με την έξοδο ΑΜ-ΔΜΒΜ και τροφοδοτούνται από κοινού σε ένα πλήρως συνδεδεμένο επίπεδο εξόδου, που ενεργοποιείται από μια σιγμοειδή συνάρτηση για την παραγωγή της τελικής σημασιολογικής τιμής: έναν πραγματικό αριθμό στην περιοχή [0, 1].

Το προεκπαιδευμένο νευρωνικό μοντέλο που χρησιμοποιείται για την αναγνώριση πολικότητας βασίζεται σε μια αρχιτεκτονική ΑΜ-ΔΜΒΜ. Ένα πλήρως συνδεδεμένο επίπεδο ενσωμάτωσης χρησιμοποιείται στο μπροστινό μέρος του δικτύου, το οποίο έχει μάθει να αντιστοιχίζει κάθε λέξη εισόδου σε μια διανυσματική αναπαράσταση 200 διαστάσεων. Αυτές οι ενσωματώσεις τροφοδοτούνται σε ένα επίπεδο ΑΜ-ΔΜΒΜ ακολουθούμενο από μια λειτουργία max-pooling. Στη συνέχεια, χρησιμοποιούνται πολλαπλά πλήρως συνδεδεμένα νευρωνικά επίπεδα που ενεργοποιούνται με τη συνάρτηση ReLU, καθώς και dropout για περαιτέρω μείωση της διάστασης της εξόδου. Η παραγόμενη πυκνή αναπαράσταση χαρακτηριστικών τροφοδοτείται στο τελικό πλήρως συνδεδεμένο επίπεδο σιγμοειδούς ενεργοποίησης που μας δίνει την τελική πραγματική τιμή συναισθήματος για την πολικότητα στην περιοχή [0, 1].

Ο πίνακας 4.1 συνοψίζει την επιτευχθείσα ακρίβεια αναγνώρισης (%) καθενός από τους τέσσερις ταξινομητές γνώμης στο σύνολο ελέγχου του αντίστοιχου συνόλου δεδομένων εκπαίδευσης.

Πίνακας 4.1: Επιτευχθείσα ακρίβεια καθενός από τους τέσσερις ταξινομητές γνώμης στο σύνολο δοκιμής του αντίστοιχου συνόλου δεδομένων εκπαίδευσης.

<i>Model</i>	<i>Accuracy</i>
Μεροληψία (Bias)	75.64%
Μεταφορικότητα (Figurativeness)	84.45%
Πολικότητα (Polarity)	88.00%
Προσβλητικότητα (Offensiveness)	84.00%

Κατά τη διάρκεια του σταδίου ελέγχου και για τα τέσσερα μοντέλα BNΔ, κάθε προεκπαιδευμένο BNΔ χρησιμοποιείται στην πραγματικότητα για τον υπολογισμό ενός διαφορετικού τιμήματος του μεμονωμένου 4Δ περιγραφέα tweet για κάθε tweet εισόδου. Η σημασιολογική τιμή εξόδου υποδηλώνει πόσο προσβλητικό/μεροληπτικό/μεταφορικό/αρνητικό κρίνεται το μήνυμα. Μια τιμή εξόδου 0 σημαίνει πολύ υψηλή πιθανότητα να είναι μη προσβλητικό, αντικειμενικό, μεταφορικό και αρνητικό, αντίστοιχα. Μια βαθμολογία εξόδου 1 σημαίνει πολύ υψηλή πιθανότητα να είναι προσβλητικό, μεροληπτικό, κυριολεκτικό και θετικό, αντίστοιχα. Μια βαθμολογία κοντά στο 0,5 θα υπονοούσε ότι το tweet κρίνεται ουδέτερο στο αντίστοιχο χαρακτηριστικό ή απλώς δεν μπορεί να ταξινομηθεί.

4.4.2.4 Υπερπαράμετροι

Οι βέλτιστες υπερπαράμετροι που χρησιμοποιούνται για την εκπαίδευση των ταξινομητών BNΔ λήφθηκαν με χειροκίνητη προσαρμογή και παρουσιάζονται στον πίνακα 4.2. Ο προεπιαδευμένος ταξινομητής πολικότητας χρησιμοποίησε τις υπερπαραμέτρους που φαίνονται στο σχετικό αποθετήριο λογισμικού⁷.

Το N_layers υποδηλώνει τον αριθμό των κρυφών επιπέδων στο ΔΜΒΜ και το N_hidden είναι το μέγεθος (αριθμός νευρώνων) αυτών των επιπέδων. Το Weight_decay υποδηλώνει το παράγοντα κανονικοποίησης L2. Το Lr_decay υποδηλώνει τον πολλαπλασιαστικό παράγοντα μάθησης. Το Wd_multiplier υποδηλώνει τον πολλαπλασιαστικό παράγοντα μείωσης βάρους. Το Batch_size υποδηλώνει τον αριθμό των tweet που υποβάλλονται σε επεξεργασία σε κάθε βήμα της βελτιστοποίησης. Τα Dropout, dropout_enc και dropout_op είναι τα ποσοστά dropout που χρησιμοποιούνται μετά τα επίπεδα ΔΜΒΜ, Ενσωμάτωσης και Εξόδου αντίστοιχα.

Πίνακας 4.2: Υπερπαράμετροι που χρησιμοποιούνται για την εκπαίδευση των ταξινομητών συναισθημάτων.

Υπερπαράμετρος	Τιμή
n_hidden	200
n_layers	2
dropout	0.5
weight_decay	1e-7
dropout_enc	0.2
dropout_op	0.5
lr_decay	0.7
wd_multiplier	6
learning_rate	5e-3
batch_size	52

⁷<https://github.com/DheerajKumar97/US-2020-Election-Campaign-Youtube-Comments-Sentiment-Analysis>

4.4.3 Βήμα 3: Συνάθροιση

Έχοντας αποκτήσει έναν μεμονωμένο 4Δ σημασιολογικό περιγραφέα πραγματικών τιμών ανά tweet, το επόμενο βήμα είναι να συναθροιστούν τα παράγωγα διανύσματα στον επιθυμητό περιγραφέα κοινής γνώμης. Για να γίνει αυτό, ο χρήστης πρέπει να επιλέξει τη **στρατηγική συνάθροισης και το στόχο συνάθροισης**.

Ο στόχος αναφέρεται στον βαθμό λεπτομέρειας της διαδικασίας συνάθροισης και επηρεάζει άμεσα τον τελικό αριθμό των συνολικών διανυσμάτων περιγραφής της κοινής γνώμης. Δύο πιθανές επιλογές είναι οι πιο απλές:

- **Ολική Συνάθροιση.** Όλοι οι μεμονωμένοι περιγραφείς tweet συγχωνεύονται σε ένα ενιαίο διάνυσμα περιγραφής της κοινής γνώμης, που αντιπροσωπεύει ολόκληρη τη δεξαμενή μηνυμάτων που εξήχθη στο Βήμα 1.
- **Χρονικά Τμηματοποιημένη Συνάθροιση.** Το συνολικό εύρος ημερομηνιών από τις οποίες εξήχθη ολόκληρο το σύνολο μηνυμάτων στο Βήμα 1 χωρίζεται σε ισόχρονες, μη επικαλυπτόμενες και διαδοχικές χρονικές περιόδους. Όλοι οι μεμονωμένοι περιγραφείς tweet που εμπίπτουν σε κάθε περίοδο συγχωνεύονται σε ένα ενιαίο διάνυσμα περιγραφής συνολικής κοινής γνώμης. Αυτό εκτελείται ξεχωριστά για κάθε περίοδο.

Με πλήρη συνάθροιση, το αποτέλεσμα είναι ένα ενιαίο 4Δ διάνυσμα. Με τη χρονικά τμηματοποιημένη συνάθροιση, το αποτέλεσμα είναι μια 4Δ χρονοσειρά. Παραδείγματα χρονικά τμηματοποιημένων στόχων συνάθροισης θα ήταν η μέρα με τη μέρα ή εβδομάδα με την εβδομάδα συνάθροιση. Ανάλογα με την εφαρμογή, μπορούν να χρησιμοποιηθούν διαφορετικοί πρόσθετοι στόχοι συνάθροισης.

Η στρατηγική συνάθροισης αναφέρεται στον τρόπο με τον οποίο ένα σύνολο μεμονωμένων 4Δ περιγραφών tweet συνδυάζονται σε έναν ενιαίο αθροιστικό 4Δ περιγραφέα. Τρεις πιθανές επιλογές είναι οι πιο απλές:

- **Διανυσματικός μέσος όρος στοιχείων.**
- **Διανυσματική διάμεσος στοιχείων.**
- **Διανυσματικός περικομμένος μέσος όρος στοιχείων.**

Και οι τρεις αυτές επιλογές μπορούν να υλοποιηθούν απλώς εκτελώντας υπολογισμούς χωριστά κατά μήκος καθεμιάς από τις τέσσερις διαστάσεις του περιγραφέα. Όπως και πριν, μπορούν επίσης να χρησιμοποιηθούν διαφορετικές πρόσθετες στρατηγικές συνάθροισης, ανάλογα με την εφαρμογή.

4.5 Αξιολόγηση

Ο προτεινόμενος μηχανισμός αξιολογήθηκε με βάση τα σύνολα tweet των Προεδρικών Εκλογών των Ηνωμένων Πολιτειών του 2016 και του 2020, χρησιμοποιώντας έναν ημέρα προς ημέρα χρονικά τμηματοποιημένο στόχο συνάθροισης. Και οι τρεις στρατηγικές συνάθροισης που περιγράφονται στην ενότητα 4.4 ακολουθήθηκαν και αξιολογήθηκαν ξεχωριστά. Λεπτομέρειες και αποτελέσματα ακολουθούν στις επόμενες υποενότητες.

4.5.1 Δεδομένα

Το σύνολο tweet για τις προεδρικές εκλογές των ΗΠΑ του 2016 από το Kaggle⁸ περιέχει 61 εκατομμύρια εγγραφές. Το συνολικό χρονικό εύρος τους είναι από 30-08-2016 έως 28-02-2017, με 20 ημέρες να λείπουν, δηλαδή 163 ημέρες συνολικά. Από αυτό το αρχικό σύνολο δεδομένων, διατηρήσαμε περίπου 32 εκατομμύρια tweets μετά την εφαρμογή μιας κοινής διαδικασίας καθαρισμού: αφαίρεση κενών σειρών, μη αγγλικού κειμένου, διπλότυπων tweet και μηνυμάτων που περιείχαν λιγότερες από 5 λέξεις μετά την προεπεξεργασία κειμένου. Αυτό είναι σημαντικό για τη σωστή σημασιολογική ανάλυση, καθώς ένας επαρκής αριθμός λέξεων ανά tweet είναι απαραίτητος για την επίτευξη υψηλής ακρίβειας εξόρυξης γνώμης. Στη συνέχεια, οι λέξεις-κλειδιά «Κλίντον», «Ομπάμα» και «Τραμπ» χρησιμοποιήθηκαν για τον διαχωρισμό των μηνυμάτων σε αυτά που αναφέρονται σε Δημοκρατικούς και σε αυτά που αναφέρονται σε Ρεπουμπλικάνους.

Συνολικά, όλη αυτή η μη αυτόματη διαδικασία ισοδυναμούσε με την εκτέλεση του Βήματος 1 του προτεινόμενου μηχανισμού. Δεν χρειάστηκε στο σύνολό του για το μικρότερο δεύτερο σύνολο δεδομένων που χρησιμοποιήσαμε, δηλαδή τα tweets για τις εκλογές στις ΗΠΑ το 2020 από το Kaggle⁹, δεδομένου ότι τα tweets του είναι προετοιμασμένα σε δύο κομματικές ομάδες (Δημοκρατικούς και Ρεπουμπλικάνους). Περιέχει 1,72 εκατομμύρια σειρές συνολικά, με συνολικό χρονικό εύρος από 15-10-2020 έως 08-11-2020, δηλαδή 25 ημέρες συνολικά. Από αυτό το αρχικό σύνολο δεδομένων, διατηρήθηκαν περίπου 720 χιλιάδες tweets μετά την εφαρμογή της διαδικασίας καθαρισμού που περιγράφεται παραπάνω.

Στη συνέχεια, ο προτεινόμενος μηχανισμός εφαρμόστηκε ξεχωριστά στις δύο ομάδες μηνυμάτων κάθε συνόλου δεδομένων. Ως αποτέλεσμα, προέκυψαν δύο μέρα με τη μέρα χρονοσειρές 4Δ περιγραφέων, που καλύπτουν συνολικά ακριβώς το ίδιο χρονικό εύρος: ένα για τους Ρεπουμπλικάνους και ένα για τους Δημοκρατικούς (ξεχωριστά για κάθε σύνολο δεδομένων - 2016 και 2020). Ενδεικτικά, για το σύνολο tweet των Προεδρικών Εκλογών των ΗΠΑ 2016, η δημιουργία των περιγραφέων για όλα τα σχετικά tweets πήρε περίπου 24 ώρες. Η

⁸<https://www.kaggle.com/paulrohan2020/2016-usa-presidential-election-tweets61m-rows>

⁹<https://www.kaggle.com/manchunhui/us-election-2020-tweets>

συνάθροιση ανα μέρα απαιτούσε 10 λεπτά για κάθε στρατηγική συνάθροισης. Τα πειράματα πραγματοποιήθηκαν σε προσωπικό υπολογιστή με CPU AMD Ryzen 5@3.2GHz, 16GB μνήμης RAM DDR4 και nVidia GeForce GTX1060 (6GB RAM) GP-GPU.

4.5.2 1η Ανάλυση: Πρόβλεψη Χρονοσειράς

Ο πρώτος τρόπος αξιολόγησης που πραγματοποιήθηκε στις παραγόμενες χρονοσειρές ήταν να αξιολογηθεί η προβλεψιμότητά τους χρησιμοποιώντας μοντέλα προβλεψης τεχνητής νοημοσύνης. Δεδομένου ότι οι δύο κατασκευασμένες χρονοσειρές περιγράφουν την κοινή γνώμη για τα δύο αντίστοιχα κόμματα κατά τη διάρκεια μιας έντονης προεκλογικής περιόδου, η πρόβλεψη έχει προφανή πολιτική χρησιμότητα: μπορεί να επιτρέψει σε έναν ενδιαφερόμενο οργανισμό να προβλέψει αλλαγές στο κοντινό μέλλον της δημόσιας εικόνας του, χρησιμοποιώντας δεδομένα του twitter. Φυσικά, σε αυτό το πλαίσιο, το «κοντινό μέλλον» συνεπάγεται έναν ορίζοντα προβλεψης το πολύ μερικών εβδομάδων.

4.5.2.1 Υλοποίηση

Οι δύο χρονοσειρές είναι καθημερινές 4Δ περιγραφές της εξελισσόμενης κοινής γνώμης για τους Δημοκρατικούς και τους Ρεπουμπλικάνους, αντίστοιχα. Δεδομένου ότι χρησιμοποιήθηκαν τρεις διαφορετικές στρατηγικές συνάθροισης (μέσος όρος, διάμεσος, περικομμένος μέσος όρος), στην πραγματικότητα προέκυψαν συνολικά έξι 4Δ χρονοσειρές, με όλες να καλύπτουν την ίδια χρονική περίοδο. Εφόσον η προβλεψη εκτελείται συνήθως σε μονομεταβλητές χρονοσειρές, στη συνέχεια έγινε ο χειρισμός κάθε καναλιού του 4Δ περιγραφέα ξεχωριστά, οδηγώντας σε συνολικά 24 διαφορετικές χρονοσειρές.

Επιλέχθηκε ένας ορίζοντας προβλεψης 7 ημερών, καθώς αυτό επιτρέπει μάλλον αξιόπιστες προβλέψεις, ενώ εξαιρούνται να είναι πρακτικά χρήσιμος. Μια πρόσφατη στοιβαγμένη αρχιτεκτονική ΔMBM υιοθετήθηκε [139], με κάθε επίπεδο ΔMBM ακολουθούμενο από ένα πλήρως συνδεδεμένο νευρωνικό επίπεδο. Το συνολικό BNΔ εκπαιδεύτηκε ξεχωριστά για κάθε χρονοσειρά, χρησιμοποιώντας Οπίσθια-Διάδωση Σφάλματος στο Χρόνο (ΟΔΣΧ) και τον Βελτιστοποιητή Συνεχούς Ρίψης Κέρματος (ΒΣΡΚ) [140]. Αυτή η εκπαιδευτική προσέγγιση επιλέχθηκε έναντι εναλλακτικών βελτιστοποιητών (Adagrad, Adam) επειδή εμφάνισε ανώτερη απόδοση στα πολλαπλά πειράματα του [139]. Το γεγονός ότι ο ΒΣΡΚ επιχειρεί να ελαχιστοποιήσει τη συνάρτηση κόστους κατά την εκπαίδευση του νευρωνικού αυτορυθμίζοντας το ρυθμό εκμάθησής, επιταχύνει τη σύγκλιση σε σύγκριση με άλλους αλγόριθμους που βασίζονται σε διαβάθμιση με σταθερό ή φθίνων ρυθμό εκμάθησης, όπου η σύγκλιση γίνεται πιο αργή κοντά στο βέλτιστο σημείο. Η χρήση του ΟΔΣΧ για την ενημέρωση των παραμέτρων του μοντέλου κατά τη διάρκεια της εκπαίδευσης είναι απαραίτητη όταν χρησιμοποιούνται ΔMBM.

Οι βέλτιστες υπερπαράμετροι για την εκπαίδευση του μοντέλου πρόβλεψης λήφθηκαν χρησιμοποιώντας τη Διαμόρφωση Αλγορίθμου βάσει Διαδοχικού Μοντέλου (ΔΑΔΜ) [141] και παρουσιάζονται στον Πίνακα 4.3. Το minibatch_size υποδηλώνει τον αριθμό των χρονοσειρών που λαμβάνονται υπόψη για κάθε πλήρη οπίσθια διάδοση στο ΔΜΒΜ. Το Epoch_size υποδηλώνει πόσες φορές το σύνολο δεδομένων διασχίζεται σε κάθε εποχή. Εποχές είναι τα 'περάσματα' του συνόλου των δεδομένων από το μοντέλο κατά την εκπαίδευση και στην περίπτωσή μας γίνονται 2 τέτοια περάσματα (num_epochs). Το l2_regularization και το gaussian_noise που προστίθεται στην είσοδο χρησιμοποιούνται για τη μείωση της υπερπροσαρμογής του μοντέλου στα δεδομένα εκπαίδευσης. Τα βάρη μονάδων ΔΜΒΜ αρχικοποιήθηκαν χρησιμοποιώντας έναν random_normal_initializer. Τέλος, οι τιμές cell_dimension και num_hidden_layers αναφέρονται στα ΔΜΒΜ που χρησιμοποιούνται στην αρχιτεκτονική αυτή.

Πίνακας 4.3: Υπερπαράμετροι που χρησιμοποιούνται για την εκπαίδευση του μοντέλου πρόβλεψης.

Υπερπαράμετρος	Τιμή
cell_dimension	20
gaussian_noise_stdev	1e-4
l2_regularization	1e-4
max_epoch_size	1
max_num_epochs	2
minibatch_size	4
num_hidden_layers	1
random_normal_initializer_stdev	1e-4

Από τα δύο σύνολα δεδομένων αξιολόγησης, μόνο το 2016 χρησιμοποιήθηκε για την εκπαίδευση του μοντέλου πρόβλεψης. Ένα τμήμα παρακρατήθηκε για δοκιμαστικούς σκοπούς από το τέλος κάθε χρονοσειράς, με μήκος ίσο με τον ορίζοντα πρόβλεψης. τα υπόλοιπα δεδομένα αποτελούσαν το σύνολο δεδομένων εκπαίδευσης. Επιπλέον, αυτό το προεκπαιδευμένο μοντέλο δοκιμάστηκε ξεχωριστά στο σύνολο δεδομένων του 2020. Τα αποτελέσματα από τις δοκιμές και στα δύο σύνολα δεδομένων παρουσιάζονται στη συνέχεια.

Η αποεποχοποίηση εφαρμόστηκε ως ένα κοινό βήμα προεπεξεργασίας, καθώς τα BNΔ είναι αδύναμα στη μοντελοποίηση της εποχικότητας [142]. Αυτό επιτεύχθηκε με την αποσύνθεση κάθε χρονοσειράς σε στοιχεία εποχικότητας, τάσης και υπολοίπων, προκειμένου στη συνέχεια να αφαιρεθεί το στοιχείο εποχικότητας. Εάν μια χρονοσειρά δεν εμφάνιζε εποχικότητα, αυτό το βήμα απλώς επέστρεφε μηδενική τιμή εποχικότητας.

Υιοθετήθηκαν σχήματα οιισθαίνοντων παραθύρων για την τροφοδοσία εισόδων στο BNΔ και την εξαγωγή των εξόδων, με το μέγεθος του παραθύρου εξόδου n να είναι ίσο με το μέγεθος του ορίζοντα πρόβλεψης $H = 7$. Το μέγεθος του παραθύρου εισόδου m ορίστηκε

εμπειρικά σε $9 = n1.25$. Κάθε χρονοσειρά εκπαίδευσης αναλύθηκε σε μπλοκ μεγέθους $m+n$, σχηματίζοντας έτσι τα ζεύγη εισόδου-εξόδου για κάθε στιγμιότυπο μονάδας ΔΜΒΜ.

Πίνακας 4.4: Πρόβλεψη αποτελεσμάτων σχετικά με το σύνολο δεδομένων tweet για τις προεδρικές εκλογές των ΗΠΑ 2016 για τις έξι κατασκευασμένες χρονοσειρές. Το 'Dem' αναφέρεται στους Δημοκρατικούς, το 'Rep' στους Ρεπουμπλικάνους, ενώ τα 'mean', 'med', 'trim' στις αντίστοιχες 3 στρατηγικές συνάθροισης: μέσου όρου, διαμέσου και περικομμένου μέσου όρου. Σε κάθε περίπτωση, οι μετρικές SMAPE/MASE έχουν υπολογιστεί παίρνοντας μια ενιαία τιμή και για τα τέσσερα κανάλια περιγραφής χρησιμοποιώντας τόσο τον μέσο όρο όσο και τον διάμεσο τελεστή. Μια χαμηλότερη τιμή είναι καλύτερη και για τις δύο μετρικές, ενώ το SMAPE είναι ένα ποσοστό.

Χρονοσειρά	Mean SMAPE	Median SMAPE	Mean MASE	Median MASE
Dem-Mean	0.1096	0.0563	1.2396	1.0799
Dem-Med	0.1380	0.0913	1.0620	1.1711
Dem-Trim	0.1798	0.0676	1.3364	1.0885
Rep-Mean	0.0529	0.0280	0.7689	0.6937
Rep-Med	0.0492	0.0330	0.5158	0.5303
Rep-Trim	0.0737	0.0314	0.6931	0.6549

Πίνακας 4.5: Πρόβλεψη αποτελεσμάτων σχετικά με το σύνολο δεδομένων tweet για τις προεδρικές εκλογές των ΗΠΑ 2020 για τις έξι κατασκευασμένες χρονοσειρές. Το 'Dem' αναφέρεται στους Δημοκρατικούς, το 'Rep' στους Ρεπουμπλικάνους, ενώ τα 'mean', 'med', 'trim' στις αντίστοιχες 3 στρατηγικές συνάθροισης: μέσου όρου, διαμέσου και περικομμένου μέσου όρου. Σε κάθε περίπτωση, οι μετρικές SMAPE/MASE έχουν υπολογιστεί παίρνοντας μια ενιαία τιμή και για τα τέσσερα κανάλια περιγραφής χρησιμοποιώντας τόσο τον μέσο όρο όσο και τον διάμεσο τελεστή. Μια χαμηλότερη τιμή είναι καλύτερη και για τις δύο μετρικές, ενώ το SMAPE είναι ένα ποσοστό.

Χρονοσειρά	Mean SMAPE	Median SMAPE	Mean MASE	Median MASE
Dem-Mean	0.1793	0.1490	1.6975	1.6943
Dem-Med	0.3431	0.2088	1.7526	1.6620
Dem-Trim	0.2847	0.1903	1.7535	1.7434
Rep-Mean	0.0961	0.0813	1.5233	1.4572
Rep-Med	0.1867	0.1303	1.7228	1.6949
Rep-Trim	0.1472	0.0999	1.5961	1.5637

4.5.2.2 Μετρικές και Αποτελέσματα

Χρησιμοποιήθηκε ένα σύνολο κοινών μετρικών αξιολόγησης πρόβλεψης χρονοσειρών για την αξιολόγηση της προβλεψιμότητας των υπολογισμένων χρονοσειρών.

Αρχικά, το Συμμετρικό Μέσο Απόλυτο Ποσοστιαίο Σφάλμα (SMAPE) ορίζεται ως εξής:

$$SMAPE = \frac{100\%}{H} \sum_{k=1}^H \frac{|F_k - Y_k|}{(|Y_k| + |F_k|)/2}, \quad (4.1)$$

όπου H , F_k , και Y_k υποδηλώνουν τον χρονικό ορίζοντα πρόβλεψης, την πρόβλεψη του μοντέλου και την πραγματική τιμή που προσπαθούμε να προβλέψουμε, αντίστοιχα.

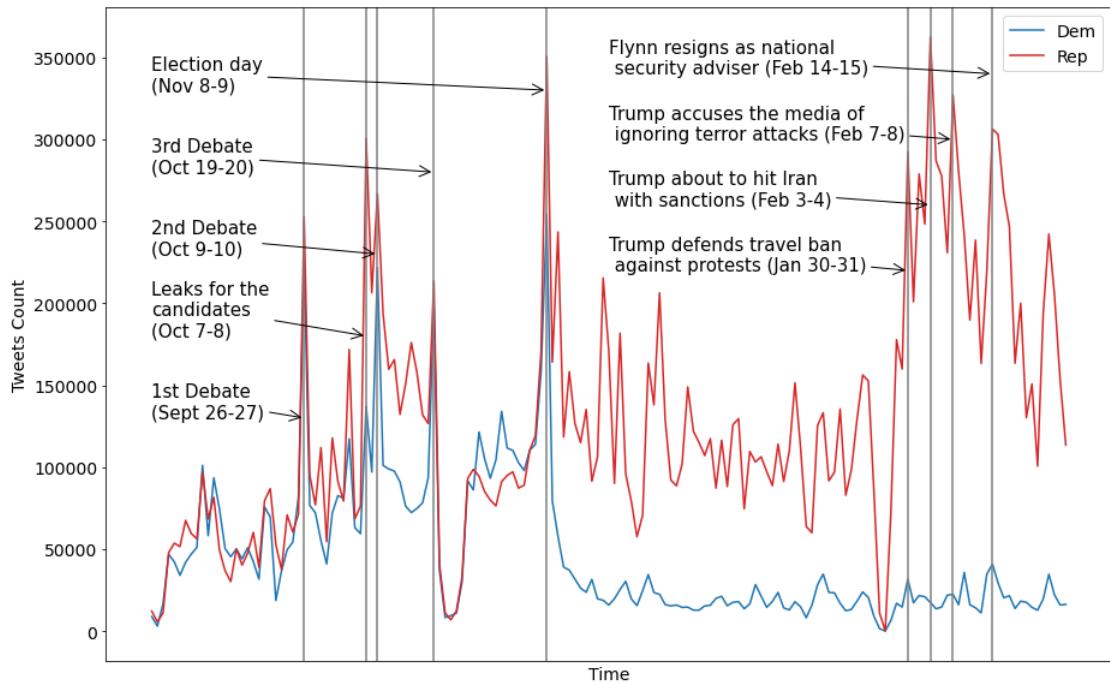
Εξαιτίας της δύσκολης ερμηνείας του SMAPE [143], το ανεξάρτητο από την ηλίμακα Μέσο Απόλυτο Κλιμακωμένο Σφάλμα (MASE) χρησιμοποιήθηκε επιπρόσθετα. Για μη περιοδικές χρονοσειρές, ορίζεται ως εξής:

$$MASE = \frac{\frac{1}{H} \sum_{k=1}^H |F_k - Y_k|}{\frac{1}{T-1} \sum_{k=2}^T |Y_k - Y_{k-1}|}, \quad (4.2)$$

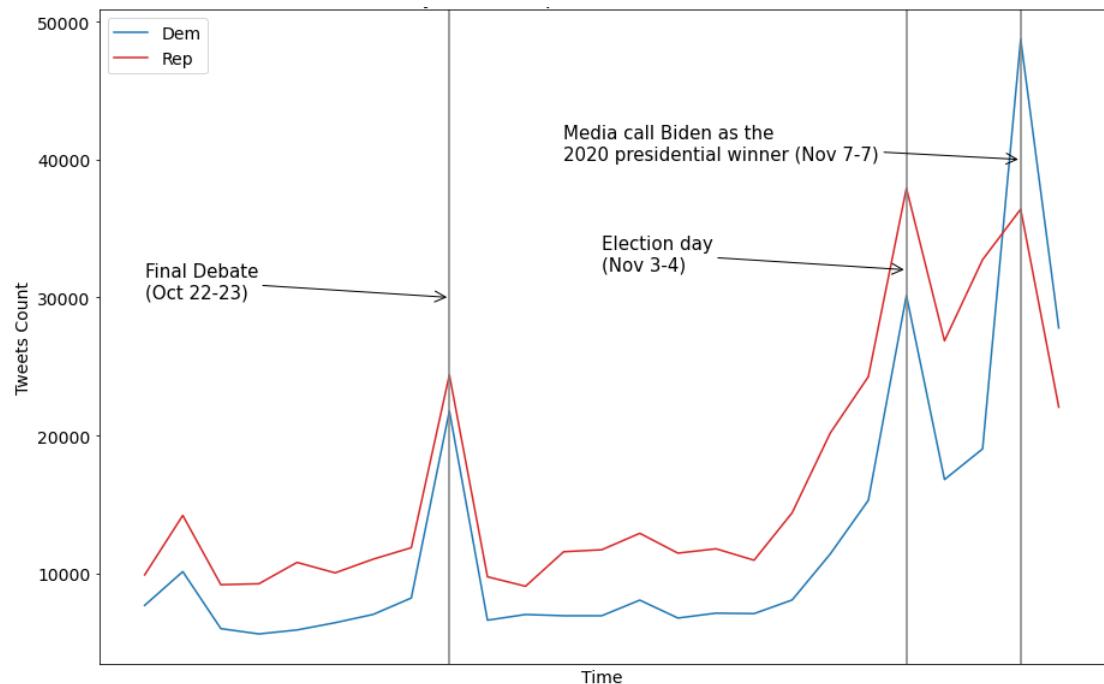
Στην εξ. (4.2), ο αριθμητής είναι ο ίδιος όπως στο SMAPE, αλλά κανονικοποιημένος από το μέσο σφάλμα πρόβλεψης ενός βήματος στο δείγμα. Μια τιμή MASE μεγαλύτερη από 1 υποδηλώνει ότι η απόδοση του δοκιμασμένου μοντέλου είναι κατά μέσο όρο χειρότερη από την Αφελή Μέθοδο Πρόβλεψης (ΑΜΠ), κατά την οποία χρησιμοποιούμε τις πιο πρόσφατες τιμές της χρονοσειράς για να προβλέψουμε μελλοντικές τιμές της (για παράδειγμα κοιτάμε την τελευταία περίοδο για να καθορίσουμε την επόμενη). Αντίθετα, μια τιμή μικρότερη από 1 υποδηλώνει απόδοση καλύτερη από την ΑΜΠ. Επομένως, αυτή η μέτρηση σφάλματος παρέχει μια άμεση σύγκριση της ακρίβειας της πρόβλεψης σε σχέση με την ΑΜΠ.

Εφόσον αυτές οι μετρήσεις υπολογίζονται για πρόβλεψη μονομεταβλητών χρονοσειρών, χρησιμοποιήσαμε τη μέση και τη διάμεση συνάθροιση στα τέσσερα κανάλια περιγραφέα για καθεμιά από τις έξι χρονοσειρές. Τα αποτελέσματα που ελήφθησαν για τα σύνολα δεδομένων του 2016 και του 2020 εμφανίζονται στους πίνακες 4.4 και 4.5, αντίστοιχα, όπου μεγαλύτερες τιμές SMAPE ή MASE υποδεικνύουν χειρότερη ακρίβεια πρόβλεψης.

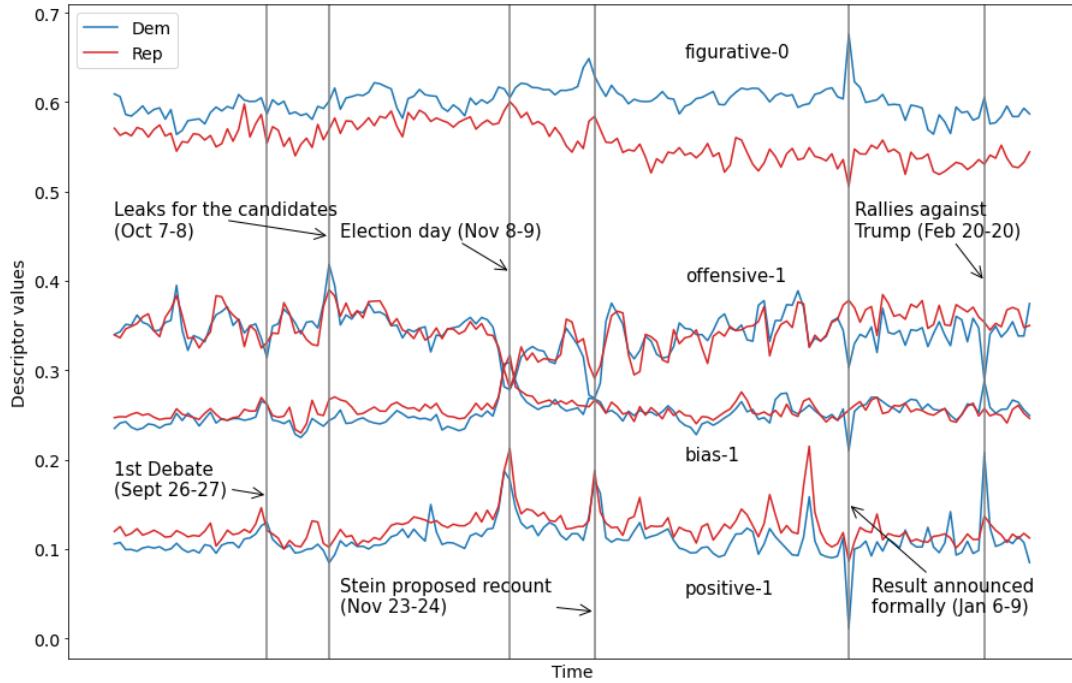
Γενικά, οι χρονοσειρές που κατασκευάζονται χρησιμοποιώντας τον προτεινόμενο μηχανισμό φαίνεται να είναι προβλέψιμες σε αποδεκτό βαθμό χρησιμοποιώντας το χρησιμοποιούμενο μοντέλο BNΔ. Επιπλέον, η πρόβλεψη συμπεριφέρεται παρόμοια και για τα δύο σύνολα δεδομένων, οδηγώντας μας στο να βγάλουμε ένα κοινό σύνολο συμπερασμάτων. Πρώτον, η στρατηγική συγκέντρωσης μέσου όρου είχε ως αποτέλεσμα τις χρονοσειρές με την καλύτερη συνολική συμπεριφορά πρόβλεψης. Δεύτερον, με βάση και τις δύο μετρικές, είναι σαφές ότι οι προβλέψεις αποδίδουν χειρότερα για τους Δημοκρατικούς παρά για τους Ρεπουμπλικάνους, υπονοώντας ότι η κοινή γνώμη σχετικά με αυτούς (όπως εκφράζεται στο twitter) ήταν λιγότερο σταθερή και προβλέψιμη κατά την εξεταζόμενη περίοδο. Τέλος, παρατηρούμε πτώση στην ακρίβεια κατά τη δοκιμή στο σύνολο δεδομένων του 2020 (μικρή σε απόλυτες τιμές), σε σύγκριση με το 2016. Αυτό είναι αναμενόμενο, καθώς το BNΔ πρόβλεψης είχε προεκπαιδευτεί μόνο στο σετ εκπαίδευσης του συνόλου δεδομένων του 2016.



Σχήμα 4.2: Καθημερινός αριθμός tweets για Δημοκρατικούς και Ρεπουμπλικάνους στο σύνολο δεδομένων του 2016. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).



Σχήμα 4.3: Καθημερινός αριθμός tweets για Δημοκρατικούς και Ρεπουμπλικάνους στο σύνολο δεδομένων του 2020. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).

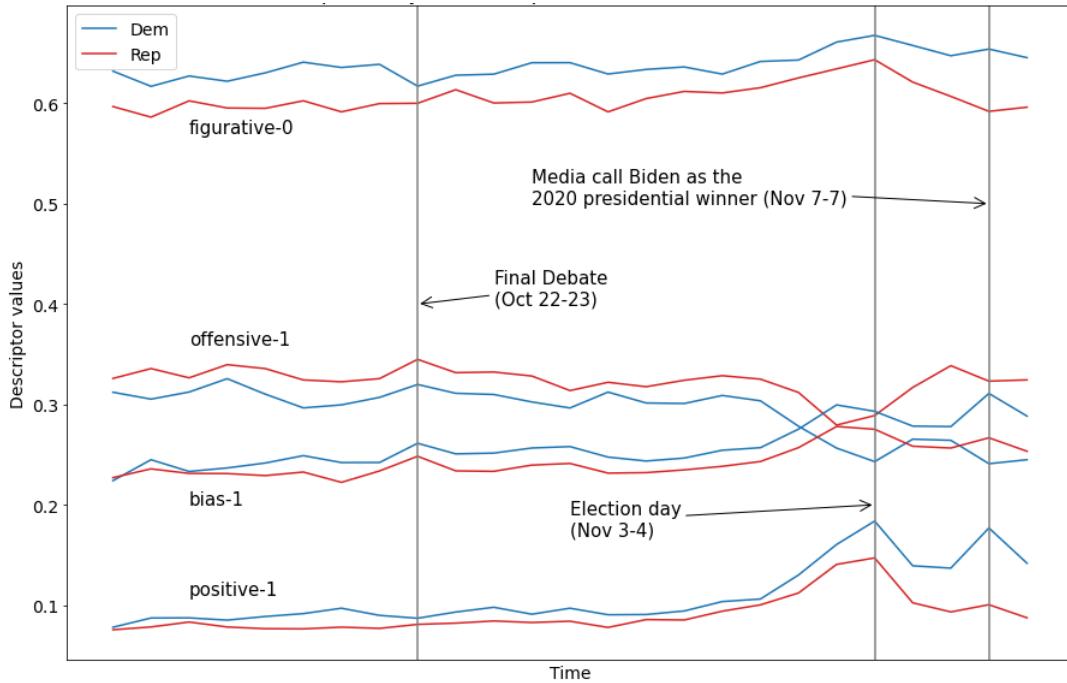


Σχήμα 4.4: Οι καθημερινές τιμές ανά κανάλι των χρονοσειρών 4Δ που κατασκευάστηκαν από τα δεδομένα του 2016 χρησιμοποιώντας τον προτεινόμενο περιγραφέα και τη μέση στρατηγική συγκέντρωσης, ξεχωριστά για Δημοκρατικούς και Ρεπουμπλικάνους. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).

4.5.3 2η Ανάλυση: Οπτικοποιήσεις και Ποιοτική Αξιολόγηση

Ένα σύνολο οπτικοποιήσεων δημιουργήθηκε από τις 4Δ χρονοσειρές που κατασκευάστηκαν χρησιμοποιώντας τον προτεινόμενο μηχανισμό, προκειμένου να διευκολυνθεί η μη αυτόματη επιθεώρηση των αποτελεσμάτων. Λεδομένων των συμπερασμάτων της υποενότητας 4.5.2, μόνο οι χρονοσειρές που προέκυψαν από τη συνάθροιση μέσου όρου χρησιμοποιήθηκαν εδώ. Αυτή η υποενότητα παρουσιάζει αυτήν τη διαδικασία ποιοτικής αξιολόγησης και τα συμπεράσματά της, μαζί με βιοηθητικές πληροφορίες σχετικά με τα αρχικά σύνολα δεδομένων για τις προεδρικές εκλογές των ΗΠΑ 2016/2020.

Πρώτον, το σχήμα 4.2 απεικονίζει τον αριθμό των tweets που δημοσιεύτηκαν κάθε μέρα στο χρονικό εύρος του πλήρους συνόλου δεδομένων (από 30-08-2016 έως 28-02-2017), ξεχωριστά για τους Δημοκρατικούς και τους Ρεπουμπλικάνους. Για τα πιο σημαντικά γεγονότα, όπως τα τρία προεδρικά ντιμπέιτ και η ημέρα των εκλογών, παρατηρείται αυξημένη επισκεψιμότητα στο twitter και για τα δύο κόμματα. Οι κατηγορίες τόσο για την Κλίντον όσο και για τον Τραμπ που σημειώθηκαν το 07-10-2016 φαίνεται να επηρέασαν περισσότερο τον τελευταίο υποψήφιο, καθώς η πλειονότητα των δημοσιεύσεων εξέφραζε άποψη για αυτόν. Ίση κίνηση παρατηρείται και για τα δύο κόμματα λίγο πριν την ημέρα των εκλογών (08-11-2016), αφού σε αυτό το στάδιο το twitter παίζει σημαντικό ρόλο στην εκστρατεία και των δύο υποψηφίων. Ωστόσο, ο αριθμός των tweets που αφορούν τους Δημοκρατικούς μειώνεται σημαντικά μετά



Σχήμα 4.5: Οι καθημερινές τιμές ανά κανάλι των χρονοσειρών 4Δ που κατασκευάστηκαν από τα δεδομένα του 2020 χρησιμοποιώντας τον προτεινόμενο περιγραφέα και τη μέση στρατηγική συγκέντρωσης, ξεχωριστά για Δημοκρατικούς και Ρεπουμπλικάνους. Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).

την ημέρα των εκλογών και την ήττα τους. Αντίθετα, ο κόσμος συνέχισε να ανεβάζει συχνά tweet για τον νικητή και επίκεντρο της προσοχής Ντόναλντ Τραμπ σχετικά με τις ενέργειές του ως πρόεδρος των ΗΠΑ, όπως η ταξιδιωτική απαγόρευση, οι κυρώσεις στο Ιράν κ.λπ.

Το σχήμα 4.3 απεικονίζει τον αριθμό των tweets που δημοσιεύτηκαν από 15-10-2020 έως 08-11-2020, ξεχωριστά για τους Δημοκρατικούς και τους Ρεπουμπλικάνους. Και πάλι, παρατηρείται αυξημένη επισκεψιμότητα στο twitter και για τα δύο κόμματα κατά τη διάρκεια των πιο σημαντικών γεγονότων. Ωστόσο, λιγότερα συμβάντα παρατηρούνται στην γραφική παράσταση του 2020, γεγονός που οφείλεται στο μικρότερο μέγεθος του συνόλου δεδομένων σε σύγκριση με το σύνολο δεδομένων του 2016. Γενικά, υπήρχαν περισσότερα tweets για τον Τραμπ μέχρι τις 7 Νοεμβρίου. Εκείνη την ημέρα, τα μέσα ενημέρωσης αποκαλούσαν τον Μπάιντεν ως τον νικητή του 2020 και η κίνηση στο twitter εξερράγη για τον Μπάιντεν, ξεπερνώντας τις αναρτήσεις του Τραμπ με σημαντική διαφορά. Αυτό προφανώς βγάζει νόημα καθώς η νίκη του Μπάιντεν ανακοινώνεται επίσημα για πρώτη φορά.

Έχοντας καθιερώσει πρωτότυπα μοτίβα επισκεψιμότητας στο twitter, οι ταυτόχρονες ημέρσιες τιμές των 4Δ χρονοσειρών που κατασκευάστηκαν χρησιμοποιώντας τον προτεινόμενο περιγραφέα και τη στρατηγική συνάθροισης μέσου όρου απεικονίζονται στο σχήμα 4.4, ξεχωριστά για κάθε κόμμα των εκλογών του 2016 στις ΗΠΑ (η συσχέτιση με το κόμμα είναι με έγχρωμη κωδικοποίηση). Σε αυτό το σχήμα, όπως και σε πολλά επόμενα, επισημαίνουμε

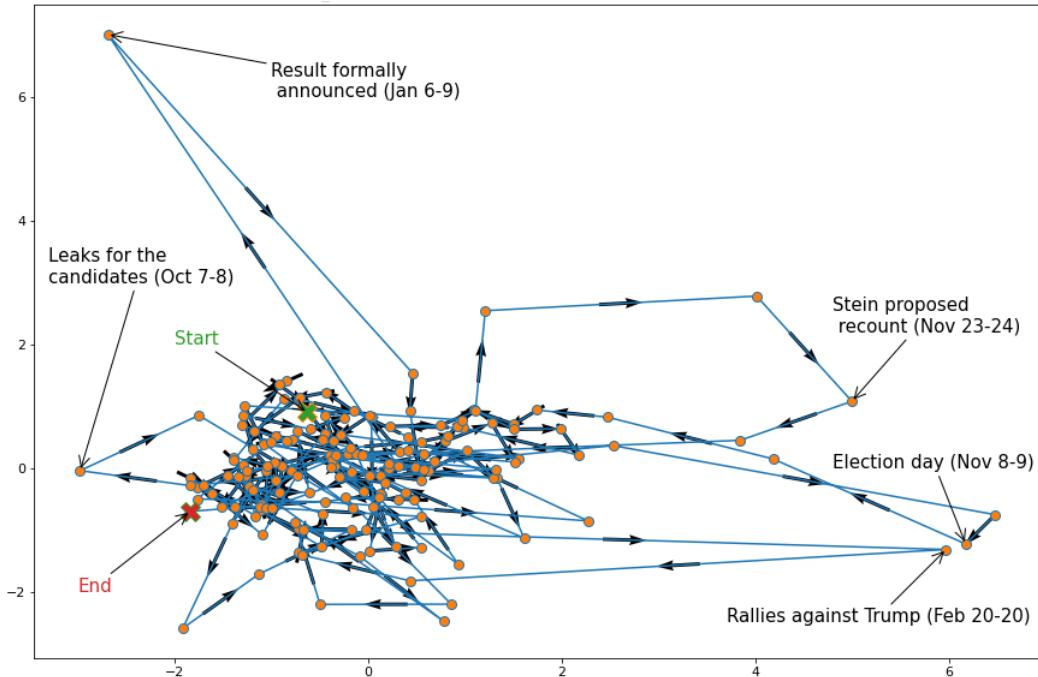
κάθε χρονοσειρά με μία από τις δύο αντίθετες ετικέτες ηλάσης (π.χ. 'θετική' ή 'μεταφορική') ακολουθούμενη από την τιμή εξόδου του ταξινομητή tweet που υποδηλώνει αυτήν την ετικέτα (π.χ. Το tweet πλήρως και αναμφίβολα ταξινομημένο ως μεταφορικό/θετικό, έχει αποδοθεί τιμή 0/1 από τον ταξινομητή μεταφορικότητας/πολικότητας, αντίστοιχα). Είναι προφανές ότι οι χρονοσειρές διατηρούν μια σταθερή ηλάση και στις τέσσερις διαστάσεις και σε όλο το χρονικό εύρος και για τα δύο κόμματα: η ηλάση τους είναι αρνητική (σε πολικότητα), αντικειμενική, μη προσβλητική και κυριολεκτική. Αυτό υποδηλώνει μια γενική στάση του κοινού απέναντι στους ανταγωνιστές πολιτικούς, που αντικατοπτρίζει ένα επικριτικό και αγανακτισμένο (αρνητικό + κυριολεκτικό) αλλά ταυτόχρονα μορφωμένο (αμερόληπτο + μη προσβλητικό) κοινό.

Το σχήμα 4.5 απεικονίζει τον μέσο ημερήσιο 4Δ περιγραφέα για κάθε κόμμα στις εκλογές του 2020 στις ΗΠΑ. Και πάλι, οι χρονοσειρές διατηρούν τις ίδιες σταθερές ηλάσεις με τα δεδομένα του 2016 και στις τέσσερις διαστάσεις και σε όλο το χρονικό εύρος και για τα δύο κόμματα. Ένα ενδιαφέρον συμπέρασμα που μπορεί να εξαχθεί συγκρίνοντας τα δεδομένα των εκλογών του 2016 και των εκλογών του 2020 είναι ότι το κοινό φαίνεται να έχει μια μάλλον σταθερή στάση απέναντι στα ανταγωνιστικά κόμματα, που μεταφέρονται από τη μία εκλογική περίοδο στην άλλη, με τον νικητή να καθορίζεται από ένα μικρό περιθώριο/διαφορά.

Επιπλέον, ελέγχοντας οπτικά τα σχήματα 4.2, 4.4 και 4.3, 4.5 για τα σύνολα δεδομένων 2016 και 2020, αντίστοιχα, μπορούμε να εντοπίσουμε μια συσχέτιση μεταξύ της εμφάνισης κρίσιμων γεγονότων και των απότομων αλλαγών (αιχμές) στον αριθμό των tweets ή της κοινής γνώμης. Όπως ήταν αναμενόμενο, αυτή η αντίδραση στο twitter γίνεται εμφανής μια μέρα μετά το γεγονός.

Συγκρίνοντας τις χρονοσειρές των δύο κομμάτων στο σχήμα 4.4, μπορεί κανείς να παρατηρήσει ότι τα tweets για τους Ρεπουμπλικάνους είναι λιγότερο αρνητικά, λιγότερο αντικειμενικά και λιγότερο κυριολεκτικά, ενώ δεν υπάρχει ξεκάθαρη διαφορά μεταξύ των δύο κομμάτων όσον αφορά στην προσβλητικότητα. Αυτές οι παρατηρήσεις ρίχνουν νέο φως στα εκλογικά αποτελέσματα της 8ης Νοεμβρίου. Μια λιγότερο αρνητική γνώμη είναι ξεκάθαρα ένα πλεονέκτημα για τους Ρεπουμπλικάνους, αλλά ο συνδυασμός της με μια πιο μεροληπτική γνώμη αντανακλά την πιθανότητα να υπήρχαν περισσότεροι οπαδοί του Τραμπ ενεργοί στο twitter. Δεδομένου ότι οι οπαδοί ενός κόμματος είναι αποφασισμένοι ψηφοφόροι που δεν αλλάζουν εύκολα τη γνώμη τους, αυτά τα συμπεράσματα που προκύπτουν από την ανάλυση των κατασκευασμένων χρονοσειρών δίνουν την εικόνα ενός σημαντικού πλεονεκτήματος των Ρεπουμπλικανών, χρησιμοποιώντας μόνο δημόσια δεδομένα του twitter.

Είναι ενδιαφέρον ότι οι αναρτήσεις για τον Τραμπ φαίνεται να είναι λιγότερο κυριολεκτικές. Παρά τις κοινές αντιλήψεις ότι η μεταφορική γλώσσα χρησιμοποιείται συχνότερα για να εκφράσει αρνητικές απόψεις, τα tweets για τους Ρεπουμπλικάνους είναι κατά μέρος λιγότερο αρνητικά. Μια πιθανή εξήγηση είναι ότι η μεταφορικότητα δεν αντανακλά την

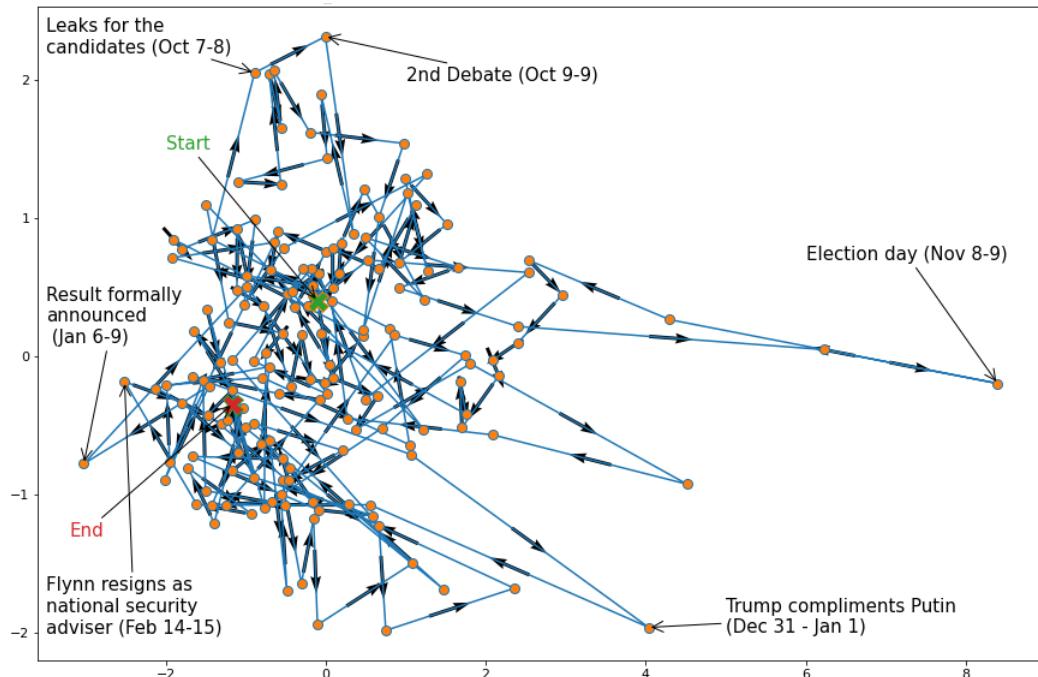


Σχήμα 4.6: Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Δημοκρατικούς, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2016 (163 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).

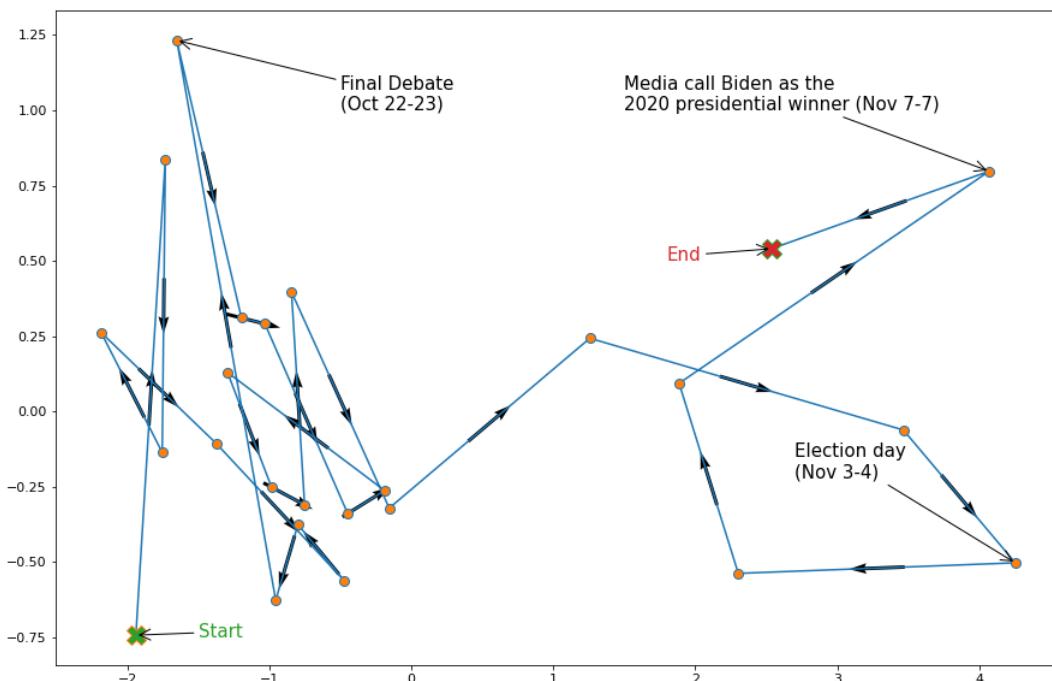
τελική απόφαση των ψηφοφόρων και χρησιμοποιείται κυρίως από τους χρήστες του twitter για να προσελκύσει μεγαλύτερη προσοχή. Επομένως, η ανάλυσή μας δείχνει ότι εάν ένας ψηφοφόρος είναι ξεκάθαρα εναντίον ενός υποψηφίου, είναι πιο πιθανό να είναι ευθύς/κυριολεκτικός στα σχόλιά του.

Αντίστοιχα συμπεράσματα μπορούν να εξαχθούν από το σχήμα 4.5 σχετικά με τις εκλογές του 2020 στις ΗΠΑ. Μπορεί κανείς να παρατηρήσει ότι τα tweets για τους Δημοκρατικούς είναι λιγότερο αρνητικά, λιγότερο αντικειμενικά, πιο κυριολεκτικά και λιγότερο προσβλητικά. Τα λιγότερο αρνητικά και λιγότερο αντικειμενικά χαρακτηριστικά μπορούν να ερμηνευθούν ως ευεργετικοί παράγοντες για το Δημοκρατικό κόμμα (όπως το 2016). Η κύρια διαφορά στα δεδομένα του 2020 είναι τα λιγότερο προσβλητικά και πιο κυριολεκτικά tweets που φαινομενικά συμβάλλουν περαιτέρω στη δημοκρατική κυριαρχία, καθώς η μεταφορική γλώσσα συνήθως υποδηλώνει αρνητικότητα [144].

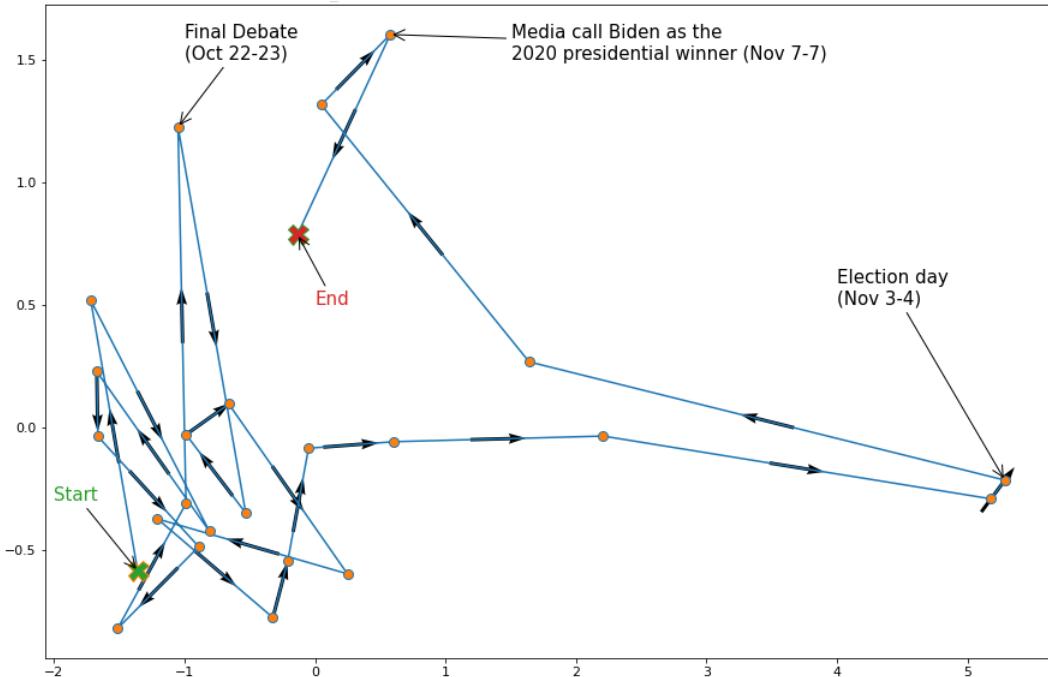
Η Ανάλυση Πρωτευουσών Συνιστωσών (ΑΠΣ) αξιοποιήθηκε για την μείωση διαστάσεων των 4Δ χρονοσειρών των Ρεπουμπλικάνων/Δημοκρατιών, έτσι ώστε να μπορούν να απεικονιστούν σε γραφικές παραστάσεις 2Δ. Τα σημεία περιγραφής 2Δ ανά κόμμα παρουσιάζονται στα σχήματα 4.6, 4.7 για το σύνολο δεδομένων 2016 και 4.8, 4.9 για το σύνολο δεδομένων 2020. Εδώ, τα απομακρυσμένα σημεία αντιστοιχούν στις αιχμές των αρχικών γραφημάτων συναρτήσει του χρόνου των σχημάτων 4.2, 4.4 και 4.3, 4.5. Έτσι, οι ακραίες τιμές στα σχήματα ΑΠΣ



Σχήμα 4.7: Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Ρεπουμπλικάνους, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2016 (163 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).



Σχήμα 4.8: Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Δημοκρατικούς, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2020 (25 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).

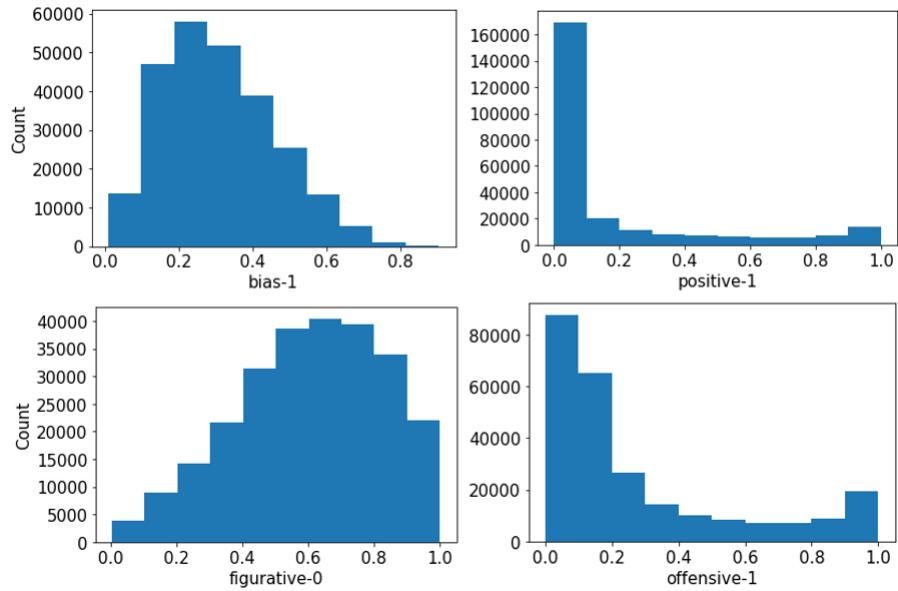


Σχήμα 4.9: Δισδιάστατη απεικόνιση βάσει ΑΠΣ των κατασκευασμένων 4Δ χρονοσειρών για τους Ρεπουμπλικάνους, χρησιμοποιώντας τη μέση στρατηγική συνάθροισης, σε ολόκληρο το χρονικό εύρος των δεδομένων του 2020 (25 ημέρες). Οι δύο ημερομηνίες που δίνονται ανά συμβάν είναι η ημερομηνία αυτού του συμβάντος (πρώτη) και η ημερομηνία της αντίστοιχης αντίδρασης στο twitter (δεύτερη).

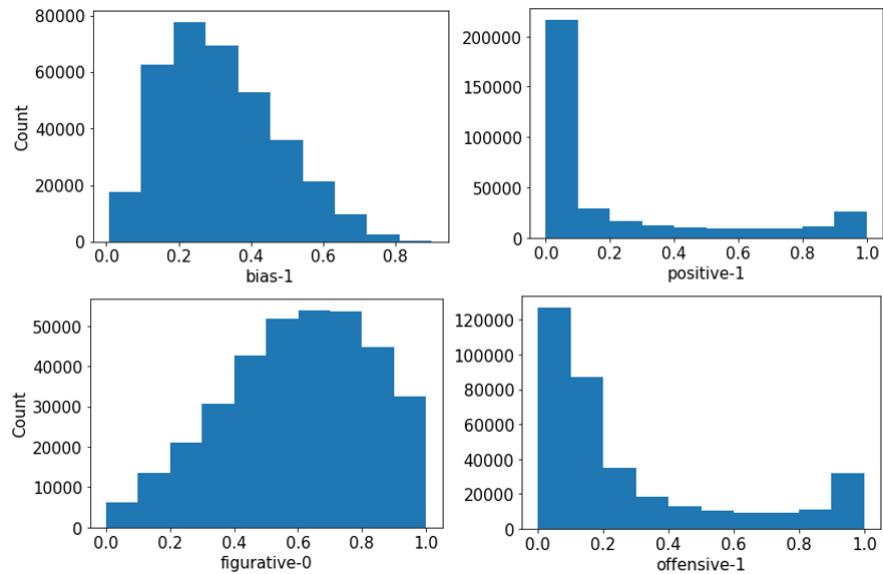
υποδεικνύουν την εμφάνιση κρίσιμων γεγονότων. Αυτή η οπτικοποίηση μπορεί να μας βοηθήσει να εντοπίσουμε περιστατικά που επηρεάζουν σημαντικά την κοινή γνώμη, αλλά δεν μας παρέχει αμέσως αντίστοιχες πληροφορίες σχετικά με τις καθεαυτές τιμές σημασιολογικής περιγραφής.

Ωστόσο, οι ακόλουθες παρατηρήσεις μπορούν να γίνουν με βάση τις ακραίες τιμές. Όσον αφορά τα δεδομένα του 2016, μπορούμε να πούμε από τα σχήματα 4.6, 4.7 ότι τα γεγονότα που επηρέασαν περισσότερο την κοινή γνώμη για τους Δημοκρατικούς ήταν οι ίδιες οι εκλογές, η επίσημη ανακοίνωση των αποτελεσμάτων και οι διαρροές σχετικά με τους υποψήφιους. Τα αντίστοιχα γεγονότα για τους Ρεπουμπλικάνους είναι πανομοιότυπα, με την προσθήκη του δεύτερου ντιμπέιτ υποψηφίων και τα κομπλιμέντα που έκανε ο Ντόναλντ Τραμπ στον Πρόεδρο Πούτιν. Αυτό πιθανώς αντανακλά την αυξημένη ανησυχία του κόσμου για την εθνική ασφάλεια. Όσον αφορά τα δεδομένα του 2020, τα σχήματα 4.8, 4.9 επιβεβαιώνουν τα τρία κρίσιμα συμβάντα που υποδεικνύονται από τις αιχμές στα γραφήματα 4.3, 4.5.

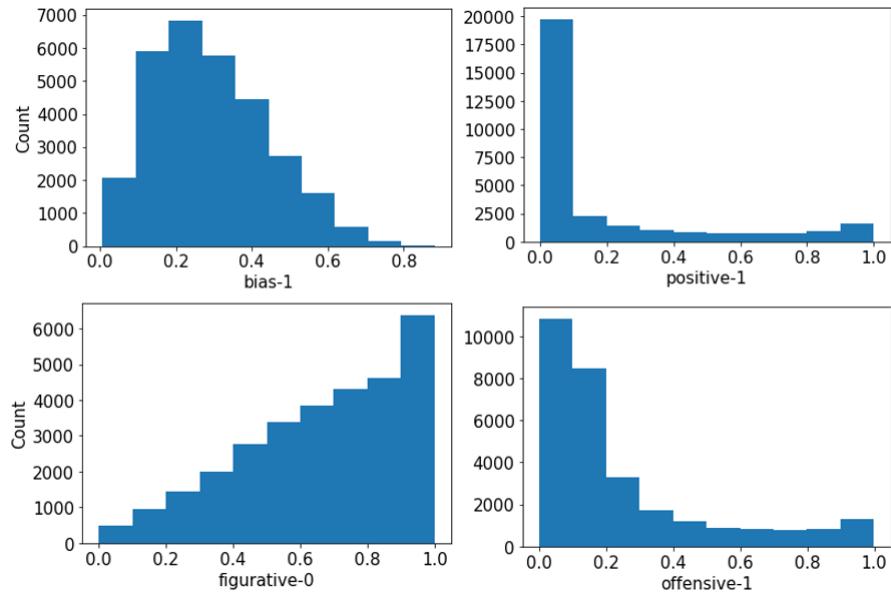
Δεδομένης της επεξηγηματικής ισχύος συγκεκριμένων ημερομηνιών που φαίνεται να είναι ακραία σημεία στις κατασκευασμένες χρονοσειρές, διερευνήθηκε επίσης ένας διαφορετικός τρόπος εκμετάλλευσης του προτεινόμενου μηχανισμού περιγραφής της κοινής γνώμης: η εστίαση σε μεμονωμένες σημαντικές ημερομηνίες. Στο πλαίσιο αυτής της μελέτης και δεδομένων των παρατηρήσεων που συζητήθηκαν προηγουμένως, η επομένη των εκλογών (9



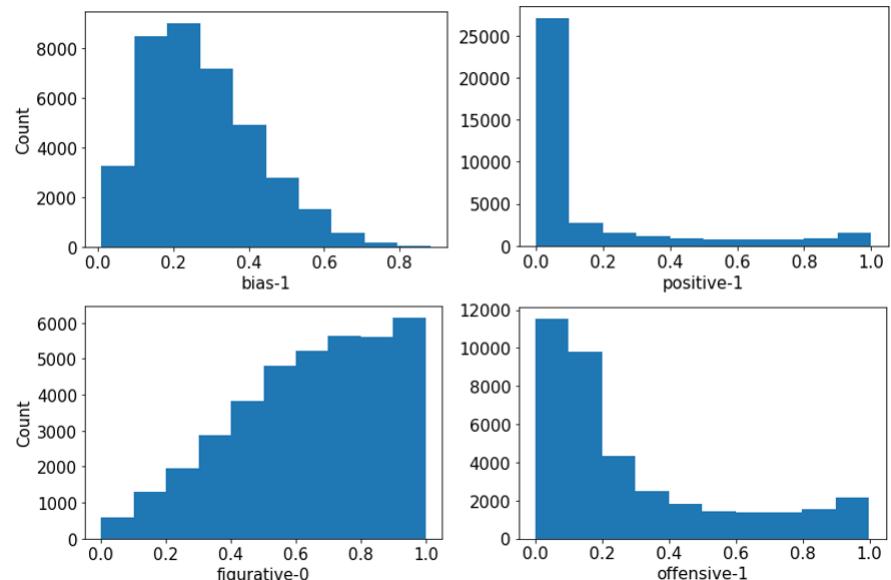
Σχήμα 4.10: Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Δημοκρατικούς στις 9 Νοεμβρίου 2016 (την επομένη των εκλογών).



Σχήμα 4.11: Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Ρεπουμπλικάνους στις 9 Νοεμβρίου 2016 (την επομένη των εκλογών).



Σχήμα 4.12: Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Δημοκρατικούς στις 4 Νοεμβρίου 2020 (την επομένη των εκλογών).



Σχήμα 4.13: Ιστογράμματα των τεσσάρων διαστάσεων του περιγραφέα, που απεικονίζουν τον τρόπο με τον οποίο κατανέμεται ο αριθμός των tweet στις εξόδους των ταξινομητών. Αυτά τα ιστογράμματα αφορούν τους Ρεπουμπλικάνους στις 4 Νοεμβρίου 2020 (την επομένη των εκλογών).

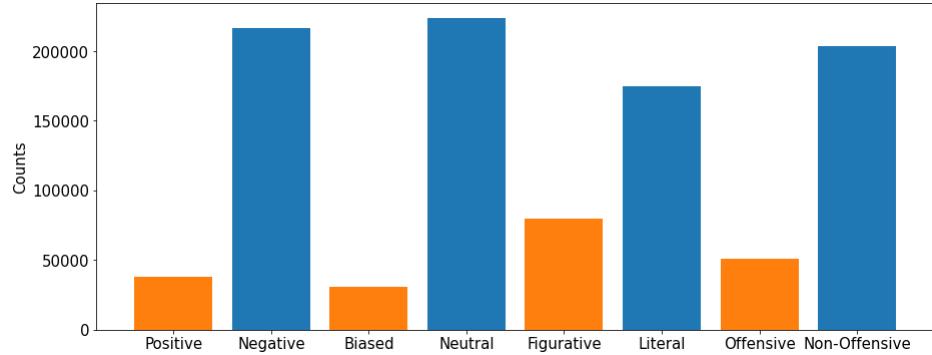
Νοεμβρίου 2016 και 4 Νοεμβρίου 2020) επιλέχθηκε ως ημερομηνία-στόχος.

Τα σχήματα 4.10 και 4.11 δείχνουν πώς τα tweets που δημοσιεύτηκαν στις 9 Νοεμβρίου 2016 κατανεμήθηκαν κατά μήκος των τεσσάρων διαστάσεων του περιγραφέα, ξεχωριστά για τα δύο κόμματα και πριν εφαρμοστεί οποιαδήποτε στρατηγική συνάθροισης. Αυτά τα ιστογράμματα των 10 bins δείχνουν την κατανομή του αριθμού των tweet (κάθετος άξονας) στις σημασιολογικές τιμές καθεμίας από τις τέσσερις διαστάσεις του περιγραφέα (οριζόντιος άξονας). Οι χρησιμοποιούμενες σημασιολογικές τιμές (που εξάγονται από τους 4 προεκπαιδευμένους ταξινομητές BNΔ) ήταν πραγματικοί αριθμοί στο διάστημα [0,1]. Επομένως, στον οριζόντιο άξονα καθένα από τα 10 βινς αντιστοιχεί σε ένα υποεύρος μήκους 0.1.

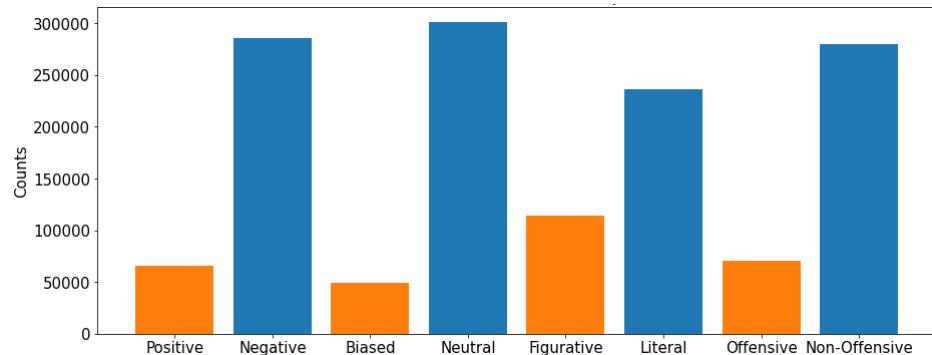
Τα ιστογράμματα είναι σχεδόν πανομοιότυπα για τους Δημοκρατικούς και τους Ρεπουμπλικάνους, μια παρατήρηση συμβατή με τη συμπεριφορά που καταγράφεται στο σχήμα 4.4. Επιπλέον, παρόμοια σχήματα ιστογράμματος μπορούν να διακριθούν για την μεροληψία-μεταφορικότητα και για τα χαρακτηριστικά πολικότητας-προσβλητικότητας. Η μεροληψία και η μεταφορικότητα είναι στην ουσία μετατοπισμένες κανονικές κατανομές, υπονοώντας ότι η στρατηγική μέσης συνάθροισης είναι πράγματι μια καλή επιλογή κατά την κατασκευή χρονοσειρών. Αντίθετα, τα ιστογράμματα πολικότητας και προσβλητικότητας είναι σημαντικά πιο πολωμένα, καθιστώντας τη στρατηγική μέσης συνάθροισης λιγότερο αξιόπιστη στην περίπτωσή τους για τη συγκεκριμένη ημερομηνία.

Το γεγονός ότι η πλειονότητα των tweets εμπίπτει στο διάστημα [0,0,2] για τις διαστάσεις πολικότητας και προσβλητικότητας, σημαίνει ότι έχουν σαφώς ταξινομηθεί ως αρνητικά και μη προσβλητικά. Ένα tweet πλήρως αρνητικό/θετικό και απολύτως μη προσβλητικό/προσβλητικό θα χαρακτηρίζεται από μια τιμή 0/1 και στις δύο διαστάσεις, αντίστοιχα. Αυτά τα ιστογράμματα σκιαγραφούν την εικόνα ενός κοινού που διαμαρτύρεται στον απόηχο των εκλογών, αλλά αποφεύγει τη χρήση προσβλητικής γλώσσας. Όσον αφορά τη σχετική απουσία ενδιάμεσων τιμών εντός του εύρους [0,2,0,8], μπορούμε να πούμε ότι η ταξινόμηση ως προς την πολικότητα και την προσβλητικότητα ήταν απλή και τα αντίστοιχα μοντέλα αρκετά σίγουρα. Αυτό σημαίνει ότι πράγματι τα περισσότερα tweets ήταν σαφώς αρνητικά ή θετικά, καθώς και σαφώς μη προσβλητικά ή προσβλητικά, χωρίς πολλοί χρήστες να είναι ουδέτεροι από αυτές τις απόψεις. Αντίθετα, η ταξινόμηση σχετικά με την μεροληψία και τη μεταφορικότητα δεν οδηγεί σε τέτοια πολωμένα αποτελέσματα. Αυτό οφείλεται στο γεγονός ότι τα μοντέλα BNΔ δυσκολεύονται να ταξινομήσουν αυτά τα tweets ως καθαρά στιγμιότυπα μιας συγκεκριμένης ηλάσης (π.χ., 'μεταφορικό' ή 'κυριολεκτικό'), που οδηγεί σε ενδιάμεσες τιμές κοντά στο 0.5. Αυτό σημαίνει ότι οι περισσότεροι χρήστες ήταν μάλλον ουδέτεροι σε σχέση με αυτές τις σημασιολογικές διαστάσεις. Ωστόσο, μπορούμε να δούμε ξεκάθαρα ότι η πλειοψηφία των tweets τείνουν να είναι αντικειμενικά και κυριολεκτικά.

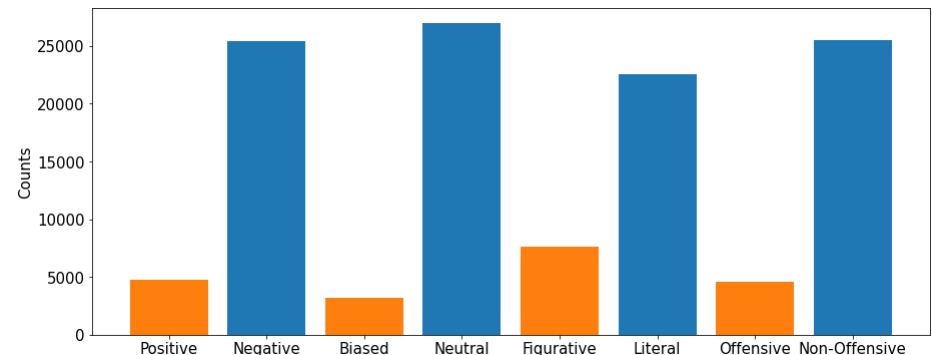
Παρόμοια σχήματα ιστογράμματος παρατηρούνται για τις αντίστοιχες διανομές tweet της 4ης Νοεμβρίου 2020 τόσο για τους Δημοκρατικούς 4.12 όσο και για τους Ρεπουμπλικάνους



Σχήμα 4.14: Αριθμός tweets που αφορούν τους Δημοκρατικούς, ξεχωριστά για κάθε κλάση των τεσσάρων διαστάσεων περιγραφής, στις 9 Νοεμβρίου 2016. Τα δύο χρώματα διακρίνουν τις αντίθετες κλάσεις κάθε σημασιολογικής διάστασης.

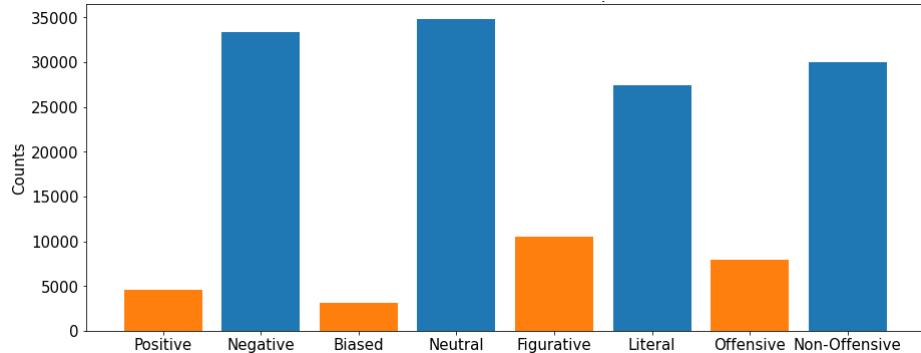


Σχήμα 4.15: Αριθμός tweets που αφορούν τους Ρεπουμπλικάνους, ξεχωριστά για κάθε κλάση των τεσσάρων διαστάσεων περιγραφής, στις 9 Νοεμβρίου 2016. Τα δύο χρώματα διακρίνουν τις αντίθετες κλάσεις κάθε σημασιολογικής διάστασης.



Σχήμα 4.16: Αριθμός tweets που αφορούν τους Δημοκρατικούς, ξεχωριστά για κάθε κλάση των τεσσάρων διαστάσεων περιγραφής, στις 4 Νοεμβρίου 2020. Τα δύο χρώματα διακρίνουν τις αντίθετες κλάσεις κάθε σημασιολογικής διάστασης.

4.13. Αυτό δεν ήταν έκπληξη δεδομένης της ομοιότητας των σχημάτων 4.4 και 4.5. Ωστόσο, υπάρχει μια αξιοσημείωτη διαφορά στη διάσταση μεταφορικότητας του περιγραφέα του 2016 και του 2020, όπου η κατανομή μετατοπίζεται προς τα δεξιά, υποδεικνύοντας πιο κυριολεκτική γλώσσα που χρησιμοποιείται στα tweet εκείνης της ημέρας. Αυτό μπορεί να επιβεβαιώθει συγκρίνοντας τα σχήματα 4.4 και 4.5.



Σχήμα 4.17: Αριθμός tweets που αφορούν τους Ρεπουμπλικάνους, ξεχωριστά για κάθε κλάση των τεσσάρων διαστάσεων περιγραφής, στις 4 Νοεμβρίου 2020. Τα δύο χρώματα διακρίνουν τις αντίθετες κλάσεις κάθε σημασιολογικής διάστασης.

Τέλος, οι εικόνες 4.14, 4.15 και 4.16, 4.17 απεικονίζουν τον αριθμό των tweet ανά κόμμα, ξεχωριστά για κάθε κλάση από τις τέσσερις διαστάσεις του περιγραφέα, τις ημέρες των εκλογών, 9 Νοεμβρίου 2016 και 4 Νοεμβρίου 2020. Αυτή η οπτικοποίηση παρέχει μια ματιά στις μη κυρίαρχες κλάσεις που εξαφανίστηκαν κατά την κατασκευή των χρονοσειρών χρησιμοποιώντας τη στρατηγική συνάθροισης μέσου όρουν.

4.6 Συζήτηση

Η αξιολόγηση που παρουσιάζεται στην ενότητα 4.5 υποδεικνύει ότι ο προτεινόμενος μηχανισμός για την αυτοματοποιημένη παρακολούθηση της κοινής γνώμης μέσω του twitter είναι ένα πολύ ισχυρό εργαλείο, μακόν να παρέχει πολύτιμες πληροφορίες για πιο αποτελεσματική λήψη αποφάσεων. Η **πολυδιάστατη φύση** του παρουσιαζόμενου περιγραφέα μεταφέρει πλούσιες ιδέες (αναλύονται στην Ενότητα 4.5) που συνήθως δεν αποτυπώνονται από τις υπάρχουσες σχετικές μεθόδους, οι οποίες εκμεταλλεύονται μόνο πολικότητα (και σπάνια μεροληψία). Αυτό φαίνεται μέσω των ποιοτικών πληροφοριών που εξάγονται στην Υποενότητα 4.5.3. Επιπλέον, σε αντίθεση με τη συντριπτική πλειονότητα των μεθόδων που έχουν δημοσιευτεί στο παρελθόν, ο προτεινόμενος μηχανισμός βασίζεται σε υπερσύγχρονα **εργαλεία ΕΦΓ βασισμένα σε αρχιτεκτονικές BNΔ**, γεγονός που εγγυάται βελτιωμένη ακρίβεια σε σύγκριση με τις υπάρχουσες συγκρίσιμες προσεγγίσεις.

Για να καταδειχθεί συνοπτικά η χρησιμότητα του προτεινόμενου μηχανισμού στην πολιτική ανάλυση, τα πιο σημαντικά ευρήματα που εξάγονται με την εφαρμογή του στα σύνολα δεδομένων της ενότητας 4.5 (σχετικά με τις προεδρικές εκλογές των ΗΠΑ του 2016 και του 2020) συνοψίζονται παρακάτω:

- Η κοινή γνώμη σχετικά με τους Δημοκρατικούς ήταν λιγότερο σταθερή και λιγότερο προβλέψιμη, σε σύγκριση με την κοινή γνώμη για τους Ρεπουμπλικάνους.

- Ωστόσο, το κοινό έχει γενικά μια σχετικά σταθερή στάση απέναντι στους ανταγωνιστές πολιτικούς: επικριτικό και αγανακτισμένο (αρνητικά + κυριολεκτικά στοιχεία) αλλά ταυτόχρονα μορφωμένο (αντικειμενικό + μη προσβλητικό).
- Οι χρονοσειρές που προκύπτουν μέσω του προτεινόμενου μηχανισμού δίνουν μια αρκετά ακριβή εικόνα του ευνοούμενου υποψηφίου. Αυτό φαίνεται μέσω της οπτικοποίησης/επιθεώρησης των χρονοσειρών.
- Το νικητήριο κόμμα αναδεικνύεται κατά την προεκλογική περίοδο σε tweets που είναι από κοινού λιγότερο αρνητικά + λιγότερο προσβλητικά + πιο μεροληπτικά. Η ισχυρή παρουσία κομματικών οπαδών στο twitter φαίνεται να συσχετίζεται σε μεγάλο βαθμό με τα υψηλά ποσοστά ψήφου.
- Τα κρίσιμα γεγονότα οδηγούν άμεσα σε απότομες αλλαγές στον ημερήσιο αριθμό των tweets (το οποίο είναι αναμενόμενο), αλλά και στην κοινή γνώμη. Αυτό δείχνει ότι οι αφοσιωμένοι/σταθεροί κομματικοί υποστηρικτές αποτελούν πάντα μειοψηφία στο αμερικανικό twitter. Επιπλέον, οι post hoc απεικονίσεις των χρονοσειρών που παράγονται αυτόματα μέσω του προτεινόμενου μηχανισμού μπορούν πραγματικά να δείξουν ποια γεγονότα ήταν τα πιο κρίσιμα για τις αλλαγές στην κοινή γνώμη.

Μερικά από αυτά τα ευρήματα επαληθεύουν παρόμοια συμπεράσματα που είχαν εξαχθεί προηγουμένως στην υπάρχουσα βιβλιογραφία: [9] (το κοινό έχει μια γενικά σχετικά σταθερή, αρνητική στάση απέναντι στους ανταγωνιστές πολιτικούς, κατά τη διάρκεια μιας συγκεκριμένης προεκλογικής περιόδου, ενώ οι χρονοσειρές για το δημόσιο αίσθημα δίνουν μια αρκετά ακριβή εικόνα του ευνοούμενου υποψηφίου) και [125] [12] [9] (κρίσιμα γεγονότα οδηγούν άμεσα σε απότομες αλλαγές στην κοινή γνώμη). Ωστόσο, η πλειονότητα των ευρημάτων μας για τις προεδρικές εκλογές των ΗΠΑ του 2016 και του 2020, όπως περιγράφονται λεπτομερώς στην Ενότητα 4.5, είναι πρωτότυπες συνεισφορές αυτής της μελέτης. Το πιο σημαντικό όμως είναι ότι ο προτεινόμενος μηχανισμός δεν συνδέεται με αυτές τις συγκεκριμένες εκλογές. Είναι μια πλήρως γενική και σχεδόν πλήρως αυτοματοποιημένη μέθοδος, η οποία επιτρέπει στους ενδιαφερόμενους χρήστες να εξάγουν εύκολα παρόμοιες πληροφορίες για οποιαδήποτε χρονική περίοδο.

4.7 Συμπεράσματα

Η αυτοματοποιημένη παρακολούθηση της κοινής γνώμης με χρήση των μέσων κοινωνικής δικτύωσης είναι ένα πολύ ισχυρό εργαλείο, ικανό να παρέχει στα ενδιαφερόμενα κόμματα πολύτιμες πληροφορίες για πιο καρποφόρα λήψη αποφάσεων. Το twitter έχει κερδίσει σημαντική προσοχή από αυτή την άποψη, καθώς οι άνθρωποι το χρησιμοποιούν για να

εκφράσουν τις απόψεις τους και οι πολιτικοί το χρησιμοποιούν για να προσεγγίσουν τους ψηφοφόρους τους.

Αυτή η μελέτη παρουσίασε έναν νέο, αυτοματοποιημένο μηχανισμό παρακολούθησης της κοινής γνώμης, που αποτελείται από έναν σύνθετο, ποσοτικό, σημασιολογικό περιγραφέα που βασίζεται σε αλγόριθμους ΕΦΓ. Ένα τετραδιάστατο διάνυσμα, δηλαδή, ένα παράδειγμα του προτεινόμενου περιγραφέα, εξάγεται πρώτα για κάθε tweet ανεξάρτητα, ποσοτικοποιώντας την πολικότητα, την προσβλητικότητα, την μεροληψία και την μεταφορικότητα του κειμένου. Στη συνέχεια, οι υπολογισμένοι περιγραφείς συνοψίζονται για πολλά tweets, σύμφωνα με μια επιθυμητή στρατηγική συνάθροισης (π.χ. αριθμητικός μέσος όρος) και στόχο συνάθροισης (π.χ. μια συγκεκριμένη χρονική περίοδο). Αυτό μπορεί να αξιοποιηθεί με διάφορους τρόπους. Για παράδειγμα, η συνάθροιση των tweet κάθε ημέρας ξεχωριστά μας επιτρέπει να κατασκευάσουμε μια πολυμεταβλητή χρονοσειρά που μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός αλγόριθμου πρόβλεψης τεχνητής νοημοσύνης, για καθημερινές προβλέψεις της κοινής γνώμης.

Προκειμένου να αξιολογηθεί η χρησιμότητα του προτεινόμενου μηχανισμού, εφαρμόστηκε στα μεγάλης κλίμακας σύνολα tweet για τις προεδρικές εκλογές των ΗΠΑ 2016/2020. Οι περιγραφείς της κοινής γνώμης που προέκυψαν χρησιμοποιήθηκαν με επιτυχία για την εκπαίδευση ενός μοντέλου πρόβλεψης κοινής γνώμης βασισμένο σε BNΔ με ορίζοντα πρόβλεψης 7 ημερών. Επιπλέον, οι κατασκευασμένες χρονοσειρές επιθεωρήθηκαν ενδελεχώς με ποιοτικό τρόπο, προκειμένου να συναχθούν πληροφορίες σχετικά με την κοινή γνώμη κατά τη διάρκεια μιας έντονης προεκλογικής περιόδου.

Διαθεσιμότητα Δεδομένων

Τα σύνολα δεδομένων που δημιουργήθηκαν κατά τη διάρκεια της τρέχουσας μελέτης είναι διαθέσιμα από τον αντίστοιχο συγγραφέα κατόπιν εύλογου αιτήματος. Το ακατέργαστο σύνολο δεδομένων tweet για τις προεδρικές εκλογές των ΗΠΑ 2016 είναι δημόσια διαθέσιμο στη διεύθυνση <https://www.kaggle.com/paulrohan2020/2016-usa-presidential-election-tweets> και το αντίστοιχο για το 2020 στη διεύθυνση <https://www.kaggle.com/datasets/manchunhui/us-election-2020-tweets>.

Κεφάλαιο 5

Μεταφορά Νευρωνικής Γνώσης για Βελτιωμένη Ανάλυση Συναισθημάτων σε Κείμενα με Μεταφορική Γλώσσα

5.1 Περίληψη

Η ΑΣ σε κείμενα, γνωστή και ως εξόρυξη γνώμης, είναι μια σημαντική εργασία ΕΦΓ, με πολλές εφαρμογές στην αυτοματοποιημένη παρακολούθηση μέσων κοινωνικής δικτύωσης, την επεξεργασία σχολίων πελατών, τη σάρωση e-mail, κ.λπ. Παρά την πρόσφατη πρόοδο λόγω της προόδου στα BNΔ, τα κείμενα που περιέχουν μεταφορική γλώσσα (π.χ. σαρκασμός, ειρωνεία, μεταφορές) εξακολουθούν να αποτελούν πρόκληση για τις υπάρχουσες μεθόδους λόγω των σημασιολογικών ασαφειών που περιέχουν. Σε αυτή τη μελέτη, προτείνεται μια νέα διάταξη νευρωνικής μεταφοράς γνώσης για ΑΣ που βασίζεται σε BNΔ αναγνώρισης μεταφορικών κειμένων. Χρησιμοποιείται για την απόσταξη της γνώσης από έναν προεκπαιδευμένο δυαδικό αναγνωριστή μεταφορικής γλώσσας σε έναν ταξινομητή συναισθημάτων πολλαπλών κλάσεων, ενώ ο τελευταίος εκπαιδεύεται σε μια ρύθμιση πολλαπλών εργασιών. Έτσι, οι υπαινιγμοί για την μεταφορικότητα βοηθούν έμμεσα στην επίλυση σημασιολογικών ασαφειών. Η αξιολόγηση σε ένα σχετικό δημόσιο σύνολο δεδομένων δείχνει ότι η προτεινόμενη μέθοδος οδηγεί σε ακρίβεια αιχμής, ξεπερνώντας όλες τις ανταγωνιστικές προσεγγίσεις.

5.2 Εισαγωγή

Η πρόσφατη πρόοδος στα BNΔ έχει αυξήσει τρομερά την απόδοση των μεθόδων αιχμής ΑΣ. Οι τρέχουσες προσεγγίσεις βασίζονται σε αρχιτεκτονικές BNΔ, όπως τα ΣΝΔ ή/και οι

παραλλαγές των ΔMBM, που έχουν εκπαιδευτεί για ταξινόμηση ή παλινδρόμηση κάτω από ένα περιβάλλον εποπτευόμενης μάθησης. Κάτω από αυτό το πλαίσιο, οι χειροποίητες αναπαραστάσεις κειμένου εισόδου έχουν αντικατασταθεί σε μεγάλο βαθμό από ενσωματώσεις λέξεων που “μαθαίνονται” από BNΔ [19] και καταφέρνουν να συλλάβουν σημαντικές σημασιολογικές ιδιότητες.

Ωστόσο, τα κείμενα που περιέχουν μεταφορική γλώσσα (π.χ. που περιλαμβάνουν σαρκασμό, ειρωνεία, μεταφορές κ.λπ.) εξακολουθούν να αποτελούν πρόκληση για τις υπάρχουσες μεθόδους [4]. Αυτό οφείλεται σε δυσκολίες στον προσδιορισμό του εάν μια μεταφορική φράση υπονοεί πράγματι αρνητική ή θετική γνώμη, δεδομένης της εγγενούς σημασιολογικής της ασάφειας. Το ζήτημα είναι ιδιαίτερα σημαντικό λόγω της πανταχού παρούσας φύσης της μεταφορικής γλώσσας στην ανθρώπινη επικοινωνία. Για παράδειγμα, ο σαρκασμός και η ειρωνεία είναι γνωστό ότι προσελκύουν μεγαλύτερη προσοχή στις διαδικτυακές κοινότητες [5], με πολλούς ανθρώπους να είναι έμπειροι στη χρήση τους και να τα εκμεταλλεύονται για να αυξήσουν τον αντίκτυπο των αναρτήσεών τους στα μέσα κοινωνικής δικτύωσης (π.χ. στο Twitter).

Σε αυτή τη μελέτη, προτείνεται μια νέα ρύθμιση της απόσταξης γνώσης [6] για ΑΣ βάσει BNΔ μεταφορικών κειμένων [4]. Χρησιμοποιείται για τη μεταφορά γνώσης από έναν δυαδικό προεκπαιδευμένο αναγνωριστή μεταφορικής γλώσσας (δάσκαλος) σε έναν ταξινομητή συναισθημάτων πολλαπλών κλάσεων (μαθητής), ενώ ο τελευταίος εκπαιδεύεται σε μια ρύθμιση πολλαπλών εργασιών. Μέχρι τώρα, η απόσταξη γνώσης εφαρμόζεται κυρίως σε ρυθμίσεις ταξινόμησης που βασίζονται στη softmax ενεργοποίηση, όπου ο δάσκαλος και ο μαθητής επιλύουν πανομοιότυπες εργασίες. Αυτή η μελέτη, αντίθετα, προτείνει την απόσταξη από έναν δάσκαλο δυαδικής ταξινόμησης με σιγμοειδές αποτέλεσμα, προκειμένου να βοηθήσει έναν μαθητή ταξινόμησης πολλαπλών κλάσεων. Από όσο γνωρίζουμε, είναι η πρώτη φορά που χρησιμοποιείται αυτή η προσέγγιση για ένα πραγματικό πρόβλημα με τη χρήση κανονικών BNΔ. Επιπλέον, είναι η πρώτη φορά που αξιοποιείται οποιαδήποτε μορφή νευρωνικής απόσταξης για την αύξηση της ακρίβειας της ΑΣ σε κείμενα με μεταφορική γλώσσα. Η αξιολόγηση σε ένα σχετικό δημόσιο σύνολο δεδομένων δείχνει ότι η προτεινόμενη μέθοδος οδηγεί σε επιδόσεις αιχμής, ξεπερνώντας όλες τις ανταγωνιστικές προσεγγίσεις.

5.3 Σχετική Βιβλιογραφία

Αυτή η ενότητα παρουσιάζει την υφιστάμενη αιχμή της τεχνολογίας που αφορά: α) Ανάλυση Συναισθήματος σε κείμενα με Μεταφορική Γλώσσα, και β) Απόσταξη Γνώσης.

5.3.1 Ανάλυση συναισθήματος σε Μεταφορική Γλώσσα

Η ΑΣ σε μεταφορική γλώσσα προτάθηκε αρχικά στο διαγωνισμό Semantic Evaluation 2015 - Task 11 (S15-T11) [4]. Δεδομένου ενός συνόλου tweets που είναι πλούσια σε μεταφορά, σαρκασμό και ειρωνεία, ο στόχος ήταν να προσδιοριστεί εάν ένας χρήστης έχει εκφράσει θετικό, αρνητικό ή ουδέτερο συναίσθημα σε καθένα από αυτά. Στους συμμετέχοντες δόθηκαν τόσο ακέραιες όσο και επικέτες πραγματικών τιμών, επιτρέποντας έτσι τη χρήση προσεγγίσεων είτε ταξινόμησης είτε παλινδρόμησης. Όλες οι δημοσιευμένες μελέτες που σχετίζονται με την ΑΣ σε μεταφορική γλώσσα αξιολογούνται στη συλλογή tweet S15-T11, λόγω της έλλειψης άλλων σχετικών δημόσιων συνόλων δεδομένων.

Οι πρώτες μέθοδοι της βιβλιογραφίας για την αντιμετώπιση του προβλήματος αυτού βασίζονταν σε ΜΔΥ, δέντρα αποφάσεων ή μοντέλα παλινδρόμησης, που λειτουργούσαν σε χειροποίητα χαρακτηριστικά (π.χ. σύνολα λέξεων ή ΣΟ-ΑΣΕ) και λεξικά (π.χ. SentiWordnet, Depeche Mood, American National Corpus, κ.λπ.) [15] [16]. Αργότερα, το [17] προσέγγισε το πρόβλημα ως πρόβλημα παλινδρόμησης και εκμεταλλεύτηκε μια αρχιτεκτονική ΣΝΔ. Οι λέξεις με περισσότερες από δύο εμφανίσεις χρησιμοποιήθηκαν ως κειμενικά χαρακτηριστικά εισόδου, ενώ τα hashtags που προσδιορίζουν το είδος της μεταφορικής έκφρασης αντικαταστάθηκαν από δυαδικούς δείκτες. Το DESC [18] ενσωμάτωσε ένα AM-ΔMBM, ένα ΔMBM με μηχανισμό προσοχής και ένα MLP, με τα δύο πρώτα να τροφοδοτούνται με προεκπαιδευμένες ενσωματώσεις λέξεων GloVe [19] και το MLP να τροφοδοτείται με χειροποίητα χαρακτηριστικά tweet, όπως ζεύγη λέξεων και τιμές ΣΟ-ΑΣΕ. Το DESC επεκτάθηκε στο [20] χρησιμοποιώντας πλαισιωμένες ενσωματώσεις εισόδου από ένα προεκπαιδευμένο μοντέλο RoBERTa [21], οι οποίες τροφοδοτούνται σε ένα επαναλαμβανόμενο συνελικτικό νευρωνικό δίκτυο (ΕΣΝΔ). Το πρόβλημα προσεγγίστηκε ως ταξινόμηση κειμένου.

5.3.2 Μεταφορά Νευρωνικής Γνώσης

Η απόσταξη γνώσης είναι η πιο διαδεδομένη μορφή νευρωνικής μεταφοράς γνώσης. Συνεπάγεται την εκπαίδευση ενός μαθητή BNΔ σύμφωνα με την απόκριση ενός προεκπαιδευμένου BNΔ (δάσκαλος) σε κάποια δεδομένα εισόδου (το λεγόμενο «σύνολο δεδομένων μεταφοράς»), που συνήθως είναι μη επισημασμένα. Συνήθως, ο μαθητής θεωρείται ότι είναι αρχικά μη εκπαιδευμένος, ενώ τα δύο δίκτυα μπορεί να έχουν διαφορετικές νευρωνικές αρχιτεκτονικές [6]. Το μαθητικό BNΔ εκπαιδεύεται με μια αντικειμενική συνάρτηση κόστους που μετρά την απόκλιση της εξόδου του από την αντίστοιχη έξοδο του αμετάβλητου δασκάλου BNΔ στο σύνολο μεταφοράς [6]. Ο στόχος είναι η μεταφορά της λεγόμενης σκοτεινής γνώσης από τον δάσκαλο στον μαθητή.

Στις εργασίες ταξινόμησης, συνήθως χρησιμοποιούνται ασυνήθιστα υψηλές θερμοκρασίες softmax για την εξαγωγή και των δύο εξόδων κατά την εκπαίδευση του μαθητή, προκειμένου να αποκτηθούν μαλακές κατανομές για τις ιλάσεις που κωδικοποιούν έμμεσα ομοιότητες

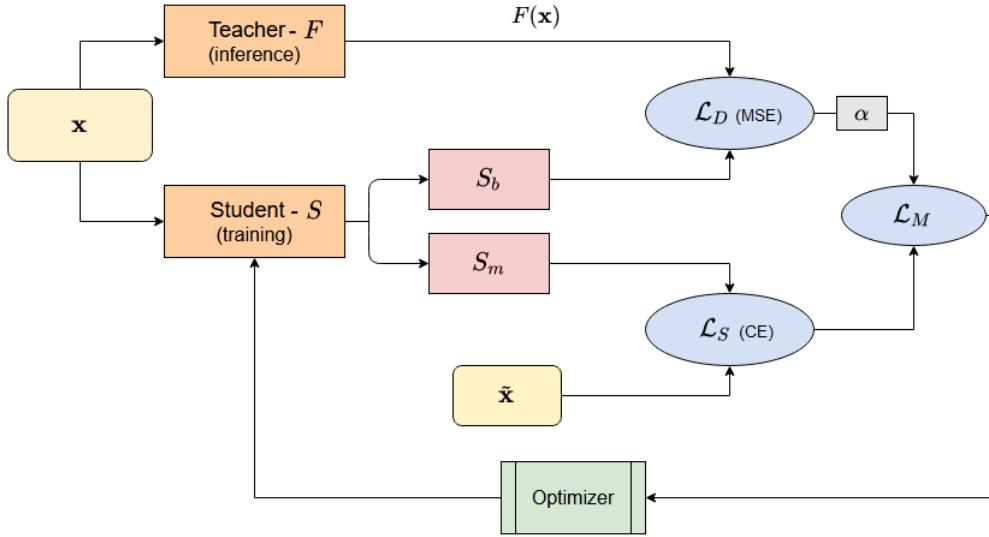
μεταξύ δειγμάτων δεδομένων εισόδου. Μετά την εκπαίδευση, ο μαθητής αναπτύσσεται με μια πιο συμβατική θερμοκρασία softmax. Εάν οι ετικέτες είναι διαθέσιμες για δυνατότητα επίβλεψης της εκπαίδευσης στο σύνολο μεταφοράς, μπορεί επίσης να προστεθεί ένας κανονικός όρος ταξινόμησης στην αντικειμενική συνάρτηση κόστους. Στην περίπτωση αυτή, ο όρος απόσταξης έχει κυρίως λειτουργία κανονικοποιητή.

Αρκετές παραλλαγές αυτού του βασικού συστήματος νευρωνικής απόσταξης έχουν εμφανιστεί όλα αυτά τα χρόνια. Για παράδειγμα, σε περίπτωση που το σύνολο μεταφοράς είναι γνωστό/επισημασμένο, οι παράμετροι/βάρη του εκπαιδευμένου δασκάλου μπορούν να χρησιμοποιηθούν για την έξυπνη αρχικοποίηση του μαθητή πριν προχωρήσουμε στην τακτική εποπτευόμενη εκπαίδευση [22]. Πληροφορία/χαρακτηριστικά από ενεργοποιήσεις ενδιάμεσων ιρυφών επιπέδων στον δάσκαλο κατά την επεξεργασία του συνόλου μεταφοράς μπορούν επίσης να χρησιμοποιηθούν για την καθοδήγηση της εκπαίδευσης των μαθητών, εκτός από το τελικό αποτέλεσμα του επιπέδου εξόδου του δασκάλου [23]. Μπορεί να χρειαστεί μείωση διαστάσεων όταν οι διαστάσεις του επιπέδου διαφέρουν μεταξύ μαθητή και δασκάλου. Για να παρακαμφθεί αυτός ο περιορισμός, η εκπαίδευση των μαθητών στο [24] καθοδηγείται από τις ομοιότητες μεταξύ των αναπαραστάσεων δειγμάτων δεδομένων συνόλου μεταφοράς που κατασκευάζονται από τον δάσκαλο. Ενδιάμεσα δίκτυα «βοηθού δασκάλου», στο πλαίσιο μιας διαδικασίας πολλαπλών βημάτων, προτάθηκαν επίσης στο [25] για σενάρια όπου τα δύο μοντέλα διαφέρουν ως προς την πολυπλοκότητα. Τέλος, παρουσιάστηκε απόσταξη από έναν βαθύ γραμμικό δάσκαλο για δυαδική ταξινόμηση στο πλαίσιο μιας καθαρά θεωρητικής ανάλυσης στο [26], ενώ παρόμοια έρευνα πραγματοποιήθηκε στο [27] για ρηχά νευρωνικά δίκτυα.

5.4 Προτεινόμενη Μέθοδος

Η προτεινόμενη μέθοδος εκμεταλλεύεται την απόσταξη γνώσης [6] για την αύξηση της ακρίβειας ενός BNΔ κατά την ΑΣ σε μεταφορικά κείμενα, που περιέχουν σαρκασμό, ειρωνεία ή/και μεταφορά. Έτσι, κατά τη διάρκεια της εκπαίδευσης, χρησιμοποιείται μια αρχιτεκτονική δασκάλου-μαθητή για να εμπλουτιστεί το μοντέλο του μαθητή με τη γνώση ενός προεκπαιδευμένου αναγνωριστή μεταφορικής γλώσσας. Ο τελευταίος είναι ένας δυαδικός ταξινομητής, ενώ ο πρώτος είναι ένας ταξινομητής πολλαπλών κλάσεων που έχει ως αποστολή την αναγνώριση συναισθημάτων στα εισερχόμενα κείμενα. Έτσι, λόγω της φύσης του δασκάλου και της διαφορετικής εργασίας που επιλύει σε σύγκριση με τον μαθητή, προτείνεται ένα άτυπο είδος απόσταξης που δεν έχει χρησιμοποιηθεί προηγουμένως σε κανονικά BNΔ για πραγματικά προβλήματα.

Παρακάτω, όλα τα νευρωνικά μοντέλα εννοείται ότι εκπαιδεύονται με την οπίσθια διάδοση σφαλμάτων και μια παραλλαγή της καθόδου στοχαστικής κλίσης. Ας υποθέσουμε επίσης ότι ένας δυαδικός ταξινομητής κειμένου που βασίζεται σε BNΔ F έχει προεκπαιδευτεί κάτω



Σχήμα 5.1: Η προτεινόμενη αρχιτεκτονική εκπαίδευσης δασκάλου-μαθητή.

από μια κανονική εποπτευόμενη ρύθμιση σε μια βάση δεδομένων που περιέχει δύο ιλάσεις: "μεταφορικό" και "κυριολεκτικό". Δεδομένου ότι είναι σύνηθες οι δυαδικοί νευρωνικοί ταξινομητές να τελειώνουν με έναν μόνο σιγμοειδή νευρώνα, υποθέτουμε ότι αυτό ισχύει για το F . Έτσι, μια βαθμωτή έξοδος πραγματικής τιμής 0/1 αντιστοιχεί σε μεταφορική/κυριολεκτική πρόβλεψη ιλάσης, αντίστοιχα, ενώ μια τυπική έξοδος $F(\mathbf{x})$ για ένα αντίστοιχο σημείο δεδομένων εισόδου \mathbf{x} θα βρίσκεται στην πραγματικότητα στο διάστημα $[0, 1]$.

Ο μαθητής S είναι το νευρωνικό μοντέλο που θέλουμε να βελτιστοποιήσουμε, σε ένα διαφορετικό σύνολο δεδομένων επισημασμένο με τα σχετικά συναισθήματα. Χωρίς απώλεια γενικότητας, υποθέτουμε ότι εκπαιδεύεται κάτω από μια εποπτευόμενη ρύθμιση ταξινόμησης κειμένου πολλαπλών ιλάσεων. Συνήθως, $N \geq 3$ ιλάσεις χρησιμοποιούνται για την εργασία ανάλυσης συναισθήματος/εξόρυξης απόψεων (θετικό, αρνητικό, ουδέτερο) και ένα τελικό επίπεδο ενεργοποίησης softmax χρησιμοποιείται για την εξαγωγή της πρόβλεψης ιλάσης. Το S εκπαιδεύεται από μια κανονική, κατάλληλη συνάρτηση κόστους \mathcal{L}_S , όπως η Διασταυρούμενη Εντροπία (ΔE).

Η προτεινόμενη μέθοδος συνίσταται στην εκπαίδευση του S με την ακόλουθη συνάρτηση κόστους πολλαπλών εργασιών:

$$\mathcal{L}_M = \mathcal{L}_S + \alpha \mathcal{L}_D, \quad (5.1)$$

όπου το \mathcal{L}_D υπολογίζεται σε κάθε επανάληψη με εκμετάλλευση του προεκπαιδευμένου F . Στην ουσία, το \mathcal{L}_D αποστάζει τη γνώση του F σχετικά με το τρέχον σημείο δεδομένων εκπαίδευσης \mathbf{x} , δηλαδή $F(\mathbf{x})$. Όπως σημειώνεται στο [26] για το βαθύ γραμμικό σενάριο, η σιγμοειδής ενεργοποίηση εξόδου για την περίπτωση δυαδικής ταξινόμησης είναι ισοδύναμη με τις μαλακές ετικέτες που χρησιμοποιούνται συνήθως για απόσταξη πολλαπλών ιλάσεων

βάσει softmax ενεργοποίησης [6]. Για να υπολογίσουμε αυτόν τον όρο κόστους, ένα παράλληλο επίπεδο εξόδου S_b που χρησιμεύει ως βοηθητική κεφαλή δυαδικής ταξινόμησης είναι αρχιτεκτονικά συνδεδεμένο στο προτελευταίο επίπεδο του S , ενώ το $F(\mathbf{x})$ αποτελεί τις επιθυμητές εξόδους (πραγματικών τιμών) για το \mathcal{L}_D . Για να αποφευχθεί η σύγχυση, η κανονική (softmax ενεργοποίησης) κεφαλή ταξινόμησης πολλαπλών κλάσεων του S συμβολίζεται παρακάτω με S_m . Έτσι, υποθέτοντας N κλάσεις συναισθήματος, το S_m/S_b είναι ένα νευρωνικό επίπεδο εξόδου που αποτελείται από $N/1$ νευρώνες, αντίστοιχα.

Χρησιμοποιώντας το μέσο τετραγωνικό σφάλμα (ΜΤΣ) για κόστος \mathcal{L}_D , η προτεινόμενη πλήρης συνάρτηση κόστους πολλαπλών εργασιών είναι:

$$\mathcal{L}_M = \mathcal{L}_S(S_m(\mathbf{x}), \tilde{\mathbf{x}}) + \alpha (S_b(\mathbf{x}) - F(\mathbf{x}))^2, \quad (5.2)$$

όπου $\tilde{\mathbf{x}}$ είναι η πραγματική κλάση συναισθήματος (αντικειμενική αλήθεια) που αντιστοιχεί στο \mathbf{x} , στο πλαίσιο της ταξινόμησης πολλαπλών κλάσεων. Είναι αξιοσημείωτο ότι δεν χρησιμοποιείται εικέτα αντικειμενικής αλήθειας σχετικά με τη μεταφορικότητα του \mathbf{x} .

Μια επισκόπηση της προτεινόμενης μεθόδου απεικονίζεται στο σχήμα 5.1. Είναι σημαντικό ότι δεν απαιτείται ή αξιοποιείται κανένας σχολιασμός γενικής αλήθειας (επισημασμένα δεδομένα) σχετικά με την παρουσία ή τον τύπο μεταφορικής γλώσσας κατά την εκπαίδευση του S για ανάλυση συναισθήματος. Φυσικά, αφού το S έχει εκπαιδευτεί πλήρως, τόσο ολόκληρο το μοντέλο F όσο και το βοηθητικό επίπεδο εξόδου-κεφαλή δυαδικής ταξινόμησης S_b μπορούν να αφαιρεθούν με ασφάλεια.

Η υποκείμενη διαίσθηση πίσω από την προτεινόμενη συνάρτηση κόστους πολλαπλών εργασιών είναι η εικασία ότι η σκοτεινή γνώση σχετικά με τον βαθμό μεταφορικότητας ενός κειμένου εισόδου θα πρέπει να βοηθήσει έναν ταξινομητή συναισθημάτων στην επίλυση ασαφειών σχετικά με το εκφραζόμενο συναίσθημα, που προκύπτουν λόγω σαρκαστικής, μεταφορικής ή ειρωνικής γλώσσας. Ο προτεινόμενος όρος κόστους απόσταξης θα πρέπει να έχει ως αποτέλεσμα τη ρύθμιση του ταξινομητή συναισθήματος πολλαπλών κλάσεων προς τον εντοπισμό και την αντιμετώπιση τέτοιων ασαφειών. Ο αναγνωριστής μεταφορικής γλώσσας F επιλέχθηκε ως δυαδικός ταξινομητής προκειμένου να μεγιστοποιήσει το ποσοστό επιτυχίας του σταδίου συμπερασμάτων σε αυτή τη βοηθητική εργασία, κάνοντας το πρόβλημα ταξινόμησης όσο το δυνατόν πιο εύκολο. Σημειώτεον, η απόσταξη γνώσης από δυαδικούς ταξινομητές με σιγμοειδή έξοδο, προκειμένου να βοηθήσει έναν ταξινομητή πολλαπλών κλάσεων σε διαφορετική εργασία, δεν έχει παρουσιαστεί προηγουμένως για BNΔ.

5.5 Ποσοτική Αξιολόγηση

5.5.1 Λεπτομέρειες Υλοποίησης

Η νευρωνική αρχιτεκτονική ROB-ΕΣΝΔ από το [20] αναδημιουργήθηκε και υιοθετήθηκε για το βασικό μοντέλο ανάλυσης συναισθήματος μαθητή S . Αυτή η νευρωνική αρχιτεκτονική χρησιμοποιεί ένα προεκπαιδευμένο μοντέλο γλώσσας RoBERTa [21], σε συνδυασμό με ένα ΕΣΝΔ [145], προκειμένου να συλλάβει αποτελεσματικά τις πληροφορίες κειμένου κατά την αναπαράσταση κάθε λέξης. Η τελική πρόβλεψη είναι η έξοδος ενός επιπέδου softmax ενεργοποίησης. Ο λόγος πίσω από την επιλογή του ROB-ΕΣΝΔ ως βάση ήταν αποκλειστικά πρακτικός. *Η προτεινόμενη μέθοδος μπορεί να χρησιμοποιηθεί για να ενισχύσει και οποιονδήποτε άλλο ταξινομητή συναισθήματος.*

Πίνακας 5.1: Αποτελέσματα αξιολόγησης στο σύνολο δεδομένων S15-T11. Το υψηλότερο-χαμηλότερο είναι καλύτερο για τη μέτρηση COS/MTΣ, αντίστοιχα. Τα καλύτερα αποτελέσματα είναι με έντονους χαρακτήρες.

Μέθοδος	COS	MTΣ
ELMo [146]	0.71	3.61
USE [147]	0.71	3.17
NBSVM [148]	0.69	3.23
FastText [149]	0.72	2.99
XLnet [150]	0.76	1.84
BERT-Cased [151]	0.72	1.97
BERT-Uncased [151]	0.79	1.54
RoBERTa [21]	0.78	1.55
UPF [15]	0.71	2.46
ClaC [152]	0.76	2.12
DESC [18]	0.82	2.48
ROB-RCNN [20]	0.82	1.92
ROB-RCNN + Proposed (\mathcal{L}_D)	0.85	1.50

Η νευρωνική αρχιτεκτονική ΣΝΔ/ΑΜ-ΔΜΒΜ OSLCfit [153] προεκπαιδεύτηκε για αναγνώριση μεταφορικής γλώσσας, αικιδευτικών τεκμηρίων και συναρπαστικών λέξεων. Το μοντέλο δασκάλου δυαδικής ταξινόμησης F που ορίζεται στο [153] και στη συνέχεια υιοθετήθηκε ως το μοντέλο δασκάλου δυαδικής ταξινόμησης F . Οι αναπαραστάσεις του κειμένου εισαγωγής προέρχονται από τη χρήση ενσωματώσεων 200 διαστάσεων από ένα προεκπαιδευμένο μοντέλο GloVe [19].

Αυτό το μοντέλο δασκάλου ήταν προεκπαιδευμένο στο επισημασμένο σύνολο δεδομένων από το [154], το οποίο περιέχει 81,4K tweets ομαδοποιημένα σε 4 διαφορετικές ετικέτες κλάσης (σαρκασμός, ειρωνεία, μεταφορά και κυριολεξία). Οι τρεις πρώτες κλάσεις συνδυάστηκαν στην γενική κλάση “μεταφορικό”, προκειμένου να εκπαιδεύσουμε το F ως δυαδικό ταξινομητή μεταφορικού κειμένου.

Ο μαθητής S εκπαιδεύτηκε χρησιμοποιώντας τον βελτιστοποιητή Adam και την ΔΕ ως την κύρια συνάρτηση κόστους ταξινόμησης πολλαπλών τάξεων \mathcal{L}_S του μαθητή.

5.5.2 Πλαίσιο Αξιολόγησης

Το σύνολο δεδομένων του διαγωνισμού S15-T11 [4] χρησιμοποιήθηκε για την αξιολόγηση της προτεινόμενης μεθόδου και τη σύγκρισή της με ανταγωνιστικές προσεγγίσεις. Περιέχει 8000/4000 tweets για εκπαίδευση/δοκιμή, αντίστοιχα, συμπεριλαμβανομένων tweets με ειρωνική, σαρκαστική και μεταφορική γλώσσα. Τα 12.000 σημεία δεδομένων ομαδοποιούνται σε 11 κατηγορίες επισημασμένες με ακέραιους αριθμούς σε κλίμακα 11 σημείων, που κυμαίνονται από -5 έως +5, που υποδηλώνουν την πολικότητα κάθε tweet, από "πολύ αρνητικό" σε "πολύ θετικό". Δεδομένου ότι είναι ένα σύνολο δεδομένων ανάλυσης συναισθήματος, δεν περιέχει σχολιασμούς/ετικέτες σχετικά με την παρουσία ή τον τύπο μεταφορικής γλώσσας.

Χρησιμοποιήθηκαν δύο μετρικές αξιολόγησης: ομοιότητα συνημιτόνου (COS, υψηλότερο είναι καλύτερο) και μέσο τετραγωνικό σφάλμα (ΜΤΣ, χαμηλότερο είναι καλύτερο). Υποθέτοντας ένα σύνολο σημείων δεδομένων δοκιμής T , και οι δύο υπολογίζονται συγκρίνοντας δύο T -διάστατα ακέραια διανύσματα, που περιέχουν τις εξόδους/προβλέψεις του εκπαιδευόμενου μοντέλου και τις ετικέτες αντικειμενικής αλήθειας αντίστοιχα. Η προτεινόμενη υλοποίηση της μεθόδου είναι στην πραγματικότητα το ROB-ΕΣΝΔ [20] επαυξημένο με \mathcal{L}_D κατά τη διάρκεια της εκπαίδευσης, ενώ η βασική γραμμή που βελτιώνουμε είναι το ROB-ΕΣΝΔ εκπαιδευμένο με απλό \mathcal{L}_S , αντί της προτεινόμενης Εξ. (5.2). Οι βέλτιστες υπερπαράμετροι υιοθετήθηκαν από το [20], ενώ η διασταύρομενη επικύρωση 5 διπλώσεων είχε ως αποτέλεσμα την καλύτερη τιμή $\alpha = 0,5$. Τα αποτελέσματα της αξιολόγησης της δοκιμαστικής φάσης παρουσιάζονται στον Πίνακα 5.1, συμπεριλαμβανομένης της ακρίβειας που επιτυγχάνεται με διάφορες ανταγωνιστικές μεθόδους. Όλα τα αναφερόμενα νούμερα ακρίβειας παρουσιάζονται και στο [20], εκτός από αυτά για το ROB-ΕΣΝΔ. Η τελευταία μέθοδος αναδημιουργήθηκε, εκπαιδεύτηκε και αξιολογήθηκε εξαρχής από εμάς, ακολουθώντας αυστηρά όλες τις λεπτομέρειες εφαρμογής και τις τιμές υπερπαραμέτρων που περιγράφονται στο [20].

Συνολικά, η προτεινόμενη υλοποίηση μεθόδου ROB-ΕΣΝΔ + \mathcal{L}_D επιτυγχάνει κορυφαίες επιδόσεις και στις δύο μετρικές, επιβεβαιώνοντας ότι την εγκυρότητα της υποκείμενης διαίσθησής μας. Αξιοσημείωτα, σε σύγκριση με το DESC, καταφέρνει να μειώσει το ΜΤΣ από 2,48 σε 1,50, ενώ ταυτόχρονα αυξάνει το COS από 0,82 σε 0,85. Αντίθετα, το βασικό ROB-ΕΣΝΔ επιτυγχάνει βελτιώσεις ΜΤΣ έναντι του DESC, χωρίς να βελτιώσει το COS.

5.6 Συμπεράσματα

Οι εργασίες ΕΦΓ, όπως η ΑΣ σε κείμενα, έχουν προχωρήσει σημαντικά χάρη στην πρόοδο στα ΒΝΔ. Ωστόσο, οποιαδήποτε παρουσία μεταφορικής γλώσσας (σαρκασμός, μεταφορά,

ειρωνεία) αυξάνει σημαντικά τη δυσκολία του έργου της ΑΣ. Αυτή η μελέτη εκμεταλλεύεται τη διαίσθηση ότι οι εκτιμήσεις σχετικά με την ύπαρξη μεταφορικής γλώσσας σε ένα κείμενο εισαγωγής μπορούν να ενισχύσουν την ακρίβεια ενός ταξινομητή συναισθημάτων, βοηθώντας τον να επιλύσει εσωτερικά τις σημασιολογικές ασάφειες. Έτσι, η προτεινόμενη μέθοδος συνίσταται σε μια νέα διάταξη απόσταξης γνώσης από έναν προεκπαιδευμένο δυαδικό αναγνωριστή μεταφορικής γλώσσας, που χρησιμοποιείται ως βοηθητικό έργο κατά την εκπαίδευση ενός νευρωνικού μοντέλου ανάλυσης συναισθημάτων πολλαπλών κλάσεων κάτω από μια ρύθμιση πολλαπλών εργασιών. Σημειωτέον, δεν απαιτείται να υπάρχει καμία επισήμανση αντικειμενικής αλήθειας σχετικά με τη μεταφορικότητα κατά την εκπαίδευση για την εργασία ανάλυσης συναισθήματος. Η αξιολόγηση σε ένα σχετικό δημόσιο σύνολο δεδομένων δείχνει ότι η προτεινόμενη μέθοδος οδηγεί σε επιδόσεις αιχμής, ξεπερνώντας όλες τις ανταγωνιστικές προσεγγίσεις.

Βιβλιογραφία

- [1] M. El Barachi, M. AlKhatib, S. Mathew, and F. Oroumchian. A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change. *Journal of Cleaner Production*, page 127820, 2021. 1.1.1, 1.2.1, 4.2, 4.3.1.2, 4.3.1.4
- [2] D. M. Romero, B. Meeder, and J. Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. In *Proceedings of the International Conference on World Wide Web*, 2011. 1.1.1, 4.2
- [3] K. Ravi and V. Ravi. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-based systems*, 89:14–46, 2015. 1.1.1, 4.2
- [4] A. Ghosh, G. Li, T. Veale, P. Rosso, E. Shutova, J. Barnden, and A. Reyes. SemEval-2015 Task 11: Sentiment analysis of figurative language in Twitter. In *Proceedings of the International Workshop on Semantic Evaluation (SemEval)*, 2015. 1.1.2, 1.2.2.1, 5.2, 5.3.1, 5.5.2
- [5] S. Muresan, R. Gonzalez-Ibanez, D. Ghosh, and N. Wacholder. Identification of nonliteral language in social media: A case study on sarcasm. *Journal of the Association for Information Science and Technology*, 67(11):2725–2737, 2016. 1.1.2, 5.2
- [6] G. Hinton, O. Vinyals, and J. Dean. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015. 1.1.2, 1.2.2.2, 1.3.2, 5.2, 5.3.2, 5.4, 5.4
- [7] B. Heredia, J. Prusa, and T. Khoshgoftaar. Exploring the effectiveness of Twitter at polling the United States 2016 presidential election. In *Proceedings of the IEEE International Conference on Collaboration and Internet Computing (CIC)*, 2017. 1.2.1, 4.3.1.2, 4.3.1.4
- [8] D. Grimaldi, J. D. Cely, and H. Arboleda. Inferring the votes in a new political landscape: the case of the 2019 Spanish presidential elections. *Journal of Big Data*, 7(1):1–19, 2020. 1.2.1, 4.3.1.2, 4.3.1.4
- [9] U. Yaqub, S. A. Chun, V. Atluri, and J. Vaidya. Sentiment based analysis of tweets during the US presidential elections. In *Proceedings of the Annual International Conference on Digital Government Research*, 2017. 1.2.1, 4.3.1.3, 4.3.1.4, 4.6

- [10] L. Tavoschi, F. Quattrone, E. D'Andrea, P. Ducange, M. Vabanesi, F. Marcelloni, and P. L. Lopalco. Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. *Human Vaccines & Immunotherapeutics*, 16(5):1062–1069, 2020. 1.2.1, 4.3.1.2, 4.3.1.4
- [11] O. Kraaijeveld and J. De Smedt. The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65:101188, 2020. 1.2.1, 4.3.1.3, 4.3.1.4
- [12] W. Shi, H. Wang, and S. He. Sentiment analysis of Chinese microblogging based on sentiment ontology: a case study of '7.23 Wenzhou Train Collision'. *Connection Science*, 25(4):161–178, 2013. 1.2.1, 4.3.1.3, 4.3.1.4, 4.6
- [13] M. Wang, H. Wu, T. Zhang, and S. Zhu. Identifying critical outbreak time window of controversial events based on sentiment analysis. *PLOS One*, 15(10):e0241355, 2020. 1.2.1, 4.3.1.2, 4.3.1.4
- [14] I. Onyenwe, S. Nwagbo, N. Mbeledogu, and E. Onyedinma. The impact of political party/candidate on the election results from a sentiment analysis perspective using# anam-bradecides2017 tweets. *Social Network Analysis and Mining*, 10(1):1–17, 2020. 1.2.1, 4.3.1.3, 4.3.1.4
- [15] F. Barbieri, F. Ronzano, and H. Saggin. UPF-taln: SemEval 2015 Tasks 10 and 11. Sentiment analysis of literal and figurative language in Twitter. In *Proceedings of the International Workshop on Semantic Evaluation (SemEval 2015)*, 2015. 1.2.2.1, 5.3.1, 5.1
- [16] C. Van Hee, E Lefever, and V. Hoste. LT3: sentiment analysis of figurative tweets: piece of cake# NotReally. In *Proceedings of the International Workshop on Semantic Evaluations (SemEval 2015)*, 2015. 1.2.2.1, 5.3.1
- [17] T. Hercig and L. Lenc. The impact of figurative language on sentiment analysis. In *Proceedings of the Recent Advances in Natural Language Processing (RANLP)*, 2017. 1.2.2.1, 5.3.1
- [18] R.-A. Potamias, G. Siolas, and A. Stafylopatis. A robust deep ensemble classifier for figurative language detection. In *Proceedings of the International Conference on Engineering Applications of Neural Networks (EANN)*. Springer, 2019. 1.2.2.1, 5.3.1, 5.1
- [19] J. Pennington, R. Socher, and C. D. Manning. Glove: Global vectors for word representation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014. 1.2.2.1, 5.2, 5.3.1, 5.5.1
- [20] R. A. Potamias, G. Siolas, and A.-G. Stafylopatis. A Transformer-based approach to irony and sarcasm detection. *Neural Computing and Applications*, 32(23):17309–17320, 2020. 1.2.2.1, 5.3.1, 5.5.1, 5.1, 5.5.2

- [21] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. Roberta: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019. 1.2.2.1, 5.3.1, 5.5.1, 5.1
- [22] T. Chen, I. Goodfellow, and J. Shlens. Net2Net: Accelerating learning via knowledge transfer. *arXiv preprint arXiv:1511.05641*, 2015. 1.2.2.2, 5.3.2
- [23] A. Romero, N. Ballas, S. E. Kahou, A. Chassang, C. Gatta, and Y. Bengio. Fitnets: Hints for thin deep nets. *arXiv preprint arXiv:1412.6550*, 2014. 1.2.2.2, 5.3.2
- [24] N. Passalis and A. Tefas. Unsupervised knowledge transfer using similarity embeddings. *IEEE Transactions on Neural Networks and Learning Systems*, 30(3):946–950, 2018. 1.2.2.2, 5.3.2
- [25] S. I. Mirzadeh, M. Farajtabar, A. Li, N. Levine, A. Matsukawa, and H. Ghasemzadeh. Improved knowledge distillation via teacher assistant. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020. 1.2.2.2, 5.3.2
- [26] M. Phuong and C. Lampert. Towards understanding knowledge distillation. In *Proceedings of the International Conference on Machine Learning*, 2019. 1.2.2.2, 5.3.2, 5.4
- [27] L. Saglietti and L. Zdeborová. Solvable model for inheriting the regularization through knowledge distillation. *arXiv preprint arXiv:2012.00194*, 2020. 1.2.2.2, 5.3.2
- [28] Jason Brownlee. What is natural language processing?, 2017. URL <https://machinelearningmastery.com/natural-language-processing/>. Accessed June 23, 2022. 2.1.1
- [29] Wikipedia contributors. Linguistics – Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Linguistics&oldid=1094492783>. [Online; accessed 24-June-2022]. 2.1.4.1
- [30] Alan M Turing. Computing machinery and intelligence. In *Parsing the turing test*, pages 23–65. Springer, 2009. 2.2.1
- [31] Wikipedia contributors. Natural language processing – Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Naturallanguageprocessing&oldid=1093868703>. [Online; accessed 23-June-2022]. 2.2.2.1, 2.2.2.2, 2.2.2.2, 2.5
- [32] W John Hutchins. The georgetown-ibm experiment demonstrated in january 1954. In *Conference of the Association for Machine Translation in the Americas*, pages 102–114. Springer, 2004. 2.2.2.1

- [33] Joseph Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, 1966. 2.2.2.1
- [34] Wikipedia contributors. Parry — Wikipedia, the free encyclopedia, 2021. URL <https://en.wikipedia.org/w/index.php?title=PARRY&oldid=1023645514>. [Online; accessed 23-June-2022]. 2.2.2.1
- [35] Wikipedia contributors. Head-driven phrase structure grammar — Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Head-drivenphrasestructuregrammar&oldid=1080201954>. [Online; accessed 23-June-2022]. 2.2.2.1
- [36] Michael Lesk. Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In *Proceedings of the 5th annual international conference on Systems documentation*, pages 24–26, 1986. 2.2.2.1
- [37] IBM Cloud Education. What is machine learning?, 2020. URL <https://www.ibm.com/cloud/learn/machine-learning>. Accessed June 23, 2022. 2.2.2.2
- [38] Wikipedia contributors. Machine learning — Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Machinelearning&oldid=1094363111>. [Online; accessed 23-June-2022]. 2.2.2.2
- [39] Wikipedia contributors. Supervised learning — Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Supervisedlearning&oldid=1093417731>. [Online; accessed 23-June-2022]. 2.2.2.2
- [40] Jason Brownlee. Supervised and unsupervised machine learning algorithms, 2016. URL <https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/>. Accessed June 23, 2022. 2.2.2.2, 2.2.2.2
- [41] Wikipedia contributors. Unsupervised learning — Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Unsupervisedlearning&oldid=1083933993>. [Online; accessed 23-June-2022]. 2.2.2.2
- [42] Wikipedia contributors. Artificial neural network — Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Artificialneuralnetwork&oldid=1092513725>. [Online; accessed 23-June-2022]. 2.2.2.3

- [43] Charles C Tappert. Who is the father of deep learning? In *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 343–348. IEEE, 2019. 2.2.2.3
- [44] et al. Ivakhnenko, Aleksei Grigorevich. *Cybernetics and forecasting techniques*, volume 8. American Elsevier Publishing Company, 1967. 2.2.2.3
- [45] Sebastian Ruder. A review of the neural history of natural language processing, 2018. URL <https://ruder.io/a-review-of-the-recent-history-of-nlp/index.html#2001neurallanguagemode1s>. Accessed June 23, 2022. 2.2.2.3
- [46] Yoshua Bengio, Réjean Ducharme, and Pascal Vincent. A neural probabilistic language model. *Advances in neural information processing systems*, 13, 2000. 2.2.2.3
- [47] Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167, 2008. 2.2.2.3
- [48] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013. 2.2.2.3, 2.4.2, 2.4.2.1
- [49] Ilya Sutskever. *Training recurrent neural networks*. University of Toronto Toronto, ON, Canada, 2013. 2.2.2.3
- [50] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*, 2014. 2.2.2.3
- [51] Yoon Kim. Convolutional neural networks for sentence classification. *CoRR*, abs/1408.5882, 2014. URL <http://arxiv.org/abs/1408.5882>. 2.2.2.3
- [52] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27, 2014. 2.2.2.3, 2.4.2.2
- [53] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014. 2.2.2.3, 2.4.2.2
- [54] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 2.2.2.3, 2.4.2.2, 2.4.2.2, 2.4.2.2
- [55] Jason Weston, Sumit Chopra, and Antoine Bordes. Memory networks. *arXiv preprint arXiv:1410.3916*, 2014. 2.2.2.3

- [56] Andrew M Dai and Quoc V Le. Semi-supervised sequence learning. *Advances in neural information processing systems*, 28, 2015. 2.2.2.3
- [57] Bryan McCann, James Bradbury, Caiming Xiong, and Richard Socher. Learned in translation: Contextualized word vectors. *Advances in neural information processing systems*, 30, 2017. 2.2.2.3, 2.4.2, 2.4.2.2
- [58] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237, New Orleans, Louisiana, June 2018. Association for Computational Linguistics. doi: 10.18653/v1/N18-1202. URL <https://aclanthology.org/N18-1202>. 2.2.2.3, 2.4.2, 2.4.2.2
- [59] Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. Improving language understanding by generative pre-training. 2018. 2.2.2.3, 2.4.2, 2.4.2.2
- [60] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. 2.2.2.3, 2.4.2, 2.4.2.2
- [61] Nello Cristianini, John Shawe-Taylor, et al. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000. 2.3.2.2
- [62] Irina Rish et al. An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pages 41–46, 2001. 2.3.2.2
- [63] Silviu Guiasu and Abe Shenitzer. The principle of maximum entropy. *The mathematical intelligencer*, 7(1):42–48, 1985. 2.3.2.2
- [64] Leo Breiman, Jerome H Friedman, Richard A Olshen, and Charles J Stone. *Classification and regression trees*. Routledge, 2017. 2.3.2.2
- [65] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001. 2.3.2.2
- [66] Naomi S Altman. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185, 1992. 2.3.2.2
- [67] Fionn Murtagh and Pedro Contreras. Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1):86–97, 2012. 2.3.2.2
- [68] Aristidis Likas, Nikos Vlassis, and Jakob J Verbeek. The global k-means clustering algorithm. *Pattern recognition*, 36(2):451–461, 2003. 2.3.2.2

- [69] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009. 2.3.2.2
- [70] Jason Brownlee. Promise of deep learning for natural language processing, 2017. URL <https://machinelearningmastery.com/promise-deep-learning-natural-language-processing/>. Accessed June 23, 2022. 2.3.3.1, 2.3.3.1
- [71] Sawan Saxena. Understanding embedding layer in keras, 2020. URL <https://medium.com/analytics-vidhya/understanding-embedding-layer-in-keras-bbe3ff1327ce>. Accessed June 23, 2022. 2.3.3.2
- [72] Marius-Constantin Popescu, Valentina E Balas, Liliana Perescu-Popescu, and Nikos Mastorakis. Multilayer perceptron and neural networks. *WSEAS Transactions on Circuits and Systems*, 8(7):579–588, 2009. 2.3.3.2
- [73] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 international conference on engineering and technology (ICET)*, pages 1–6. Ieee, 2017. 2.3.3.2
- [74] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986. 2.3.3.2
- [75] Alex Sherstinsky. Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica D: Nonlinear Phenomena*, 404:132306, 2020. 2.3.3.2
- [76] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. 2.3.3.2, 3.2.4.2
- [77] Christopher Olah. Understanding lstm networks, 2015. URL <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Accessed June 23, 2022. 2.3.3.2, 3.2.4.2
- [78] Jason Brownlee. Cnn long short-term memory networks, 2017. URL <https://machinelearningmastery.com/cnn-long-short-term-memory-networks/>. Accessed June 23, 2022. 2.3.3.2
- [79] Jeet. One hot encoding of text data in natural language processing., 2020. URL <https://medium.com/analytics-vidhya/one-hot-encoding-of-text-data-in-natural-language-processing-2242>. Accessed June 23, 2022. 2.4.1.1

- [80] Wikipedia contributors. Bag-of-words model – Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Bag-of-wordsmodel&oldid=1087243933>. [Online; accessed 23-June-2022]. 2.4.1.2
- [81] Wikipedia contributors. Tf-idf – Wikipedia, the free encyclopedia, 2022. URL <https://en.wikipedia.org/w/index.php?title=Tf%E2%80%93idf&oldid=1092578440>. [Online; accessed 24-June-2022]. 2.4.1.3
- [82] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014. 2.4.2, 2.4.2.1
- [83] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146, 2017. 2.4.2, 2.4.2.1
- [84] Jay Alammar. The illustrated word2vec, 2019. URL <https://jalammar.github.io/illustrated-word2vec/>. Accessed June 23, 2022. 2.4.2.1
- [85] Jay Alammar. The illustrated transformer, 2018. URL <https://jalammar.github.io/illustrated-transformer/>. Accessed June 23, 2022. 2.4.2.2
- [86] Jay Alammar. The illustrated bert, elmo, and co. (how nlp cracked transfer learning), 2018. URL <https://jalammar.github.io/illustrated-bert/>. Accessed June 23, 2022. 2.4.2.2, 3.2.4.5
- [87] Wikipedia contributors. Sentiment analysis – Wikipedia, the free encyclopedia, 2022. URL https://en.wikipedia.org/w/index.php?title=Sentiment_analysis&oldid=1091286443. [Online; accessed 23-June-2022]. 3.1.1
- [88] Harika Bonthu. Rule-based sentiment analysis in python, 2021. URL <https://www.analyticsvidhya.com/blog/2021/06/rule-based-sentiment-analysis-in-python/>. Accessed June 23, 2022. 3.2.2.1
- [89] Steven Loria. Textblob: Simplified text processing, 2020. URL <https://textblob.readthedocs.io/en/dev/>. Accessed June 23, 2022. 3.2.2.1
- [90] Clayton Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 216–225, 2014. 3.2.2.1

- [91] Andrea Esuli and Fabrizio Sebastiani. SENTIWORDNET: A publicly available lexical resource for opinion mining. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy, May 2006. European Language Resources Association (ELRA). URL <http://www.lrec-conf.org/proceedings/lrec2006/pdf/384pdf.pdf>. 3.2.2.1
- [92] Nishtha Arora. Pre-processing of text data in nlp, 2021. URL <https://www.analyticsvidhya.com/blog/2021/06/pre-processing-of-text-data-in-nlp/>. Accessed June 23, 2022. 3.2.3
- [93] Harshith. Text preprocessing in natural language processing, 2019. URL <https://towardsdatascience.com/text-preprocessing-in-natural-language-processing-using-python-61> Accessed June 23, 2022. 3.2.3.1
- [94] Stanford University. Stemming and lemmatization, 2008. URL <https://nlp.stanford.edu/IR-book/html/htmledition/stemming-and-lemmatization-1.html>. Accessed June 23, 2022. 3.2.3.1
- [95] Aravindpai Pai. What is tokenization in nlp? here's all you need to know, 2020. URL <https://www.analyticsvidhya.com/blog/2020/05/what-is-tokenization-nlp/>. Accessed June 23, 2022. 3.2.3.1
- [96] Mohd Zeeshan Ansari, MB Aziz, MO Siddiqui, H Mehra, and KP Singh. Analysis of political sentiment orientations on twitter. *Procedia Computer Science*, 167:1821–1828, 2020. 3.2.4.2
- [97] LILIS Kurniasari and ARIEF Setyanto. Sentiment analysis using recurrent neural network-lstm in bahasa indonesia. *J. Eng. Sci. Technol.*, 15(5):3242–3256, 2020. 3.2.4.2
- [98] Yi Cai, Qingbao Huang, Zeyun Lin, Jingyun Xu, Zhenhong Chen, and Qing Li. Recurrent neural network with pooling operation and attention mechanism for sentiment analysis: A multi-task learning approach. *Knowledge-Based Systems*, 203:105856, 2020. 3.2.4.2
- [99] Dipak Gaikar, Ganesh Sapare, Akanksha Vishwakarma, and Apurva Parkar. Twitter sentimental analysis for predicting election result using lstm neural network. *IRJET*, 2019. 3.2.4.2
- [100] Baidya Nath Saha, Apurbalal Senapati, and Anmol Mahajan. Lstm based deep rnn architecture for election sentiment analysis from bengali newspaper. In *2020 International Conference on Computational Performance Evaluation (ComPE)*, pages 564–569. IEEE, 2020. 3.2.4.2

- [101] Mei Wang, Yangyang Zhu, Shulin Liu, Chunfeng Song, Zheng Wang, Pai Wang, and Xuebin Qin. Sentiment analysis based on attention mechanisms and bi-directional lstm fusion model. In *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBD-Com/IOP/SCI)*, pages 865–868. IEEE, 2019. 3.2.4.2
- [102] Razvan Pascanu, Caglar Gulcehre, Kyunghyun Cho, and Yoshua Bengio. How to construct deep recurrent neural networks. *arXiv preprint arXiv:1312.6026*, 2013. 3.2.4.2
- [103] Jason Brownlee. Stacked long short-term memory networks, 2017. URL <https://machinelearningmastery.com/stacked-long-short-term-memory-networks/>. Accessed June 23, 2022. 3.2.4.2
- [104] Mike Schuster and Kuldip K Paliwal. Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, 45(11):2673–2681, 1997. 3.2.4.2
- [105] Jason Brownlee. How to develop a bidirectional lstm for sequence classification in python with keras, 2017. URL <https://machinelearningmastery.com/develop-bidirectional-lstm-sequence-classification-python-keras/>. Accessed June 23, 2022. 3.2.4.2
- [106] Cezanne Camacho. Cnns for text classification, 2019. URL <https://cezannec.github.io/CNNTextClassification/>. Accessed June 23, 2022. 3.2.4.3
- [107] Ye Zhang and Byron Wallace. A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification. *arXiv preprint arXiv:1510.03820*, 2015. 3.2.4.3
- [108] Chenquan Gan, Lu Wang, Zufan Zhang, and Zhangyi Wang. Sparse attention based separable dilated convolutional neural network for targeted sentiment analysis. *Knowledge-Based Systems*, 188:104827, 2020. 3.2.4.3
- [109] Marco Pota, Massimo Esposito, Marco A Palomino, and Giovanni L Masala. A subword-based deep learning approach for sentiment analysis of political tweets. In *2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, pages 651–656. IEEE, 2018. 3.2.4.3
- [110] Anwar Ur Rehman, Ahmad Kamran Malik, Basit Raza, and Waqar Ali. A hybrid cnn-lstm model for improving accuracy of movie reviews sentiment analysis. *Multimedia Tools and Applications*, 78(18):26597–26613, 2019. 3.2.4.4

- [111] P Kaladevi and K Thyagarajah. Retracted article: Integrated cnn-and lstm-dnn-based sentiment analysis over big social data for opinion mining. *Behaviour & Information Technology*, 40(9):XI–XIX, 2021. 3.2.4.4
- [112] Muhammad Umer, Imran Ashraf, Arif Mehmood, Saru Kumari, Saleem Ullah, and Gyu Sang Choi. Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model. *Computational Intelligence*, 37(1):409–434, 2021. 3.2.4.4
- [113] Shervin Minaee, Elham Azimi, and AmirAli Abdolrashidi. Deep-sentiment: Sentiment analysis using ensemble of cnn and bi-lstm models. *arXiv preprint arXiv:1904.04206*, 2019. 3.2.4.4
- [114] Fazeel Abid, Muhammad Alam, Muhammad Yasir, and Chen Li. Sentiment analysis through recurrent variants latterly on convolutional neural network of twitter. *Future Generation Computer Systems*, 95:292–308, 2019. 3.2.4.4
- [115] Mingjie Ling, Qiaohong Chen, Qi Sun, and Yubo Jia. Hybrid neural network for sina weibo sentiment analysis. *IEEE Transactions on Computational Social Systems*, 7(4):983–990, 2020. 3.2.4.4
- [116] J. C. Baumgartner, J. B. Mackay, J. S. Morris, E. E. Otenyo, L. Powell, M. M. Smith, N. Snow, F. I. Solop, and B. C. Waite. *Communicator-in-chief: How Barack Obama used new media technology to win the White House*. Lexington Books, 2010. 4.2
- [117] D. Lorenzi, J. Vaidya, B. Shafiq, S. Chun, N. Vigesna, Z. Alzamil, N. Adam, S. Wainer, and V. Atluri. Utilizing social media to improve local government responsiveness. In *Proceedings of the Annual International Conference on Digital Government Research*, 2014. 4.2
- [118] J. Ramteke, S. Shah, D. Godhia, and A. Shaikh. Election result prediction using Twitter sentiment analysis. In *Proceedings of the IEEE International Conference on Inventive Computation Technologies (ICICT)*, 2016. 4.2
- [119] A. Agarwal, R. Singh, and D. Toshniwal. Geospatial sentiment analysis using Twitter data for UK-EU referendum. *Journal of Information and Optimization Sciences*, 39(1):303–317, 2018. 4.2
- [120] C. Chatfield. *Time-series forecasting*. CRC Press, 2000. 4.2
- [121] L. F. Bright, K. L. Sussman, and G. B. Wilcox. Facebook, trust and privacy in an election year: Balancing politics and advertising. *Journal of Digital & Social Media Marketing*, 8(4):332–346, 2021. 4.3.1

- [122] J. Lee and Y.-S. Lim. Gendered campaign tweets: the cases of Hillary Clinton and Donald Trump. *Public Relations Review*, 42(5):849–855, 2016. 4.3.1
- [123] L. Buccoliero, E. Bellio, G. Crestini, and A. Arkoudas. Twitter and politics: Evidence from the US presidential elections 2016. *Journal of Marketing Communications*, 26(1): 88–114, 2020. 4.3.1
- [124] D. Grimaldi. Can we analyse political discourse using twitter? evidence from Spanish 2019 presidential election. *Social Network Analysis and Mining*, 9(1):1–9, 2019. 4.3.1.1
- [125] Michael Cornfield. Empowering the party-crasher: Donald J. Trump, the first 2016 GOP presidential debate, and the Twitter marketplace for political campaigns. *Journal of Political Marketing*, 2017. 4.3.1.1, 4.6
- [126] F. Qi, C. Yang, Z. Liu, Q. Dong, M. Sun, and Z. Dong. Openhownet: An open sememe-based lexical knowledge base. *arXiv preprint arXiv:1901.09957*, 2019. 4.3.1.4
- [127] S. Zhang, Z. Wei, Y. Wang, and T. Liao. Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary. *Future Generation Computer Systems*, 81:395–403, 2018. 4.3.1.4
- [128] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas. Sentiment strength detection in short informal text. *Journal of the American society for information science and technology*, 61(12):2544–2558, 2010. 4.3.1.4
- [129] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, 2013. 4.3.1.4
- [130] C. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2014. 4.3.1.4
- [131] H. Maqsood, I. Mehmood, M. Maqsood, M. Yasir, S. Afzal, F. Aadil, M. M. Selim, and K. Muhammad. A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management*, 50:432–451, 2020. 4.3.2
- [132] J. Kordonis, S. Symeonidis, and A. Arampatzis. Stock price forecasting via sentiment analysis on Twitter. In *Proceedings of the Pan-Hellenic Conference on Informatics*, 2016. 4.3.2
- [133] M. Oussalah and A. Zaidi. Forecasting weekly crude oil using Twitter sentiment of US foreign policy and oil companies data. In *Proceedings of the IEEE International Conference on Information Reuse and Integration (IRI)*, 2018. 4.3.2

- [134] M. Arias, A. Arratia, and R. Xuriguera. Forecasting with Twitter data. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):1–24, 2014. 4.3.2
- [135] R. Kiran, P. Kumar, and B. Bhasker. OSLCFit (organic simultaneous LSTM and CNN Fit): a novel deep learning based solution for sentiment polarity classification of reviews. *Expert Systems with Applications*, 157:113488, 2020. 4.4.2, 4.4.2.3
- [136] M. Zampieri, S. Malmasi, P. Nakov, S. Rosenthal, N. Farra, and R. Kumar. Semeval-2019 task 6: Identifying and categorizing offensive language in social media (offenseval). *arXiv preprint arXiv:1903.08983*, 2019. 4.4.2.1
- [137] J. Ling and R. Klinger. An empirical, quantitative analysis of the differences between sarcasm and irony. In *European Semantic Web Conference*. Springer, 2016. 4.4.2.1
- [138] J. Pennington, R. Socher, and C. D. Manning. Glove: Global vectors for word representation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014. 4.4.2.3
- [139] H. Hewamalage, C. Bergmeir, and K. Bandara. Recurrent Neural Networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1):388–427, 2021. 4.5.2.1
- [140] F. Orabona and T. Tommasi. Training deep networks without learning rates through coin betting. *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, 30:2160–2170, 2017. 4.5.2.1
- [141] F. Hutter, H. H. Hoos, and K. Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *Proceedings of the International Conference on Learning and Intelligent Optimization*. Springer, 2011. 4.5.2.1
- [142] O. Claveria, E. Monte, and S. Torra. Data preprocessing for neural network-based forecasting: does it really matter? *Technological and Economic Development of Economy*, 23(5):709–725, 2017. 4.5.2.1
- [143] R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006. 4.5.2.2
- [144] D. K. Tayal, S. Yadav, K. Gupta, B. Rajput, and K. Kumari. Polarity detection of sarcastic political tweets. In *Proceedings of the International Conference on Computing for Sustainable Global Development (INDIACoM)*. IEEE, 2014. 4.5.3
- [145] S. Lai, L. Xu, K. Liu, and J. Zhao. Recurrent Convolutional Neural Networks for text classification. In *Proceedings of AAAI conference on artificial intelligence*, 2015. 5.5.1

- [146] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*, 2018. 5.1
- [147] D. Cer, Y. Yang, S. Kong, N. Hua, N. Limtiaco, R. John, N. Constant, M. Guajardo-Céspedes, S. Yuan, and C. Tar. Universal sentence encoder. *arXiv preprint arXiv:1803.11175*, 2018. 5.1
- [148] S. I. Wang and C. D. Manning. Baselines and bigrams: Simple, good sentiment and topic classification. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2012. 5.1
- [149] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov. FastText.zip: Compressing text classification models. *arXiv preprint arXiv:1612.03651*, 2016. 5.1
- [150] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le. Xlnet: Generalized autoregressive pretraining for language understanding. *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, 2019. 5.1
- [151] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pretraining of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. 5.1
- [152] C. Ozdemir and S. Bergler. CLaC-SentiPipe: Semeval2015 Subtasks 10 b, e, and Task 11. In *Proceedings of the International Workshop on Semantic Evaluation (SemEval 2015)*, 2015. 5.1
- [153] R. Kiran, P. Kumar, and B. Bhasker. OSLCFit (Organic Simultaneous LSTM and CNN Fit): a novel deep learning-based solution for sentiment polarity classification of reviews. *Expert Systems with Applications*, 157:113488, 2020. 5.5.1
- [154] J. Ling and R. Klinger. An empirical, quantitative analysis of the differences between sarcasm and irony. In *Proceedings of the European Semantic Web Conference*. Springer, 2016. 5.5.1