



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών
Τομέας Σημάτων, Ελέγχου και Ρομποτικής

Αναγνώριση Προθέσεως και Προσωπικότητας Ομιλητών σε Διαλόγους Χρησιμοποιώντας Βαθιά Νευρωνικά Δίκτυα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΗΝΕΛΟΠΗ ΠΑΠΑΛΑΜΠΙΔΗ

Επιβλέπων : Αλέξανδρος Ποταμιάνος
Αναπληρωτής Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2018



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών
Τομέας Σημάτων, Ελέγχου και Ρομποτικής

Αναγνώριση Προθέσεως και Προσωπικότητας Ομιλητών σε Διαλόγους Χρησιμοποιώντας Βαθιά Νευρωνικά Δίκτυα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΗΝΕΛΟΠΗ ΠΑΠΑΛΑΜΠΙΔΗ

Επιβλέπων : Αλέξανδρος Ποταμιάνος
Αναπληρωτής Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 25η Ιουνίου 2018.

.....
Αλέξανδρος Ποταμιάνος
Αναπληρωτής Καθηγητής Ε.Μ.Π.

.....
Γιώργος Στάμου
Αναπληρωτής Καθηγητής Ε.Μ.Π.

.....
Joakim Gustafsson
Καθηγητής ΚΤΗ

Αθήνα, Ιούνιος 2018

.....
Πηγελόπη Παπαλαμπίδη

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Πηγελόπη Παπαλαμπίδη, 2018.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στα πλαίσια αυτής της διατριβής, εξετάζουμε δύο σημαντικά ζητήματα αποσκοπώντας στη βελτίωση της απόδοσης των συστημάτων διαλόγου στο γενικότερο πλαίσιο της αλληλεπίδρασης ανθρώπου-μηχανής. Το πρώτο ζήτημα, η ταξινόμηση των διαλογικών πράξεων, είναι συνήθως το πιο θεμελιώδες στην επεξεργασία του διαλόγου. Επιτρέπει στα συστήματα διαλόγου να επιλέγουν την καταλληλότερη απάντηση από ένα σύνολο πιθανών απαντήσεων βασισμένα στις επικοινωνιακές προθέσεις του χρήστη εκείνη τη στιγμή. Το δεύτερο ζήτημα, η αναγνώριση της προσωπικότητας, στοχεύει στην περαιτέρω βελτίωση των συστημάτων διαλόγου, προκειμένου να προσαρμόζονται με επιτυχία στη συμπεριφορά ενός μεμονωμένου χρήστη. Αυτή η προσαρμογή συμβάλλει στην ανάπτυξη εξατομικευμένων συστημάτων διαλόγου.

Οι διαλογικές πράξεις θεωρούνται ως οι ελάχιστες γλωσσικές μονάδες επικοινωνιακών προθέσεων στο πλαίσιο του διαλόγου. Ωστόσο, τα χρησιμοποιούμενα σχήματα επισήμανσης για την αναγνώριση διαλογικών πράξεων εξαρτώνται από το πρόβλημα και τον τομέα κάθε συνόλου δεδομένων. Σε αυτήν την εργασία, προτείνουμε μια μέθοδο για την επιτυχή αναγνώριση των προθέσεων ενός ομιλητή ανεξάρτητα από το πρόβλημα του διαλόγου και τη λεπτομερειακότητα του συνόλου των ετικετών των διαλογικών πράξεων που χρησιμοποιήθηκαν. Για το σκοπό αυτό, εφαρμόζουμε ένα μοντέλο βαθιάς μάθησης για την αναγνώριση των διαλογικών πράξεων με βάση τη σημασιολογική εκπροσώπηση της τρέχουσας φράσης καθώς και την ιστορία του διαλόγου μέχρι στιγμής. Σε αυτήν την αρχιτεκτονική, προτείνουμε την επέκταση των γενικών διανυσμάτων λέξεων (word embeddings), που χρησιμοποιούνται για την αρχικοποίηση του στρώματος ενσωμάτωσης (embedding layer) του αντίστοιχου νευρικού δικτύου, με πληροφορίες σχετικά με τις διαλογικές πράξεις. Συγκεκριμένα, πρώτα εξαγάγουμε αυτόματα ένα σύνολο λέξεων-κλειδιών που θεωρούνται αντιπροσωπευτικές για κάθε ετικέτα διαλογικής πράξης, σχηματίζοντας ένα σημασιολογικό χώρο για κάθε ετικέτα. Στη συνέχεια, υπολογίζουμε τη σημασιολογική ομοιότητα μεταξύ κάθε λέξης και κάθε ετικέτας διαλογικής πράξης, υπολογίζοντας την ομοιότητα μεταξύ κάθε λέξης και του αντίστοιχου συνόλου λέξεων-κλειδιών. Τέλος, επεκτείνουμε τα γενικά διανύσματα λέξεων με τα διανύσματα λέξεων που δημιουργήσαμε και τα οποία περιέχουν πληροφορία σχετική με τις διαλογικές πράξεις. Τα τελικά ενισχυμένα διανύσματα λέξεων τροφοδοτούνται στο νευρωνικό μοντέλο μας. Αξιολογούμε την προσέγγισή μας σε ένα σύνολο δεδομένων που χρησιμοποιείται συνήθως για την ταξινόμηση των διαλογικών πράξεων και επιτυγχάνουμε συγκρίσιμα αποτελέσματα με το state-of-the-art μοντέλο.

Στη συνέχεια, αντιμετωπίζουμε το πρόβλημα της αυτόματης αναγνώρισης της προσωπικότητας. Τα χαρακτηριστικά της προσωπικότητας περιγράφονται με το Big Five μοντέλο, που προέρχεται από ψυχολογικές μελέτες. Αυτό το μοντέλο θεωρείται επαρκές για να περιγράψει την ανθρώπινη προσωπικότητα σε διάφορες γλώσσες και πολιτισμούς. Σε αυτή την εργασία υιοθετούμε την υπόθεση ότι τα χαρακτηριστικά της προσωπικότητας εξαρτώνται και από το πλαίσιο μιας δεδομένης κατάστασης και επομένως σχετίζονται με τη συμπεριφορική και συναισθηματική κατάσταση καθώς και με τις προθέσεις του ατόμου. Στην πραγματικότητα, συσχετίζουμε για πρώτη φορά το πρόβλημα της αναγνώρισης προθέσεων των ομιλητών με το πρόβλημα της αναγνώρισης της προσωπικότητας. Στόχος μας είναι να ενσωματώσουμε τη γνώση συναισθημάτων και προθέσεων στο πρόβλημα της αυτόματης αναγνώρισης της προσωπικότητας. Προτείνουμε μια καινοφανή προσαρμογή δύο γνωστών μεθόδων

μεταφοράς νευρωνικής μάθησης για την ενσωμάτωση πληροφοριών σχετικών με τα συναισθημάτα και τις προθέσεις των ατόμων σε επίπεδο προτάσεων στην αναγνώριση της προσωπικότητας τους σε επίπεδο εγγράφου. Τα μοντέλα μας βασίζονται σε δίκτυα ιεραρχικής προσοχής. Κατ' αρχάς, εκπαιδεύουμε ένα μοντέλο σε ένα πρόβλημα πηγής (source task) (συγκεκριμένα, συνείσθημα, πρόθεση ή και τα δύο μέσω της μάθησης πολλαπλών εργασιών). Στη συνέχεια, χρησιμοποιούμε τον κωδικοποιητή του προεκπαιδευμένου μοντέλου για την τελειοποίηση του προβλήματος στόχου (target task) ή ως εξολκέα χαρακτηριστικών επιπέδου προτάσεων. Η προτεινόμενη προσέγγιση επιτυγχάνει state-of-the-art αποτελέσματα σε δύο σύνολα δεδομένων αναγνώρισης προσωπικότητας. Επίσης, αξιολογούμε την ενσωμάτωση των ψυχολογικών χαρακτηριστικών, εξαγόμενων από λεξικά, στο μοντέλο μας, όπως έχει ήδη προταθεί στη βιβλιογραφία. Τέλος, διεξάγουμε μια ανάλυση σχετικά με τη συμβολή των διαφόρων πηγών πληροφοριών στο πρόβλημα και επικυρώνουμε την αρχική μας υπόθεση σχετικά με την συμβολή του προβλήματος της αναγνώρισης πρόθεσης στο πρόβλημα στόχο.

Λέξεις κλειδιά

διαλογική πράξη, βαθιά εκμάθηση, αναδρομικά νευρωνικά δίκτυα, σημασιολογική ομοιότητα, αναγνώριση προσωπικότητας, δίκτυο ιεραρχικής προσοχής, μεταφορά μάθησης, συστήματα διαλόγου

Abstract

In this work, we address two important issues related to the improvement of the performance of dialogue systems in terms of human-machine interaction. The first issue, dialogue act classification, is typically the most fundamental in dialogue processing. It allows the dialogue systems to select the most appropriate response from a set of possible generated replies based on the communicative intentions of the user at that moment. The second issue, personality recognition, aims at the further improvement of the dialogue systems in order to successfully adapt to an individual user’s behavior. This adaptation assists in developing personalized dialogue systems.

Dialogue acts are considered as the minimal linguistic units of communicative intentions in terms of dialogues. However, the utilized annotation schemes for dialogue act recognition are dependent on the task and domain of each dataset. In this work, we propose a method for successfully recognizing the intentions of an interlocutor independently from the task of the dialogue and the granularity of the dialogue act tag-set utilized. For this purpose, we implement a deep learning model for recognizing the dialogue acts based on the semantic representation of the current utterance as well as the history of the dialogue so far. In this architecture, we propose the expansion of generic word embeddings, that are used for initializing the embedding layer of the respective neural network, with dialogue act-specific information. Specifically, first we automatically extract a set of keywords that is considered representative for each dialogue act tag, forming a semantical subspace for each tag. Next, we compute the semantic similarity between each word and each dialogue act tag, by computing the similarity between each word and the respective set of keywords. Finally, we concatenate the generic word embeddings with the custom word vectors and fed them to our neural model. We evaluate our approach in a dataset commonly used for dialogue act classification and achieve results comparable with the state-of-the-art.

Next, we address the problem of automatic personality recognition. Personality traits are described with the Big Five model, derived from psychological studies. This model is considered sufficient for outlining the human personality across different languages and cultures. In this work, we adopt the hypothesis that the personality traits are nevertheless dependent on the context of a given situation and hence, related to the behavioral and emotional state as well as the intention of the individual. In fact, we first introduce the relevance of intent recognition to the personality recognition problem. We aim at incorporating emotion and intent knowledge to the automatic personality recognition problem. We propose a novel adaptation of two well-known neural transfer learning methods for incorporating sentence-level emotion and intent information to document-level personality recognition. Our models are based on hierarchical attention networks. First, we train a model on a sentence-level source task (i.e. emotion, intent or both via multi-task learning). Next, we utilize the encoder of the pretrained model for fine-tuning on the target task or as a sentence-level feature extractor. The suggested approach achieves state-of-the-art results in two personality datasets. We also evaluate the incorporation of lexicon-based psycholinguistic features to our model, as already suggested in the literature. Finally, we conduct an analysis on the contribution of different information sources to the problem and validate our initial assumption.

Key words

dialogue act, deep learning, recurrent neural networks, semantic similarity, personality recognition, hierarchical attention network, transfer learning, dialogue systems

Ευχαριστίες

Πρώτα απ' όλα, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, Αλέξανδρο Ποταμιάνο, για την καθοδήγησή του κατά τα τελευταία χρόνια που διεκπεραιώνω τη διπλωματική εργασία μου. Οι πολύτιμες συμβουλές του με βοήθησαν να βελτιώσω την έρευνά μου και να δημοσιεύσω τη δουλειά μου.

Στη συνέχεια, θα ήθελα να ευχαριστήσω τους καθηγητές Γιώργο Στάμου και Joakim Gustafsson, που συμμετέχουν στην εξεταστική επιτροπή της διπλωματικής μου εργασίας και τους συναδέλφους μου στο εργαστήριο SLP-NTUA με τους οποίους μοιράστηκα ερευνητικές ιδέες και δημιουργικές συζητήσεις (ευχαριστώ ιδιαίτερος τον Χρήστο Μπαζιώτη για την υπερ-προσπάθεια που καταβάλαμε για την υποβολή ενός paper στο EMNLP 2018).

Τέλος, ευχαριστώ βαθιά την οικογένειά μου για την υποστήριξή και τη βοήθεια της όλα αυτά τα χρόνια ώστε να μπορέσω να πετύχω τους στόχους μου. Ευχαριστώ επίσης τους φίλους μου, με τους οποίους μοιραστήκαμε τόσο χαρούμενες όσο και αγχωτικές στιγμές όλα αυτά τα χρόνια.

Πηνελόπη Παπαλαμπίδη,

Αθήνα, 25η Ιουνίου 2018

Στην οικογένειά μου.

Περιεχόμενα

Περίληψη	5
Abstract	7
Ευχαριστίες	9
Περιεχόμενα	13
Κατάλογος πινάκων	15
Κατάλογος σχημάτων	17
1. Εισαγωγή	21
1.1 Αναγνώριση Προθέσεων	22
1.2 Αναγνώριση Προσωπικότητας	22
1.3 Διάρθρωση Διπλωματικής Εργασίας	24
2. Θεωρητικό Υπόβαθρο	25
2.1 Αναγνώριση Προθέσεων	25
2.1.1 Ψυχολογική Ανάλυση της Σκοπιμότητας (intentionality)	25
2.1.2 Ορισμός των Επικοινωνιακών Πράξεων (Communicative Acts - CAs)	26
2.1.3 Ορισμός των Πράξεων Ομιλίας (Speech Acts) & των Διαλογικών Πράξεων (Dialogue Acts - DAs)	28
2.1.4 Εφαρμογές & Ταξινόμηση των DAs	29
2.1.5 Σύγχρονες Εφαρμογές Συστημάτων Διαλόγου	30
2.2 Αναγνώριση Προσωπικότητας	31
2.2.1 Θεωρίες της Προσωπικότητας	31
2.2.2 Καταλληλότητα του Big Five Μοντέλου	36
2.3 Εκμάθηση Μηχανών για Επεξεργασία Φυσικής Γλώσσας	36
2.3.1 Επεξεργασία Φυσικής Γλώσσας	36
2.3.2 Ορισμός της Μάθησης Μηχανών	38
2.3.3 Ορισμός Διαφορετικών Μεθόδων ML	39
2.3.4 Χαρακτηριστικά	40
2.3.5 Παραδοσιακές Προσεγγίσεις ML	43
2.3.6 Διανύσματα Λέξεων (Word Embeddings)	46
2.4 Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks)	48
2.4.1 Εισαγωγή	48
2.4.2 Τεχνητά Νευρωνικά Δίκτυα	48
2.4.3 Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks - RNNs)	55
2.4.4 Μεταφορά Μάθησης (Transfer Learning - TL)	58

3. Αναπαράσταση και Ταξινόμηση Διαλογικών Πράξεων Χρησιμοποιώντας Βαθιά Νευρωνικά Δίκτυα	61
3.1 Εισαγωγή	61
3.2 Επισκόπηση Βιβλιογραφίας	62
3.3 Προτεινόμενο Μοντέλο	63
3.3.1 Μοντέλο Πρότασης	63
3.3.2 Μοντέλο Διάρθρωσης του Διαλόγου (Discourse Model)	65
3.4 Πειραματικό Σύνολο Δεδομένων	66
3.5 Ρύθμιση Παραμέτρων	67
3.5.1 Επιλογή Λέξεων-κλειδιών	67
3.5.2 Παράμετροι του LSTM	68
3.5.3 Αποτελέσματα και Αξιολόγηση	69
3.5.4 Συμπεράσματα	71
4. Μεταφορά Μάθησης από τα Προβλήματα Αναγνώρισης Πρόθεσης και Συναισθήματος στο Πρόβλημα Αναγνώρισης Προσωπικότητας σε Επίπεδο Εγγράφου	73
4.1 Εισαγωγή	73
4.2 Επισκόπηση της Βιβλιογραφίας	74
4.3 Baseline Μοντέλο	76
4.4 Μεταφορά Μάθησης	76
4.4.1 Προεκπαίδευση του SE	77
4.4.2 Αρχικοποίηση (INIT-TL) του SE	77
4.4.3 Hypercolumns-TL (HTL) του SE	77
4.5 Πειράματα & Αποτελέσματα	78
4.5.1 Πειραματικά Σύνολα Δεδομένων	78
4.5.2 Πειραματική Ρύθμιση των Μοντέλων	80
4.5.3 Αποτελέσματα	82
4.6 Συμπεράσματα	84
5. Συμπεράσματα	85
Βιβλιογραφία	89
Παράρτημα	101
A. Συντομογραφίες	101

Κατάλογος πινάκων

3.1	Σύνολο δεδομένων Switchboard-DAMSL.	67
3.2	Παραδείγματα των πιο συχνών DAs.	67
3.3	Σχετική συχνότητα (%) των DAs.	68
3.4	Παραδείγματα αυτόματα επιλεγμένων λέξεων-κλειδιών (εμφανίζονται για τις πιο συχνές DAs).	69
3.5	Απόδοση υπερπαραμέτρων LSTM με βάση πειραματισμούς.	70
3.6	Απόδοση του baseline και του προτεινόμενου μοντέλου.	70
3.7	Απόδοση του προτεινόμενου μοντέλου και άλλων μεθόδων από την βιβλιογραφία.	71
4.1	Περιγραφή μοντέλου Big Five	75
4.2	Στατιστικά ετικετών πρόθεσης στο DailyDialog	78
4.3	Στατιστικά ετικετών συναισθημάτων στο DailyDialog	79
4.4	Στατιστικά στοιχεία για το σύνολο δεδομένων YouTube	79
4.5	Στατιστικά στοιχεία για το σύνολο δεδομένων stream-of-consciousness essay	79
4.6	Αποτελέσματα στο σύνολο δεδομένων stream-of-consciousness essay. Οι βαθμολογίες υπολογίζονται με τη χρήση της μέτρησης ακρίβεια ταξινόμησης (<i>accuracy</i>) χρησιμοποιώντας 10-fold cross-validation.	80
4.7	Αποτελέσματα για το σύνολο δεδομένων YouTube. Οι βαθμολογίες υπολογίζονται χρησιμοποιώντας τη μέτρηση <i>RMSE</i> , με 10-fold cross-validation. Τα προσαρμοσμένα οπτικοακουστικά χαρακτηριστικά περιλαμβάνονται σε όλα τα μοντέλα (δείτε το Σχήμα 4.1a).	81

Κατάλογος σχημάτων

2.1	Επισκόπηση του μοντέλου διεργασίας συναλλαγών	27
2.2	Η κατανομή των χαρακτηριστικών προσωπικότητας του Eysenck	34
2.3	Τα χαρακτηριστικά της προσωπικότητας του Cattell	35
2.4	Γραμμική ταξινόμηση χρησιμοποιώντας SVM	44
2.5	Διαδικασία εκπαίδευσης του τυχαίου δασικού ταξινομητή	46
2.6	Μέθοδοι CBOW και skip-gram για την κατάρτιση word2vec embeddings	47
2.7	Επισκόπηση ενός βιολογικού νευρώνα	49
2.8	Επισκόπηση των λειτουργιών ενός τεχνητού νευρώνα	49
2.9	Επισκόπηση της αρχιτεκτονικής ενός ANN	50
2.10	Sigmoid λειτουργία	51
2.11	Υπερβολική λειτουργία εφαπτομένης	52
2.12	Λειτουργία ReLU	52
2.13	Λειτουργία ReLU διαρροής	53
2.14	Βρόχοι σε RNNs	56
2.15	Ξετυλιγμένη έκδοση ενός RNN	56
2.16	Δομή ενός κυττάρου LSTM	57
3.1	Επισκόπηση του προτεινόμενου μοντέλου.	63
3.2	Επισκόπηση του baseline μοντέλου για την δημιουργία του διανύσματος αναπαράστασης της φράσης s	64
3.3	Επισκόπηση του μοντέλου διάρθρωσης του διαλόγου που προβλέπει την ετικέτα DA_{z_i} της φράσης s_i	66
4.1	Επισκόπηση των προτεινόμενων μοντέλων	75
4.2	Συμβολή των πηγών πληροφόρησης για κάθε μέθοδο TL στο σύνολο δεδομένων stream-of-consciousness essay. Η μέτρηση αξιολόγησης είναι <i>ακρίβεια ταξινόμησης (accuracy)</i> , πράγμα που σημαίνει ότι οι μεγαλύτερες επιφάνειες υποδηλώνουν καλύτερη απόδοση. Οι προτεινόμενες προσεγγίσεις είναι ανταγωνιστικές με το μοντέλο HAN ενισχυμένο με τα χαρακτηριστικά, που έχουν εξαχθεί βάσει του λεξικού LIWC.	80
4.3	Συμβολή των πηγών πληροφοριών για κάθε μέθοδο TL στο σύνολο δεδομένων του YouTube. Η μέτρηση αξιολόγησης είναι <i>RMSE</i> , πράγμα που σημαίνει ότι οι μικρότερες επιφάνειες υποδηλώνουν καλύτερη απόδοση. Και οι δύο προσεγγίσεις υπερβαίνουν το μοντέλο HAN ενισχυμένο με τα χαρακτηριστικά, που έχουν εξαχθεί βάσει του λεξικού LIWC για τις περισσότερες ρυθμίσεις.	81

Κεφάλαιο 1

Εισαγωγή

Η ανθρώπινη επικοινωνία συνήθως αποσκοπεί στη συμφωνία μεταξύ συνομιλητών σχετικά με συγκεκριμένους ορισμούς μιας δεδομένης κατάστασης. Ωστόσο, συχνά η συμφωνία αυτή δεν είναι άμεση και ως εκ τούτου χρειάζεται διάλογος για να μπορέσουν οι συνομιλητές να λάβουν μια απόφαση. Κατά τη διάρκεια του διαλόγου, οι συνομιλητές δημιουργούν ένα κοινό εννοιολογικό έδαφος, στο οποίο συμφωνούν για κοινές πεποιθήσεις, αμοιβαία γνώση, κοινούς στόχους και κοινές προθέσεις και είναι σε θέση να αναγνωρίζουν τα επικοινωνιακά κίνητρα του άλλου. Προκειμένου να αναγνωριστούν τα επικοινωνιακά κίνητρα ενός ομιλητή, οι Διαλογικές Πράξεις (Dialogue Acts - DAs), οι οποίες θεωρούνται ως η ελάχιστη μονάδα γλωσσικής επικοινωνίας και συνδέονται άμεσα με τις επικοινωνιακές προθέσεις του ομιλητή, πρέπει να παρακολουθούνται καθώς ο διάλογος εξελίσσεται. Επιπλέον, η συμπεριφορά κάθε ατόμου κατά τη διάρκεια της αλληλεπίδρασης εξαρτάται από τη συμπεριφορά και τα χαρακτηριστικά του συνομιλητή του. Ως εκ τούτου, η ικανότητα αναγνώρισης των προσωπικών χαρακτηριστικών ενός ατόμου (π.χ. Εξωστρέφεια, Ευκολία, Ανοιχτότητας σύμφωνα με το Big Five μοντέλο για την ανάλυση της ανθρώπινης προσωπικότητας) μπορεί να βοηθήσει στη δημιουργία ενός κοινού εννοιολογικού εδάφους στην ανθρώπινη επικοινωνία. Οι κύριοι στόχοι της διατριβής είναι η αναγνώριση των προθέσεων και της προσωπικότητας ομιλητών, καθώς είναι δύο από τα θεμελιώδη στοιχεία σε κάθε είδους επικοινωνία, είτε στην αλληλεπίδραση μεταξύ ανθρώπου-ανθρώπου ή είτε μεταξύ ανθρώπου-μηχανής.

Με βάση τους κανόνες που διέπουν την ανθρώπινη επικοινωνία, ένα σημαντικό βήμα επεξεργασίας για τη βελτίωση της αλληλεπίδρασης ανθρώπου-μηχανής στο πλαίσιο Συστημάτων Ομιλούμενου Διαλόγου (Spoken Dialogue Systems - SDS) είναι η ταξινόμηση DA. Η αναγνώριση των DA μπορεί να βοηθήσει στην κατανόηση της εισόδου του χρήστη. Συνήθως, αυτό εφαρμόζεται με την εκχώρηση ετικετών στις δηλώσεις χρηστών που περιγράφουν (λεξικά) τις αντίστοιχες πράξεις. Επιπλέον, η ικανότητα ενός SDS να εντοπίσει και να προσαρμοστεί στην προσωπικότητα του ατόμου μπορεί να βελτιώσει τον βαθμό ενδιαφέροντος των ανθρώπων κατά την αλληλεπίδραση με μια μηχανή, καθώς και τη δυνατότητα επίτευξης συμφωνίας στο τέλος του διαλόγου. Επιπλέον, τα SDS μπορούν να υιοθετήσουν συγκεκριμένα χαρακτηριστικά προσωπικότητας προκειμένου να παράγουν απαντήσεις που θα εξαρτώνται από αυτά τα χαρακτηριστικά.

Με βάση τον παραπάνω ορισμό και την ανάλυση της ανθρώπινης επικοινωνίας, ο κύριος στόχος αυτής της εργασίας είναι να βελτιώσει την αλληλεπίδραση ανθρώπου-μηχανής (π.χ. SDS) εξοπλίζοντας την “μηχανή” με θεμελιώδεις ικανότητες στην επικοινωνία: αναγνώριση των ανθρωπίνων επικοινωνιακών κινήτρων και των χαρακτηριστικών της προσωπικότητας. Συγκεκριμένα, εξετάζουμε πρώτα το πρόβλημα της αυτόματης αναγνώρισης προθέσεων, καθώς θεωρείται το βασικό στοιχείο κάθε ανθρώπινης επικοινωνίας. Στη συνέχεια, προχωρούμε στο πρόβλημα της αυτόματης αναγνώρισης της προσωπικότητας. Σε αυτό το πρόβλημα υιοθετούμε την υπόθεση που υποστηρίζεται στη βιβλιογραφία ότι η αναγνώριση της προσωπικότητας εξαρτάται από το πλαίσιο της εξεταζόμενης κατάστασης, για παράδειγμα τη συμπεριφορά και την συναισθηματική κατάσταση καθώς και τις προθέσεις του ατόμου. Στην πραγματικότητα, εξ όσων γνωρίζουμε, εισάγουμε για πρώτη φορά την αναγνώριση προθέσεων ως ένα πρόβλημα το οποίο σχετίζεται με την αναγνώριση της προσωπικότητας.

1.1 Αναγνώριση Προθέσεων

Το πρώτο ερευνητικό ερώτημα της διατριβής είναι η αναγνώριση των προθέσεων των συνομιλητών καθώς προχωρά ο διάλογος, δεδομένου ότι η αναγνώριση DA είναι παραδοσιακά το πρώτο βήμα στην επεξεργασία του διαλόγου.

Μια πρόκληση στον τομέα αυτό είναι ο ορισμός μιας ταξινόμησης των DAs που μπορούν να γενικευτούν και να χρησιμοποιηθούν σε διαφορετικούς τομείς και διάλογους διαφορετικής φύσης. Παρόλο που έχουν γίνει προσπάθειες για την καθιέρωση ενός γενικού σχεδίου επισήμανσης στην ταξινόμηση DA, εξακολουθούν να υπάρχουν περιορισμένα δεδομένα επισήμασμένα με ένα τέτοιο σχήμα. Ως εκ τούτου, μια άλλη πρόκληση είναι η εφαρμογή ενός μοντέλου ταξινόμησης DA που μπορεί να προβλέψει με επιτυχία τις DAs στο πλαίσιο ενός διαλόγου, ανεξάρτητα από το πρόβλημα και τον τομέα του διαλόγου, καθώς και την ταξινόμηση DA που χρησιμοποιείται για την επισήμανση του συγκεκριμένου συνόλου δεδομένων.

Οι προηγούμενες προσεγγίσεις βασίζονται στη μηχανική των χαρακτηριστικών (feature engineering) προκειμένου να εντοπιστούν σωστά οι DAs στο πλαίσιο του διαλόγου. Συνήθως εξάγουν λεξιλογικά, συντακτικά και σημασιολογικά χαρακτηριστικά που σχετίζονται με τα συναισθήματα των ομιλητών, καθώς και χαρακτηριστικά που σχετίζονται με το ιστορικό του διαλόγου (π.χ., προηγούμενες φράσεις). Οι πιο πρόσφατες προσεγγίσεις έχουν εκμεταλλευτεί τα βαθιά νευρωνικά δίκτυα (Deep Neural Networks - DNN) προκειμένου να καταγράψουν τις πληροφορίες που αποκρύπτονται τόσο σε επίπεδο φράσης όσο και σε επίπεδο διαλόγου. Για το σκοπό αυτό, στη βιβλιογραφία έχουν προταθεί αρκετές αρχιτεκτονικές, συμπεριλαμβανομένης της μακροχρόνιας βραχυπρόθεσμης μνήμης (Long Short-Term Memory - LSTM) και των συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks - CNNs). Ωστόσο, αυτά τα μοντέλα χάνουν σημασιολογικές πληροφορίες που σχετίζονται με το συγκεκριμένο πρόβλημα της αναγνώρισης DA, παρουσιάζοντας ως αποτέλεσμα μια διακύμανση της απόδοσής τους, ανάλογα με τα διαφορετικά χρησιμοποιούμενα σύνολα δεδομένων και το αντίστοιχο σύνολο των DA που χρησιμοποιούνται για την επισήμανση.

Σε αυτή την εργασία, προτείνουμε την επέκταση των γενικών ενσωματώσεων λέξεων (word embeddings) με διανύσματα λέξεων που υποδηλώνουν τη σημασιολογική ομοιότητα κάθε λέξης με κάθε μία ετικέτα DA. Αυτές οι ενσωματωμένες λέξεις υπολογίζονται βάσει σημασιολογικών υποχώρων που αντιστοιχούν σε κάθε ετικέτα DA, που έχουν κατασκευαστεί με βάση την αυτόματη εξαγωγή λέξεων-κλειδιών που θεωρούνται αντιπροσωπευτικές αυτών των ετικετών. Στη συνέχεια, η συνένωση των γενικών ενσωματώσεων λέξεων με τα διανύσματα τα οποία υπολογίστηκαν βάσει της ομοιότητας των λέξεων με τις ετικέτες DA τροφοδοτείται σε ένα DNN για την πρόβλεψη του DA για κάθε έκφραση κάθε διαλόγου. Οι ενισχυμένες ενσωματωμένες λέξεις που προτείνουμε υπολογίζονται αυτόματα και είναι σε θέση να προσαρμοστούν σε σύνολα δεδομένων διαφορετικών τομέων και σχημάτων επισήμανσης DA διαφορετικής λεπτομερειακότητας. Στα πειράματά μας, χρησιμοποιούμε μια πολύ γνωστή βάση δεδομένων, που χρησιμοποιείται συνήθως για το σκοπό αυτό στη βιβλιογραφία, για την αξιολόγηση της προσέγγισής μας και επιτυγχάνουμε αποτελέσματα συγκρίσιμα με τα αντίστοιχα state-of-the-art.

1.2 Αναγνώριση Προσωπικότητας

Το δεύτερο ερευνητικό ερώτημα που εξετάζεται στη διατριβή και που στοχεύει στη βελτίωση της απόδοσης των SDS είναι η αναγνώριση των χαρακτηριστικών της προσωπικότητας ενός ατόμου. Η αυτόματη ανίχνευση των χαρακτηριστικών προσωπικότητας ενός ατόμου έχει πολλές εφαρμογές τόσο στην πραγματική ζωή όσο και σε άλλα προβλήματα τεχνητής νοημοσύνης (Artificial Intelligence - AI).

Στα συστήματα συστάσεων, η ομαδοποίηση των χαρακτηριστικών της προσωπικότητας των χρηστών σε μεγάλες ομάδες μπορεί να βελτιστοποιήσει τις συστάσεις που απευθύνονται σε κάθε άτομο και να επωφεληθεί από τις προτιμήσεις ατόμων με παρόμοια προσωπικότητα. Στη διάγνωση ψυχικών ασθενειών, ορισμένες διαγνώσεις συσχετίζονται με ορισμένα χαρακτηριστικά της προσωπικότητας. Ωστόσο, οι εφαρμογές αναγνώρισης της προσωπικότητας δεν περιορίζονται σε τέτοια προβλήματα. Η απόδοση των Συστημάτων Διαλόγου (Dialogue Systems - DS) μπορεί επίσης να ενισχυθεί εάν το DS είναι σε θέση να εντοπίσει και να προσαρμοστεί στην προσωπικότητα του ατόμου. Επιπλέον, τα χαρακτηριστικά γνωρίσματα της προσωπικότητας μπορούν να υιοθετηθούν από το DS προκειμένου να παράγουν αποκρίσεις που εξαρτώνται από αυτά τα χαρακτηριστικά.

Το κυρίαρχο μοντέλο που προέρχεται από την ψυχολογία για την περιγραφή της ανθρώπινης προσωπικότητας είναι το Big Five μοντέλο (βλ. Ενότητα). Το σύστημα αυτό θεωρείται επαρκές για την περιγραφή της ανθρώπινης προσωπικότητας στη βιβλιογραφία. Συγκεκριμένα, όπως υποστηρίζεται από το [1], “αν χρησιμοποιείται μεγάλος αριθμός κλιμάκων αξιολόγησης και εάν το εύρος των αριθμών αυτών είναι πολύ ευρύ, η ανθρώπινη προσωπικότητα μπορεί να περιγραφεί πλήρως με βάση πέντε ισχυρούς παράγοντες”. Τα βασικά συστατικά του συστήματος προσωπικότητας χαρακτηρίζονται ως “basic tendencies”, “characteristic adaptations” και “self-concept” [2]. Οι χαρακτηριστικές προσαρμογές, ωστόσο, επηρεάζονται από εξωτερικές επιρροές, όπως η κατάσταση ή το πλαίσιο μέσα στο οποίο εξετάζεται η προσωπικότητα. Επομένως, αν και τα χαρακτηριστικά της προσωπικότητας θεωρούνται σταθερά διαχρονικά σε κάποιο βαθμό, συνδέονται ωστόσο με τις συνθήκες υπό τις οποίες εξετάζονται οι συμπεριφορές του ατόμου [3]. Αυτή η θεωρητική παραδοχή αποτελεί την βάση για την ενσωμάτωση διαφόρων πηγών πληροφόρησης που σχετίζονται με την αναγνώριση της προσωπικότητας στο στόχο που επιδιώκεται. Επιπλέον, μια άλλη σημαντική πρόκληση στην αυτόματη αναγνώριση της προσωπικότητας είναι ο περιορισμός που παρουσιάζεται στα επισημασμένα δεδομένα. Δεδομένου ότι η επισήμανση των χαρακτηριστικών προσωπικότητας σε δεδομένα είναι χρονοβόρα διαδικασία, όπου οι προαναφερθείσες επισημάνσεις γίνονται σε ατομικό επίπεδο, τα διαθέσιμα σύνολα δεδομένων για την εκπαίδευση και τον έλεγχο των αυτόματων μεθόδων αναγνώρισης της προσωπικότητας περιέχουν περιορισμένο αριθμό δειγμάτων.

Προηγούμενες επιτυχημένες προσεγγίσεις σχετικά με την αναγνώριση της προσωπικότητας προσδίδουν λεξιλογικές πληροφορίες χρησιμοποιώντας παραδοσιακές και βαθιές αρχιτεκτονικές μάθησης [4, 5, 6]. Ειδικότερα, για το πρόβλημα της αναγνώρισης της προσωπικότητας σε επίπεδο εγγράφου (δηλαδή όταν η πρόβλεψη των χαρακτηριστικών προσωπικότητας ενός ατόμου βασίζεται σε ένα εκτεταμένο έγγραφο, όπως ένας διάλογος, ένας μονολόγος ή ένα δοκίμιο), τα state-of-the-art μοντέλα ενσωματώνουν επίσης ψυχολογικά και συναισθηματικά χαρακτηριστικά, που τυπικά εξάγονται από βάσεις γνώσεων ή λεξικά, σε επίπεδο εγγράφου.

Στο πλαίσιο της διατριβής, προτείνουμε μια προσέγγιση βαθιάς μάθησης χωρίς χειροποίητα χαρακτηριστικά και λεξικά. Συγκεκριμένα, με βάση την υπόθεση ότι η αναγνώριση της προσωπικότητας εξαρτάται από το πλαίσιο μιας δεδομένης κατάστασης, προτείνουμε μια νέα προσαρμογή δύο γνωστών μεθόδων μεταφοράς (Transfer Learning - TL) για την ενσωμάτωση πληροφοριών σχετικών με την αναγνώριση της προσωπικότητας (δηλαδή πρόθεση, συναίσθημα, ψυχολογολογία) στο πρόβλημα-στόχος (π.χ. αναγνώριση της προσωπικότητας). Μεταφέρουμε τις γνώσεις από πηγαία προβλήματα επιπέδου πρότασης στο πρόβλημα-στόχο επιπέδου εγγράφου και διεξάγουμε μια ανάλυση σχετικά με τη συμβολή των διαφόρων πηγών πληροφοριών στην αναγνώριση κάθε χαρακτηριστικού προσωπικότητας συγγραφέων και ομιλητών. Επιτυγχάνουμε state-of-the-art αποτελέσματα κάτω από τα περισσότερα πειραματικά αποτελέσματα και καταλήγουμε στο ποια πηγή πληροφόρησης κυριαρχεί περισσότερο στην αναγνώριση κάθε χαρακτηριστικού προσωπικότητας για διαφορετικές βάσεις δεδομένων.

1.3 Διάρθρωση Διπλωματικής Εργασίας

Η εργασία διαρθρώνεται ως εξής. Στο κεφάλαιο 2 παρουσιάζεται η θεωρία του υποβάθρου. Συγκεκριμένα, παρουσιάζονται λεπτομερώς οι έννοιες της σκοπιμότητας, της πράξης ομιλίας και της DA και παρουσιάζονται οι διάφορες ταξινομίες που έχουν κατασκευαστεί για την ανάλυσή τους. Στη συνέχεια, εισάγονται οι θεωρίες για την ανάλυση του συστήματος προσωπικότητας. Μια επισκόπηση των βασικών τεχνικών και τάσεων στην επεξεργασία φυσικής γλώσσας (Natural Language Processing - NLP) και τεχνητής μάθησης (Machine Learning - ML) ακολουθεί. Το τελευταίο μέρος του Κεφαλαίου 2 είναι αφιερωμένο στην εισαγωγή στην βαθιά μάθηση (Deep Learning - DL) με ιδιαίτερη έμφαση στα αναδρομικά νευρωνικά δίκτυα (Recurrent Neural Networks - RNNs), τα δίκτυα αμφίδρομης βραχυπρόθεσμης μνήμης (Bidirectional Long Short-Term Memory - BiLSTM) και τη μάθηση μεταφοράς (Transfer Learning - TL). Μετά την παρουσίαση του θεωρητικού υπόβαθρου, στα κεφάλαια 3 και 4 παρουσιάζονται οι προτεινόμενες προσεγγίσεις που χρησιμοποιούνται για την αντιμετώπιση των δύο ερευνητικών ερωτημάτων: ταξινόμηση DAs και αναγνώριση προσωπικότητας. Συγκεκριμένα, στο κεφάλαιο 3 παρουσιάζεται η μεθοδολογία και τα μοντέλα που εφαρμόζονται για την αναπαράσταση και την ταξινόμηση DA, παράλληλα με τα αντίστοιχα πειραματικά αποτελέσματα και συμπεράσματα. Στη συνέχεια, στο Κεφάλαιο 4, προτείνεται μια προσέγγιση που εκμεταλλεύεται τις σημασιολογικές πληροφορίες που προέρχονται από αναγνώριση προθέσεων και συναισθημάτων μέσω TL για αναγνώριση προσωπικότητας. Παρουσιάζονται τα αποτελέσματα των διεξαγόμενων πειραμάτων καθώς και τα συμπεράσματα. Τέλος, το κεφάλαιο 5 ολοκληρώνει γενικά και προτείνει μερικές ιδέες ως μελλοντικές εργασίες.

Κεφάλαιο 2

Θεωρητικό Υπόβαθρο

2.1 Αναγνώριση Προθέσεων

2.1.1 Ψυχολογική Ανάλυση της Σκοπιμότητας (intentionality)

Το πρώτο μέρος της ανάλυσης της σκοπιμότητας (intentionality) είναι ο ορισμός της ανθρώπινης επικοινωνίας. Η ανθρώπινη επικοινωνία είναι μια ουσιαστικά συνεργατική επιχείρηση, η οποία λειτουργεί πιο φυσικά και ομαλά όταν υπάρχει

- κοινό εννοιολογικό έδαφος
- κοινά και συνεργατικά επικοινωνιακά κίνητρα

Η ανθρώπινη συνεργασία είναι δομημένη από αυτό που ορισμένοι σύγχρονοι φιλόσοφοι αποκαλούν κοινά σκοπιμότητας ή «εμείς» πρόθεση [7]. Σε γενικές γραμμές, η κοινή πρόθεση είναι το απαραίτητο στοιχείο για τις ανθρώπινες μορφές συνεργατικής δραστηριότητας στην οποία εμπλέκεται ένα πολυπληθές θέμα «εμείς»: οι κοινές επιδιώξεις, οι κοινές προθέσεις, η αμοιβαία γνώση, οι κοινές πεποιθήσεις. Επομένως, η συνεταιριστική επικοινωνία προέκυψε ως ένας τρόπος για πιο αποτελεσματικό συντονισμό αυτών των συνεργατικών δραστηριοτήτων, αρχικά κληρονόμησε και έπειτα συνέβαλε στην οικοδόμηση μιας κοινής ψυχολογικής υποδομής κοινής σκοπιμότητας.

Στην ανθρώπινη επικοινωνία, σύμφωνα με τον Tomasello [8], υπάρχουν τρεις συγκεκριμένες υποθέσεις:

- Η ανθρώπινη συνεργατική επικοινωνία εμφανίστηκε πρωταρχικά κατά την εξέλιξη στις φυσικές, αυθόρμητες χειρονομίες παντομίμας.
- Η ανθρώπινη συνεργατική επικοινωνία επαφίεται σημαντικά σε μια ψυχολογική υποδομή που συμπεριλαμβάνει την κοινή σκοπιμότητα, η οποία ξεκίνησε εξελικτικά για την υποστήριξη συνεργατικών δραστηριοτήτων και η οποία περιλαμβάνει:
 - κοινωνικογνωστικές δεξιότητες για τη δημιουργία μαζί με άλλους κοινών προθέσεων και κοινής προσοχής (και άλλων μορφών κοινού εννοιολογικού εδάφους)
 - προ-κοινωνικά κίνητρα (και ακόμη και κανόνες) για τη βοήθεια και την κοινή χρήση με άλλους
- Η συμβατική επικοινωνία, όπως ενσωματώνεται σε μία ή άλλη ανθρώπινη γλώσσα, είναι δυνατή μόνο όταν οι συμμετέχοντες έχουν ήδη:
 - τις φυσικές χειρονομίες και την κοινή υποδομή σκοπιμότητας
 - δεξιότητες πολιτιστικής μάθησης και απομίμησης για τη δημιουργία και τη μετάβαση σε κοινά κατανοητές επικοινωνιακές συμβάσεις και κατασκευές

Η έννοια της λέξης '*intentionality*', ή σκοπιμότητα, δεν πρέπει να συγχέεται με τη συνήθη έννοια της λέξης «πρόθεση» (*intention*). Όπως αναφέρει η λατινική ετυμολογία του '*intentionality*', η σχετική ιδέα της αμεσότητας ή της έντασης προκύπτει από την προσκόμιση ή την παραμονή σε κάποιο στόχο. Στη μεσαιωνική λογική και τη φιλοσοφία, η λατινική λέξη *intentio* χρησιμοποιήθηκε για ό,τι οι σύγχρονοι φιλοσόφοι και λογικοί σήμερα ονομάζουν μια «έννοια» ή μια «πρόθεση»: κάτι που μπορεί να ισχύει και για τα μη νοητικά πράγματα και τις ιδιότητες - τα πράγματα και τις ιδιότητες που βρίσκονται έξω από μυαλό - και να παρουσιαστεί στο μυαλό. Με την παραδοχή ότι μια ιδέα είναι η ίδια κάτι πνευματικό, ένα *intentio* μπορεί επίσης να είναι αλήθεια για τα ψυχικά πράγματα. Αν και η έννοια του «*intentionality*» στη σύγχρονη φιλοσοφία σχετίζεται με τις έννοιες αυτών των λέξεων όπως «πρόθεση» και «πρόθεση», παρ' όλα αυτά δεν πρέπει να συγχέεται με κανένα από αυτά.

Στη σύγχρονη αγγλική γλώσσα, τα '*intensional*' και '*intensionality*' σημαίνουν '*non-extensional*' και '*non-extensionality*', όπου τόσο η επέκταση όσο και η ένταση είναι λογικά χαρακτηριστικά λέξεων και προτάσεων. Από την άλλη πλευρά, η πρόθεση είναι συγκεκριμένη αντίληψη ότι, αντίθετα με τις πεποιθήσεις, τις κρίσεις, τις ελπίδες, τις επιθυμίες ή τους φόβους, διαδραματίζουν ξεχωριστό ρόλο στην αιτιολογία των πράξεων. Αντίθετα, η σκοπιμότητα είναι ένα διαδομένο χαρακτηριστικό πολλών διαφορετικών ψυχικών καταστάσεων: οι πεποιθήσεις, οι ελπίδες, οι κρίσεις, οι προθέσεις, η αγάπη και το μίσος εκδηλώνουν κάθε σκοπιμότητα. Στην πραγματικότητα, ο Brentano έκρινε ότι η σκοπιμότητα είναι το σήμα κατατεθέν του ψυχισμού [9].

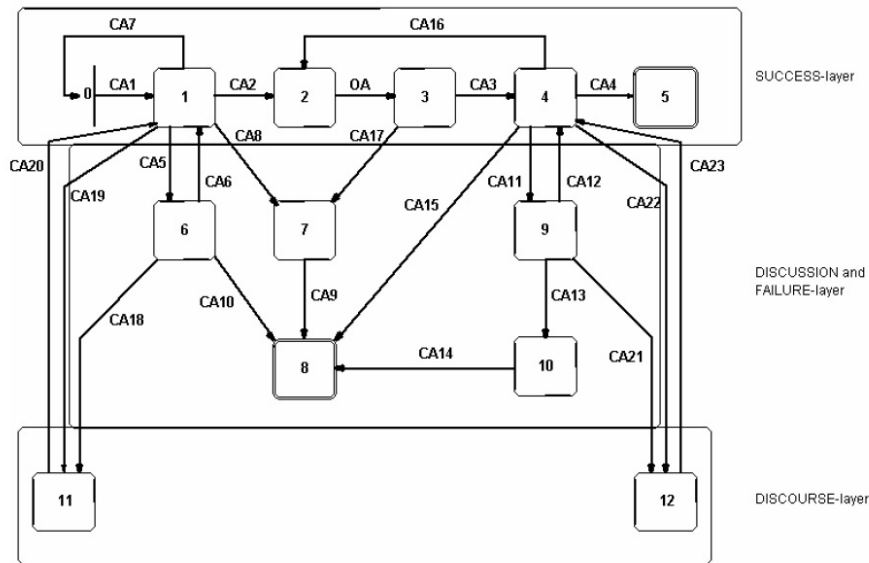
Επιπλέον, αξίζει να γίνει διάκριση μεταξύ επιπέδων σκοπιμότητας. Πολλές από τις ψυχολογικές καταστάσεις ενός ατόμου με σκοπιμότητα (π.χ. πεποιθήσεις) αφορούν μη νοητικά πράγματα, ιδιότητες και καταστάσεις υποθέσεων. Πολλές αφορούν επίσης τις ψυχολογικές καταστάσεις κάποιου άλλου (π.χ. πεποιθήσεις άλλου). Οι πεποιθήσεις για τις πεποιθήσεις των άλλων δείχνουν αυτό που είναι γνωστό ως '*higher-order intentionality*'.

Είναι γεγονός, λοιπόν, ότι ορισμένες ψυχικές καταστάσεις παρουσιάζουν πρόθεση. Αυτό θα συνέβαλε στην εξήγηση του γιατί η συμπεριφορά ενός ατόμου συνίσταται σε ένα ζεύγος του οποίου οι συντεταγμένες είναι αντίστοιχα μια εσωτερική κατάσταση και μία σωματική κίνηση που το κάνει το πρώτο. Ο Dretske [10] έχει υιοθετήσει μια συνθετική άποψη της συμπεριφοράς σύμφωνα με την οποία η συμπεριφορά ενός ατόμου δεν πρέπει να προσδιοριστεί, όπως στη λειτουργιστική αντίληψη, με την σωματική του κίνηση: η συμπεριφορά είναι η διαδικασία κατά την οποία κάποιο σωματικό κίνημα του ατόμου προκαλείται από την εσωτερική του κατάσταση. Όταν η εσωτερική κατάσταση έχει πρόθεση, η συμπεριφορά του ατόμου είναι σκόπιμη. Από τη συνθετική άποψη της συμπεριφοράς, η σκοπιμότητα της ψυχικής κατάστασης ενός ατόμου δεν είναι σχετική με την αιτιώδη συνάφεια μιας συγκεκριμένης κίνησης στο χρόνο *t*. Αντίθετα, είναι σχετικό με το γιατί τα είδη των κινήσεων προκαλούνται τακτικά από τύπους εκ προθέσεως καταστάσεων.

2.1.2 Ορισμός των Επικοινωνιακών Πράξεων (Communicative Acts - CAs)

Η επικοινωνία είναι μια διαδικασία που στοχεύει στη συμφωνία σχετικά με προκαθορισμένα χαρακτηριστικά μιας κατάστασης. Όταν η συμφωνία δεν είναι άμεση, απαιτείται συζήτηση για την επίλυση των σημείων διαφωνίας χρησιμοποιώντας επιχειρηματολογία. Με άλλα λόγια, η έννοια της λειτουργίας μιας επικοινωνιακής δράσης ορίζεται από την άποψη της ροής πληροφοριών μεταξύ του ομιλητή και του παραλήπτη. Αυτός είναι ο λόγος για τον οποίο η θεωρία είναι σημαντική για να κατανοήσουμε πώς η επικοινωνία πρόκειται να εξελιχθεί. Η θεωρία της επιχειρηματολογίας είναι η διεπιστημονική μελέτη του τρόπου με τον οποίο μπορούν να επιτευχθούν συμπεράσματα μέσω της λογικής. Αυτό επιτυγχάνεται μέσω ισχυρισμών που βασίζονται σε υποθέσεις. Περιλαμβάνει την τέχνη και τις επιστήμες της πολιτικής συζήτησης, του διαλόγου, της λογικής και των διαδικαστικών κανόνων τόσο σε τεχνητό όσο και σε πραγματικό περιβάλλον. Τα αποτελέσματα αυτής της θεωρίας

σε συνδυασμό με την προσέγγιση γλώσσας/δράσης (Language/Action Perspective -LAP) μας παρέχουν ένα μοντέλο, το οποίο αναπτύσσεται περαιτέρω σε αντιπαράθεση με το γνωστό σύστημα Issue Based Information System (IBIS) του Conklin [11], σύμφωνα με το οποίο δεν υπάρχει τέλεια λύση για ένα κακό πρόβλημα. Με βάση την προσέγγιση LAP, η γλώσσα θεωρείται ως η πρωταρχική διάσταση της ανθρώπινης συνεταιριστικής δραστηριότητας. Το τελικό μοντέλο είναι ένα μοντέλο που ονομάζεται 3-box και προτείνεται ως επέκταση του μοντέλου διαδικασίας συναλλαγών (Transaction Process Model - TPM) του Reijswoud [12] και οι σημαντικότερες ενέργειές του είναι η προσθήκη αξιώσεων, επιχειρημάτων, πλεονεκτημάτων και μειονεκτημάτων. Το μοντέλο του Reijswoud είναι το ακόλουθο, όπου οι CA είναι επικοινωνιακές δράσεις. Για παράδειγμα, το CA1 είναι ένα αίτημα και η CA2 την αποδοχή του αιτήματος. Η εναλλακτική λύση για το CA2 είναι η CA5 (αιτιολόγηση της αίτησης), μέσω της οποίας ο Hearer μετακινείται στη στρώση "Συζήτηση και αποτυχία".



Σχήμα 2.1: Επισκόπηση του μοντέλου διεργασίας συναλλαγών

Έτσι, όπως παρατηρήθηκε στο Σχήμα 2.1, το TPM έχει τρία επίπεδα επικοινωνίας, το στρώμα επιτυχίας, το στρώμα συζήτησης και αποτυχίας και το στρώμα λόγου, με το οποίο μπορούν να αναζωογονηθούν ή να προσαρμοστούν τα πρότυπα και οι κανόνες. Στο LAP, η σημασιολογία των συνομιλιών συνήθως περιγράφεται με όρους πράξεων ομιλίας και των αποτελεσμάτων αυτών από άποψη πεποιθήσεων, προθέσεων και υποχρεώσεων [13].

Ωστόσο, οι ενέργειες επικοινωνίας αλληλεπιδρούν με το κοινό έδαφος [14, 15]. Συγκεκριμένα, η σημασιολογία των επικοινωνιακών ενεργειών δίνεται με όρους αξιώσεων και οι ισχυρισμοί αυτοί υποστηρίζονται από το κοινό υπόβαθρο της κοινότητας. Οι επιπτώσεις των επικοινωνιακών ενεργειών προσελκύουν και το κοινό έδαφος: μόλις γίνει δεκτή η αξίωση, το περιεχόμενό της γίνεται αμοιβαία αποδεκτό. Ως αποτέλεσμα, η επιχειρηματολογία που λαμβάνει χώρα κατά τη διάρκεια της επικοινωνίας μπορεί να θεωρηθεί ως η διαδικασία δημιουργίας μιας γέφυρας ανάμεσα σε αυτό που έχει συμφωνηθεί στην κοινότητα (ή τη σχέση) με τη μορφή κοινής βάσης (κοινά πρότυπα, υπό την έννοια του Stamper [16]) και την κατάσταση στο χέρι ("γείωση").

Η πρόκληση που σχετίζεται με τις πράξεις επικοινωνίας είναι η ανάγκη άμεσης διαδικασίας και αντίδρασης του συστήματος σε αυτό που λέει ο ομιλητής. Η δυναμική λογική των κατηγορούμενων από τους Jeroen Groenendijk και Martin Stokhof [17] είναι χρήσιμη για την online σημασιολογική ερμηνεία. Το σύστημα αυτό έρχεται ως αποτέλεσμα της διατύπωσης και διερεύνησης μιας δυναμικής σημασιολογικής ερμηνείας της γλώσσας της λογικής πρώτης τάξης και θεωρείται ως ένα πρώτο βήμα

προς μια σύνθεση, μη αντιπροσωπευτική θεωρία της σημασιολογίας του λόγου.

2.1.3 Ορισμός των Πράξεων Ομιλίας (Speech Acts) & των Διαλογικών Πράξεων (Dialogue Acts - DAs)

Η ιδέα ότι η ανθρώπινη συζήτηση περιέχει πράξεις ομιλίας προέρχεται από τους κοινωνικο-γλωσσικούς θεωρητικούς [18, 19]. Η θεωρία των διαλογικών πράξεων (DAs) υποδεικνύει ότι οι άνθρωποι όχι μόνο γνωστοποιούν τις πραγματικές πληροφορίες μέσα από τις εκφράσεις της φυσικής γλώσσας, αλλά επίσης εκφράζουν συχνά την υποκείμενη προβλεπόμενη δράση.

Το πρώτο βήμα της επεξεργασίας του διαλόγου είναι η επισήμανση των DAs, η ανάθεση μιας λειτουργικής ετικέτας στην είσοδο του χρήστη με στόχο την αναπαράσταση των επικοινωνιακών προθέσεων πίσω από κάθε έκφραση. Ωστόσο, σύμφωνα με την προσέγγιση που βασίζεται στην εξατομίκευση, οι προθέσεις μπορούν επίσης να βασιστούν στην ανάλυση του πλήρους διαλόγου, παρά σε μια ενιαία έκφραση, ώστε να βρεθεί μια συνεπής σημασιολογική εκπροσώπηση που να καταγράφει το νόημα του διαλόγου. Είτε έτσι είτε αλλιώς αυτό το πρώτο βήμα είναι κρίσιμο για ένα αυτοματοποιημένο σύστημα, ώστε να είναι σε θέση να παράγει μια κατάλληλη απάντηση. Υπάρχει ένα ευρύ φάσμα χρήσεων των DAs, που περιλαμβάνουν παραστάσεις της πραγματικής σημασίας των ρημάτων στις θεωρίες διαλόγου [20, 21, 22, 23], δομικές μονάδες διαλόγου [24, 25], ετικέτες για σχολιασμό κορμού [26, 27], γλώσσες επικοινωνίας μεταξύ αυτόματων συστημάτων [28, 29], αντικείμενα ανάλυσης σε συστήματα διαλόγου [30, 31] και στοιχεία της λογικής θεωρίας της ορθολογικής αλληλεπίδρασης [32]. Ωστόσο, εξακολουθεί να υπάρχει δυσκολία στη δημιουργία μιας ταξινόμησης των DAs που μπορούν να κατανοηθούν και να χρησιμοποιηθούν από άλλους ερευνητές εκτός από τους σχεδιαστές της συγκεκριμένης ταξινόμησης των DAs. Αυτή η δυσκολία απορρέει από τις διαφορετικές ερμηνείες που αναθέτουν οι ερευνητές στις διαφορετικές ταξινομήσεις των DAs. Αυτό το είδος σύγχυσης έχει οδηγήσει ορισμένους (π.χ. [33, 29]) να προτείνουν πρότυπες θεωρίες που θα μπορούσαν να προσδιοριστούν καλά, να κατανοηθούν και να χρησιμοποιηθούν σε ομάδες, ενώ άλλοι (π.χ. [34, 35]) προτιμούν να θεωρούν τον προσδιορισμό του διαλόγου ως δευτερεύουσα σημασία, μέσα σε μια γενικότερη θεωρία της ορθολογικής αλληλεπίδρασης, χρησιμοποιώντας τις έννοιες ως πρωτόγονα. Ένα άλλο σημαντικό ζήτημα είναι η αναγνώριση των DAs σε ένα διάλογο μεταξύ ενός συστήματος και ενός ατόμου. Η ακριβής αναγνώριση των DAs από ένα σύστημα διαλόγου απαιτεί ένα καλά σχεδιασμένο σύστημα κατανόησης γλωσσών. Προκειμένου να σχεδιαστεί ένα τέτοιο σύστημα, θα πρέπει να ληφθούν υπόψη οι ακόλουθοι παράμετροι: η σύνταξη, δηλαδή οι σχέσεις μεταξύ των ρημάτων και της δομής των φράσεων, η σημασιολογία, δηλαδή η αναφορά και η ελλειψία, και οι πραγματισμοί, δηλαδή η ανάλυση των διαλόγων ανταλλαγής πληροφοριών επικοινωνιακών δράσεων.

Το ερώτημα που τίθεται τώρα είναι πώς χρησιμοποιούνται οι πράξεις στην πράξη για την εφαρμογή ενός συστήματος συνομιλίας. Οι πράξεις θεωρούνται μεταβάσεις από καταστάσεις σε καταστάσεις, ενώ ο διάλογος λειτουργεί ως ειδικές περιπτώσεις ενεργειών. Οι θεωρίες της δράσης που προτείνονται από έρευνες της τεχνητής νοημοσύνης (Artificial Intelligence - AI) γενικά συνδέουν διάφορα σύνολα με δράσεις: ένα σύνολο επιδράσεων (περιορισμοί στην προκύπτουσα κατάσταση), ένα σύνολο προκαταρκτικών συνθηκών (περιορισμοί στην αρχική κατάσταση) και αποσυνθέσεις (συνιστούν τη δράση).

Με βάση τον παραπάνω ορισμό της δράσης, θα πρέπει να καθοριστούν οι πτυχές της κατάστασης που σχετίζονται με τις πιθανές συνθήκες για τον καθορισμό των επιδόσεων των DAs και εκείνων που επηρεάζονται άμεσα. Οι όροι και οι συνέπειες των DAs που χρησιμοποιούνται είναι οι εξής:

- η έννοια της κατάστασης διαλόγου, όπως κωδικοποιείται ως κατάσταση στη γραμματική του διαλόγου [24, 36].

- τις ψυχικές καταστάσεις (π.χ. πεποίθηση, πρόθεση) του ομιλητή και του παραλήπτη [37, 38], που είναι η πιο δημοφιλής στην προσέγγιση σχεδιασμού
- τις κοινωνικές υποχρεώσεις και τις δεσμεύσεις που ανέλαβαν οι συμμετέχοντες στο διάλογο [39, 23, 37]

2.1.4 Εφαρμογές & Ταξινόμηση των DAs

Οι εφαρμογές όπου απαιτούνται συστήματα διαλόγου κυμαίνονται από απλές εργασίες όπως η λειτουργία μιας οικιακής συσκευής ή η κράτηση μιας πτήσης έως πιο πολύπλοκα καθήκοντα όπως ο έλεγχος ενός έξυπνου δωματίου ή η διαχείριση της οδικής κυκλοφορίας. Λόγω της πολυπλοκότητας της διαχείρισης των γλωσσικών διεπαφών και της ισχυρής τους εξάρτησης από το πλαίσιο αλληλεπίδρασης, κάθε εφαρμογή ή σύνολο εφαρμογών απαιτεί την ανάπτυξη ενός συγκεκριμένου μοντέλου. Αυτός είναι ο λόγος για τον οποίο οι περισσότερες σύγχρονες μέθοδοι δημιουργίας πρωτοτύπων περιορίζονται στην ανάπτυξη συστημάτων διαλόγου που εργάζονται σε μια ενιαία εφαρμογή ή σε ένα μικρό σύνολο εφαρμογών. Το πρωτότυπο του διαλόγου αντιπροσωπεύει επομένως ένα σημαντικό κομμάτι στην αναπτυξιακή διαδικασία του διαδραστικού συστήματος, ειδικά για εκείνους με φωνητική διεπαφή: υπάρχει μια ισχυρή ανάγκη για μια αποτελεσματική Rapid Dialogue Prototype Methodology (RPDM) [40]. Το RDPM περιλαμβάνει πέντε βασικά βήματα: 1. παραγωγή ενός μοντέλου εργασίας για την στοχευμένη εφαρμογή 2. Απόκτηση ενός αρχικού μοντέλου διαλόγου από το μοντέλο εργασίας που ελήφθη 3. πραγματοποίηση πειράματος Wizard-of-Oz για τη βελτίωση του αρχικού μοντέλου διαλόγου 4. Διεξαγωγή εσωτερικού δοκιμαστικού πεδίου για περαιτέρω βελτίωση του μοντέλου διαλόγου (αναδιατύπωση των μηνυμάτων του συστήματος, βελτιωμένη ανάδραση κ.λπ.) και επικύρωση της διαδικασίας αξιολόγησης (συνοχή, κατανόηση) 5. Διεξαγωγή εξωτερικού δοκιμαστικού πεδίου για την αξιολόγηση του τελικού μοντέλου διαλόγου σύμφωνα με τη διαδικασία αξιολόγησης που καθορίστηκε κατά τη διάρκεια της εσωτερικής δοκιμής πεδίου.

Υπάρχει ένα ευρύ φάσμα εφαρμογών των DAs που βασίζονται σε αυτόματη διαπίστωση διαλογικών πράξεων. Ανάμεσά τους, τα σημαντικότερα είναι: συστήματα διαλόγου, μηχανική μετάφραση, αυτόματη αναγνώριση ομιλίας (Automatic Speech Recognition - ASR), ταυτοποίηση θέματος και κινούμενη εικόνα του κεφαλιού συνομιλίας. Σε συστήματα διαλόγου, οι DAs μπορούν να χρησιμοποιηθούν για να αναγνωρίσουν την πρόθεση του χρήστη, για παράδειγμα όταν ο χρήστης ζητά κάποια πληροφορία και τον περιμένει ή όταν το σύστημα προσπαθεί να ερμηνεύσει την ανατροφοδότηση από το χρήστη. Ένα παράδειγμα συστήματος διαχείρισης διαλόγου που χρησιμοποιεί DA ταξινόμηση είναι το σύστημα VERBMOBIL.

Στη μηχανική μετάφραση, οι πράξεις διαλόγου μπορεί να είναι χρήσιμες για την επιλογή της βέλτιστης λύσης όταν υπάρχουν αρκετές μεταφράσεις. Ειδικότερα, η γραμματική μορφή μιας έκφρασης μπορεί να εξαρτάται από την πρόθεσή της. Η αυτόματη ανίχνευση των πράξεων διαλόγου μπορεί να χρησιμοποιηθεί στην ASR για να αυξήσει την ακρίβεια αναγνώρισης λέξεων. Μια κεφαλή ομιλίας είναι ένα μοντέλο ανθρώπινης κεφαλής που αναπαράγει την ομιλία ενός ομιλητή σε πραγματικό χρόνο. Μπορεί επίσης να κάνει εκφράσεις του προσώπου που σχετίζονται με την τρέχουσα κατάσταση του λόγου. Η αξιοποίηση της αναγνώρισης DA σε αυτό το πλαίσιο μπορεί να κάνει την κίνηση πιο φυσική, για παράδειγμα αυξάνοντας τα φρύδια όταν ζητηθεί μια ερώτηση. Μια άλλη πιο εύκολη επιλογή είναι να δειχθούν αυτές οι συμπληρωματικές πληροφορίες με σύμβολα και χρώματα κοντά στο κεφάλι.

Λόγω των διαφορών στις απαιτήσεις αυτών των εφαρμογών όσον αφορά το απαραίτητο σύνολο ετικετών DA, ο ορισμός του συνόλου των χρησιμοποιούμενων DAs είναι ένα σημαντικό, αλλά δύσκολο βήμα. Αυτό συμβαίνει επειδή προκύπτει ως συμβιβασμός μεταξύ τριών αντικρουόμενων απαιτήσεων:

- Οι ετικέτες DA πρέπει να είναι αρκετά γενικές ώστε να είναι χρήσιμες για διαφορετικές εργασίες ή τουλάχιστον ανθεκτικές στην απρόβλεπτη μεταβλητότητα και εξέλιξη της εφαρμογής-στόχου.
- Οι ετικέτες DA πρέπει να είναι αρκετά συγκεκριμένες ώστε να κωδικοποιούν λεπτομερή και εκμεταλλεύσιμα χαρακτηριστικά της εφαρμογής-στόχου.
- Οι ετικέτες DA πρέπει να είναι σαφείς και εύκολα να διαχωρίζονται, προκειμένου να μεγιστοποιηθεί η συμφωνία μεταξύ των ανθρωπίνων ετικετών.

Πολλά διαφορετικά σετ επισήμανσης των DAs μπορούν να βρεθούν στη βιβλιογραφία. Πρόσφατα, μερικά τέτοια σετ φαίνεται να εμφανίζονται ως μια κοινή γραμμή βάσης, από την οποία προέρχονται οι ετικέτες των DAs που σχετίζονται με την εκάστοτε εφαρμογή. Τα σετ αυτά είναι το Dialogue Act Markup in Several Layers (DAMSL) [41], το Switchboard DAMSL (SWBD-DAMSL) [42], το Meeting Recorder [43], το VERBMOBIL [44] και το Map-Task [45].

Το DAMSL σχεδιάστηκε αρχικά για να είναι καθολικό. Η επισήμανση αυτού του σετ αποτελείται από τέσσερα επίπεδα (ή διαστάσεις): επικοινωνιακή κατάσταση, επίπεδο πληροφόρησης, προσανατολισμό προς τα εμπρός και λειτουργίες προς τα πίσω. Γενικά, αυτές οι διαστάσεις θεωρούνται ορθογώνιες και πρέπει να είναι δυνατή η δημιουργία παραδειγμάτων για κάθε πιθανό συνδυασμό αυτών. Η επικοινωνιακή κατάσταση δηλώνει εάν η έκφραση είναι ακατανόητη, έχει εγκαταλειφθεί ή είναι αυτο-ομιλία. Αυτή η λειτουργία δεν χρησιμοποιείται για τις περισσότερες δηλώσεις. Το επίπεδο πληροφοριών παρέχει έναν αφηρημένο χαρακτηρισμό του περιεχομένου της έκφρασης. Αποτελείται από τέσσερις κατηγορίες: εργασία, διαχείριση εργασιών, επικοινωνία-διαχείριση και άλλο επίπεδο. Οι προσανατολισμένες προς το μέλλον λειτουργίες οργανώνονται σε μια ταξινόμηση, με παρόμοιο τρόπο με τις δράσεις στην παραδοσιακή θεωρία των ενεργειών ομιλίας (speech acts). Οι αναδρομικές λειτουργίες δείχνουν τη σχέση μεταξύ της τρέχουσας έκφρασης και των προηγούμενων πράξεων διαλόγου, όπως η αποδοχή μιας πρότασης ή η απάντηση στην ερώτηση. Το SWBD-DAMSL είναι η προσαρμογή του DAMSL στον τομέα των τηλεφωνικών συνομιλιών. Οι περισσότερες από τις ετικέτες SWBD-DAMSL αντιστοιχούν στις ετικέτες DAMSL.

Το σετ Meeting Recorder (MRDA) βασίζεται στην ταξινόμηση SWBD-DAMSL. Το σώμα MRDA περιέχει περίπου 72 ώρες φυσικά εξελισσόμενων πολυ-ατομικών συναντήσεων με χειροκίνητες επισημάνσεις των DA ετικετών. Οι συναντήσεις περιλαμβάνουν περιοχές με υψηλή αλληλεπικάλυψη ομιλητών, συναισθηματική παραλλαγή, σύνθετες δομές αλληλεπίδρασης, διακοπείσες φράσεις και άλλα ενδιαφέροντα φαινόμενα στροφορμής και λόγου.

Η ιεραρχία DA στο VERBMOBIL οργανώνεται ως δέντρο απόφασης. Αυτή η δομή έχει επιλεγεί για να διευκολύνει τη διαδικασία επισήμανσης και για να διευκρινίσει τις σχέσεις μεταξύ διαφορετικών DAs.

Οι ετικέτες των DAs στη βάση δεδομένων του Map-Task είναι δομημένες σε τρία επίπεδα, τις υψηλότερες συναλλαγές μοντελοποίησης, όπου κάθε συναλλαγή πραγματοποιεί ένα σημαντικό βήμα στο σχέδιο ομιλητών. Οι συναλλαγές αποτελούνται στη συνέχεια από παιχνίδια συνομιλίας, τα οποία υποδηλώνουν την κανονικότητα μεταξύ ερωτήσεων / απαντήσεων, δηλώσεων / αρνήσεων ή αποδοχής κ.ο.κ. Τα παιχνίδια τελικά αποτελούνται από συνομιλητικές κινήσεις, οι οποίες ταξινομούν διαφορετικά είδη παιχνιδιών σύμφωνα με τους σκοπούς τους.

2.1.5 Σύγχρονες Εφαρμογές Συστημάτων Διαλόγου

Διαλογικά συστήματα προσανατολισμένα σε συγκεκριμένο πρόβλημα. Αυτά είναι τα διαλογικά συστήματα τα οποία είναι σχεδιασμένα για ένα συγκεκριμένο πρόβλημα και προγραμματισμένα

ώστε να έχουν σύντομες συνομιλίες για να λάβουν πληροφορίες από τον χρήστη με σκοπό την ολοκλήρωση της εργασίας. Σε αυτά τα συστήματα συμπεριλαμβάνονται οι ψηφιακοί βοηθοί που βρίσκονται τώρα σε κάθε κινητό τηλέφωνο ή στους ελεγκτές κατοικίας (Siri, Cortana, Alexa, Google Now / Home κ.λπ.) των οποίων οι διαλογικοί βοηθοί μπορούν να προσφέρουν ταξιδιωτικές οδηγίες, εγχώριες συσκευές ελέγχου, εύρεση εστιατορίων, βοήθεια στην πραγματοποίηση τηλεφωνικών κλήσεων ή αποστολή κειμένων. Οι εταιρείες αναπτύσσουν βοηθούς συνομιλίας με βάση τους στόχους των εκάστοτε ιστοτόπων ώστε να βοηθήσουν τους πελάτες, να απαντούν σε ερωτήσεις ή να αντιμετωπίζουν προβλήματα. Οι συνομιλητές παίζουν ένα σημαντικό ρόλο ως διεπαφή στα ρομπότ. Και έχουν ακόμη και αιτήσεις για το κοινωνικό καλό. Το DoNotPay είναι ένας “robot lawyer” που βοηθά τους ανθρώπους να αμφισβητήσουν λανθασμένα πρόστιμα για παρκάρισμα, να υποβάλουν αίτηση για στέγαση έκτακτης ανάγκης ή να ζητήσουν άσυλο εάν είναι πρόσφυγες.

Chatbots είναι συστήματα τα οποία είναι σχεδιασμένα για εκτεταμένες συζητήσεις και προγραμματισμένα να μιμούνται αδόμητες συζητήσεις ή «συνομιλίες» χαρακτηριστικές της αλληλεπίδρασης ανθρώπου-ανθρώπου, αντί να επικεντρώνονται σε μια συγκεκριμένη εργασία όπως η κράτηση αεροπορικών πτήσεων. Αυτά τα συστήματα συχνά έχουν μια ψυχαγωγική αξία, όπως το σύστημα XiaoIce της Microsoft, το οποίο συνομιλεί με ανθρώπους σε πλατφόρμες ανταλλαγής μηνυμάτων κειμένου. Τα chatbots συχνά επιχειρούν να περάσουν διάφορες μορφές του Turing test. Ακόμα ξεκινώντας από το πρώτο σύστημα, το ELIZA [46], τα chatbots έχουν επίσης χρησιμοποιηθεί για πρακτικούς σκοπούς, όπως η δοκιμή ψυχολογικής συμβουλευτικής. Σημειώστε ότι η λέξη «chatbot» χρησιμοποιείται συχνά στα μέσα μαζικής ενημέρωσης και στη βιομηχανία ως συνώνυμο του βοηθού συνομιλίας.

2.2 Αναγνώριση Προσωπικότητας

2.2.1 Θεωρίες της Προσωπικότητας

Υπάρχει μια διακύμανση μεταξύ των διαφορετικών ορισμών που έχουν δοθεί για την ανθρώπινη προσωπικότητα. Σύμφωνα με τον Allport [47], «η προσωπικότητα είναι η δυναμική οργάνωση στο άτομο εκείνων των ψυχοφυσικών συστημάτων που καθορίζουν τα χαρακτηριστικά της συμπεριφοράς και της σκέψης», ενώ ο Weinbeirg και ο Gould υποστηρίζουν ότι η προσωπικότητα αποτελείται από τα χαρακτηριστικά ή το μείγμα χαρακτηριστικών που κάνουν μοναδικό πρόσωπο. Και οι δύο αυτοί ορισμοί υπογραμμίζουν τη μοναδικότητα του ατόμου και κατά συνέπεια υιοθετούν μια ιδεογραφική άποψη. Η ιδεογραφική άποψη υποθέτει ότι κάθε άτομο έχει μια μοναδική ψυχολογική δομή, ότι ορισμένα γνωρίσματα κατέχονται από ένα μόνο άτομο και ότι υπάρχουν φορές που είναι αδύνατο να συγκρίνουμε ένα άτομο με άλλους. Η άποψη αυτή τείνει να χρησιμοποιεί μελέτες περιπτώσεων για τη συλλογή πληροφοριών. Η νοοθετική άποψη [48], από την άλλη, τονίζει τη συγκρισιμότητα μεταξύ των ατόμων. Αυτή η άποψη βλέπει τα χαρακτηριστικά να έχουν το ίδιο ψυχολογικό νόημα σε όλους. Αυτή η προσέγγιση τείνει να χρησιμοποιεί ερωτήματα προσωπικής αναφοράς, ανάλυση παραγόντων κλπ. Οι άνθρωποι διαφέρουν στις θέσεις τους κατά μήκος μιας συνεχούς κλίμακας στο ίδιο σύνολο χαρακτηριστικών.

Ο [49], από την άλλη, τονίζει τη συγκρισιμότητα μεταξύ των ατόμων. Αυτή η άποψη βλέπει τα χαρακτηριστικά να έχουν το ίδιο ψυχολογικό νόημα σε όλους. Αυτή η προσέγγιση τείνει να χρησιμοποιεί ερωτήματα προσωπικής αναφοράς, ανάλυση παραγόντων κλπ. Οι άνθρωποι διαφέρουν στις θέσεις τους κατά μήκος ενός συνεχούς στο ίδιο σύνολο χαρακτηριστικών.

Οι θεωρίες της προσωπικότητας υποδηλώνουν η προσωπικότητα είναι βιολογικό χαρακτηριστικό, ενώ οι κρατικές θεωρίες, όπως το [50] της Bandura, τονίζουν το ρόλο της γαλουχίας και της περιβαλλοντικής επιρροής. Η ψυχοδυναμική θεωρία της προσωπικότητας του Sigmund Freud υποθέτει ότι υπάρχει αλληλεπίδραση μεταξύ της φύσης (έμφυτα ένστικτα) και της γαλουχίας (γονικές επιρροές).

Θεωρία του Freud: Τριμερής Θεωρία της Προσωπικότητας

Σύμφωνα με τον Freud, η προσωπικότητα περιλαμβάνει πολλούς παράγοντες:

- Ενστικτώδεις δυνάμεις
- Υποσυνείδητες διεργασίες
- Επιρροές της πρώιμης παιδικής ηλικίας

Η ανάπτυξη της προσωπικότητας εξαρτάται από την αλληλεπίδραση του ενστίκτου και του περιβάλλοντος κατά τα πρώτα πέντε χρόνια της ζωής. Η γονική συμπεριφορά είναι ζωτικής σημασίας για την κανονική και την ανώμαλη ανάπτυξη.

Ο Freud [51] είδε την προσωπικότητα δομημένη σε τρία μέρη (δηλ., Τριμερή) που όλα αναπτύσσονται σε διαφορετικά στάδια της ζωής μας και ξεχωριστά από τον εγκέφαλο:

- Το **id** είναι το πρωτόγονο και ενστικτώδες στοιχείο της προσωπικότητας. Αποτελείται από όλες τις κληρονομικές (δηλαδή βιολογικές) συνιστώσες της προσωπικότητας, όπως το ένστικτο φύλου (ζωής) και το επιθετικό (θάνατο) ένστικτο.
- Το **ego** αναπτύσσεται προκειμένου να μεσολαβήσει μεταξύ του μη ρεαλιστικού id και του εξωτερικού πραγματικού κόσμου (όπως ένας διαιτητής). Είναι η συνιστώσα της λήψης αποφάσεων της προσωπικότητας.
- Το **superego** ενσωματώνει τις αξίες και τα ηθικά της κοινωνίας που μαθαίνονται από τους γονείς και άλλους. Είναι παρόμοια με τη συνείδηση, η οποία μπορεί να τιμωρήσει το ego προκαλώντας συναισθήματα ενοχής.

Προσέγκυση της Προσωπικότητας

Αυτή η προσέγγιση υποθέτει ότι η συμπεριφορά καθορίζεται από σχετικά σταθερά χαρακτηριστικά τα οποία αποτελούν τις θεμελιώδεις μονάδες της προσωπικότητας κάποιου. Τα χαρακτηριστικά προδιαθέτουν κάποιον να ενεργεί με κάποιο τρόπο, ανεξάρτητα από την κατάσταση. Αυτό σημαίνει ότι τα γνωρίσματα πρέπει να παραμένουν συνεπή σε καταστάσεις και με την πάροδο του χρόνου, αλλά μπορεί να διαφέρουν μεταξύ των ατόμων. Υποτίθεται ότι τα άτομα διαφέρουν ως προς τα χαρακτηριστικά τους λόγω γενετικών διαφορών. Αυτές οι θεωρίες μερικές φορές αναφέρονται σε ψυχομετρικές θεωρίες, λόγω της έμφασης που δίνουν στην μέτρηση της προσωπικότητας με τη χρήση ψυχομετρικών εξετάσεων. Οι βαθμολογίες των χαρακτηριστικών είναι συνεχείς (ποσοτικές) μεταβλητές. Ένα άτομο λαμβάνει αριθμητική βαθμολογία για να υποδείξει πόσο από ένα χαρακτηριστικό που κατέχουν. Όλες οι άλλες θεωρίες που αναπτύχθηκαν (βλέπε παρακάτω υποενότητες) βασίζονται σε αυτή την παραδοχή.

Η θεωρία της Προσωπικότητας του Eysenck

Ο Eysenck [52, 53, 54] πρότεινε μια θεωρία της προσωπικότητας που βασίζεται σε βιολογικούς παράγοντες, υποστηρίζοντας ότι τα άτομα κληρονομούν έναν τύπο νευρικού συστήματος που επηρεάζει την ικανότητά τους να μαθαίνουν και να προσαρμόζονται στο περιβάλλον. Κατά τη διάρκεια της δεκαετίας του 1940 ο Eysenck εργάστηκε στο ψυχιατρικό νοσοκομείο Maudsley στο Λονδίνο. Η δουλειά του ήταν να κάνει μια αρχική αξιολόγηση για κάθε ασθενή προτού διαγνωστεί από ψυχίατρο η ψυχική του διαταραχή. Μέσα από αυτή τη θέση, συνέταξε μια σειρά από ερωτήσεις σχετικά με τη συμπεριφορά, την οποία ακολούθως εφάρμοσε σε 700 στρατιώτες οι οποίοι υποβάλλονταν σε

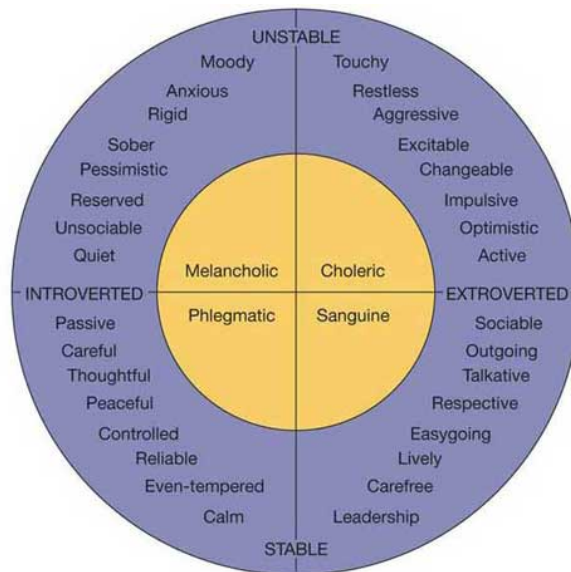
θεραπεία για νευρωτικές διαταραχές στο νοσοκομείο και διαπίστωσε ότι οι απαντήσεις των στρατιωτών φαινόταν να συνδέονται φυσικά μεταξύ τους, υποδηλώνοντας ότι υπήρχαν διάφορα χαρακτηριστικά γνωρίσματα προσωπικότητας τα οποία αποκαλύπτονταν από τις απαντήσεις του στρατιώτη και ονόμασε αυτά τα χαρακτηριστικά προσωπικότητας πρώτης τάξης Χρησιμοποίησε μια τεχνική που ονομάζεται ανάλυση παράγοντα. Αυτή η τεχνική μειώνει τη συμπεριφορά σε διάφορους παράγοντες οι οποίοι μπορούν να ομαδοποιηθούν σε ξεχωριστές επικεφαλίδες, που ονομάζονται διαστάσεις. Ο Eysenck [55] διαπίστωσε ότι η συμπεριφορά τους θα μπορούσε να εκπροσωπείται από δύο διαστάσεις: Εσωστρέφεια / Εξωστρέφεια (E); Νευροθεσία / Σταθερότητα (N). Ο Eysenck χαρακτήρισε αυτά τα χαρακτηριστικά προσωπικότητας δεύτερης τάξης. Κάθε πτυχή της προσωπικότητας (εξωστρέφεια, νευρωτισμός και ψυχωσιμός) μπορεί να ανιχνευθεί πίσω σε μια διαφορετική βιολογική αιτία. Η προσωπικότητα εξαρτάται από την ισορροπία μεταξύ διεγέρσεως και αναστολής του αυτόνομου νευρικού συστήματος (Autonomous Neural System - ANS).

- **Εσωστρέφεια/Εξωστρέφεια (Extraversion/introversion):** Οι εξωστρεφείς είναι κοινωνικοί και επιθυμούν τον ενθουσιασμό και την αλλαγή, και έτσι μπορούν να βαρεθούν εύκολα. Τείνουν να είναι ανέμελοι, αισιόδοξοι και παρορμητικοί. Είναι πιο πιθανό να βρεθούν σε κινδύνους και να αναζητήσουν συγκίνηση. Ο Eysenck υποστηρίζει ότι αυτό οφείλεται στο γεγονός ότι κληρονομούν ένα υποβραχιζόμενο νευρικό σύστημα και έτσι επιδιώκουν τη διέγερση για να αποκαταστήσει το επίπεδο της βέλτιστης διέγερσης. Οι εσωστρεφείς, από την άλλη πλευρά, βρίσκονται στο άλλο άκρο αυτής της κλίμακας, είναι ήσυχοι και δεσμευμένοι. Αυτά είναι ήδη υπερβολικά προκλητοί και αποπνέουν αίσθηση και διέγερση. Οι εσωστρεφείς επιφυλάσσονται, σχεδιάζουν τις ενέργειές τους και ελέγχουν τα συναισθήματά τους. Τείνουν να είναι σοβαροί, αξιόπιστοι και απαισιόδοξοι.
- **Νευρωτισμός/Σταθερότητα (Neuroticism/stability):** Το επίπεδο νευρωτισμού ενός ατόμου καθορίζεται από την αντιδραστικότητα του συμπαθητικού νευρικού συστήματος. Το νευρικό σύστημα ενός σταθερού ατόμου θα είναι γενικά λιγότερο αντιδραστικό στις αγχωτικές καταστάσεις, παραμένοντας ήρεμο και σε έλεγχο. Ένας άνθρωπος με υψηλό επίπεδο νευρωτισμού, από την άλλη πλευρά, θα είναι πολύ πιο ασταθής και επιρρεπής στην υπερβολική αντίδραση στα εξωτερικά ερεθίσματα και μπορεί να εκφράζει έντονα ανησυχία, θυμό ή φόβο. Είναι υπερβολικά συναισθηματικός και δυσκολεύεται να ηρεμήσει μόλις αναστατωθεί. Τα νευρολογικά άτομα έχουν ένα ANS που ανταποκρίνεται γρήγορα στο άγχος.
- **Ψυχωτισμός/Κανονικότητα (Psychoticism/normality):** Ο Eysenck [56] προσέθεσε αργότερα ένα τρίτο χαρακτηριστικό / διάσταση - Ψυχωτισμός - π.χ. έλλειψη ενσυναίσθησης, σκληρός, μοναχικός, επιθετικός και ενοχλητικός. Αυτό έχει σχέση με τα υψηλά επίπεδα τεστοστερόνης. Όσο υψηλότερη είναι η τεστοστερόνη, τόσο υψηλότερο είναι το επίπεδο ψυχωτισμού, με χαμηλά επίπεδα που σχετίζονται με πιο φυσιολογική ισορροπημένη συμπεριφορά.

Σύμφωνα με τον Eysenck, οι δύο διαστάσεις του νευρωτισμού (σταθερώ έναντι ασταθούς) και της εξωστρέφειας/εσωστρέφειας συνδυάζονται για να σχηματίσουν μια ποικιλία χαρακτηριστικών προσωπικότητας.

Η θεωρία του Μαθήματος 16PF της Cattell

Ο Cattell [57] διαφώνησε με την άποψη του Eysenck ότι η προσωπικότητα μπορεί να γίνει κατανοητή εξετάζοντας μόνο δύο ή τρεις διαστάσεις συμπεριφοράς. Αντ' αυτού, υποστήριξε ότι είναι απαραίτητο να εξετάσουμε ένα πολύ μεγαλύτερο αριθμό χαρακτηριστικών για να αποκτήσουμε μια πλήρη εικόνα της προσωπικότητας κάποιου. Ενώ ο Eysenck θεμελίωσε τη θεωρία του με βάση τις



Σχήμα 2.2: Η κατανομή των χαρακτηριστικών προσωπικότητας του Eysenck

απαντήσεις των νοσοκομειακών στρατιωτών, ο Cattell συγκέντρωσε δεδομένα από μια σειρά ανθρώπων μέσω τριών διαφορετικών πηγών δεδομένων.

- L-data: αυτό είναι δεδομένα από τα αρχεία ζωής, όπως βαθμοί σχολείου, απουσία από την εργασία κ.λπ.
- Q-data: αυτό ήταν ένα ερωτηματολόγιο σχεδιασμένο να αξιολογεί την προσωπικότητα ενός ατόμου (γνωστό ως 16PF).
- T-data: αυτά είναι δεδομένα από αντικειμενικές δοκιμές που έχουν σχεδιαστεί για να «χτυπάνε» σε μια κατασκευή προσωπικότητας.

Ο Cattell ανέλυσε τα δεδομένα T και τα δεδομένα Q χρησιμοποιώντας μια μαθηματική τεχνική που ονομάζεται ανάλυση παράγοντα για να εξετάσει ποιοι τύποι συμπεριφοράς τείνουν να ομαδοποιούνται μεταξύ τους στους ίδιους ανθρώπους. Προσδιόρισε 16 χαρακτηριστικά γνωρίσματα / παράγοντες κοινά για όλους τους ανθρώπους. Ο Cattell προέβη σε διάκριση μεταξύ των χαρακτηριστικών προέλευσης και επιφάνειας. Τα επιφανειακά γνωρίσματα είναι πολύ προφανή και μπορούν εύκολα να ταυτοποιηθούν από άλλους ανθρώπους, ενώ τα χαρακτηριστικά πηγής είναι λιγότερο ορατά σε άλλους ανθρώπους και φαίνονται να αποτελούν τη βάση πολλών διαφορετικών παραμέτρων συμπεριφοράς. Ο Cattell θεώρησε ότι τα χαρακτηριστικά πηγής είναι πιο σημαντικά στην περιγραφή της προσωπικότητας από τα χαρακτηριστικά της επιφάνειας.

Θεωρία των Χαρακτηριστικών του Allport

Η θεωρία της προσωπικότητας του Allport τονίζει τη μοναδικότητα του ατόμου και τις εσωτερικές γνωστικές και κινητικές διαδικασίες που επηρεάζουν τη συμπεριφορά. Για παράδειγμα, νοημοσύνη, ιδιοσυγκρασία, συνήθειες, δεξιότητες, συμπεριφορές και χαρακτηριστικά. Ο Allport [58] πιστεύει ότι η προσωπικότητα είναι βιολογικά καθορισμένη κατά τη γέννηση και διαμορφώνεται από την περιβαλλοντική εμπειρία ενός ατόμου. Σύμφωνα με την Allport, τα χαρακτηριστικά της προσωπικότητας μπορούν να ταξινομηθούν σε τρία γενικά επίπεδα:

- **Πρωτεύοντα Χαρακτηριστικά (Cardinal Traits):** κυριαρχούν στο σύνολο της ζωής ενός ατόμου.

Factor	Low Score	High Score
Warmth	cold, selfish	supportive, comforting
Intellect	Instinctive, unstable	cerebral, analytical
Emotional Stability	Irritable, moody	level headed, calm
Aggressiveness	Modest, docile	controlling, tough
Liveliness	somber, restrained	wild, fun loving
Dutifulness	untraditional, rebellious	conformity, traditional
Social Assertiveness	shy, withdrawn	uninhibited, bold
Sensitivity	coarse, tough	touchy, soft
Paranoia	trusting, easy going	wary, suspicious
Abstractness	practical, regular	strange, imaginative
Introversion	open, friendly	private, quiet
Anxiety	confident, self-assured	fearful, self-doubting
Open-mindedness	close-minded, set-in-ways	curious, self-exploratory
Independence	outgoing, social	loner, crave solitude
Perfectionism	Disorganized, messy	orderly, thorough
Tension	relaxed, cool	stressed, unsatisfied

Σχήμα 2.3: Τα χαρακτηριστικά της προσωπικότητας του Cattell

- **Κεντρικά Χαρακτηριστικά (Central Traits):** είναι γενικά χαρακτηριστικά που χρησιμοποιούνται για να περιγράψουν ένα άλλο άτομο (π.χ. ευγενικός, ειλικρινής και ευτυχισμένος).
- **Δευτερεύοντα Χαρακτηριστικά (Secondary Traits):** είναι εκείνα τα χαρακτηριστικά που εξέρχονται μόνο κάτω από ορισμένες καταστάσεις.

Το Big Five: Μοντέλο Πέντε Παραγόντων

Ως αποτέλεσμα μιας διεξοδικής έρευνας σχετικά με τις θεωρίες της προσωπικότητας του Cattell και του Eysenck, η θεωρία των Big Five διατυπώθηκε στο [59, 60, 61]. Σε αυτό το μοντέλο αναφέρεται ότι υπάρχουν 5 βασικά χαρακτηριστικά που συνεργάζονται για να σχηματίσουν μια ενιαία προσωπικότητα. Αυτά περιλαμβάνουν:

- **Εξωστρέφεια (Extraversion)** (εξωστρεφής/ενεργητικός vs. μοναχικός/επιφυλακτικός): τάση να είναι ενεργός, κοινωνικός, προσανατολισμένος στον άνθρωπο, ομιλητικός, αισιόδοξος, ενσυναισθητικός.
- **Ανοιχτός σε Εμπειρίες (Openness to Experience)** (εφευρετικός/περίεργος vs. συνεπής/προσεκτικός): τάση να είναι ευφάνταστος, περίεργος, δημιουργικός και μπορεί να έχει μη συμβατικές πεποιθήσεις και αξίες.
- **Τερπνότητα (Agreeableness)** (φιλικός/συμπνετικός vs. προκλητικός/αποκομμένος): τάση να είναι καλοπροαίρετος, ευγενικός, εξυπηρετικός, αλτρουιστικός και εμπιστευτικός.
- **Ευσυνειδησία (Conscientiousness)** (αποδοτικός/οργανωτικός vs. εύκολος/ανέμελος): τάση να είναι εργατικός, αξιόπιστος, φιλόδοξος, ακριβής και αυτοκατευθυνόμενος.
- **Νευρωτισμός (Neuroticism)** (ευαίσθητος/νευρικός vs. ασφαλής/σίγουρος): τάση να γίνεται συναισθηματικά ασταθής και μπορεί ακόμη και να αναπτύξει ψυχολογική δυσφορία. Αυτό το χαρακτηριστικό προσωπικότητας αναφέρεται επίσης ως *Ψυχολογική σταθερότητα (Emotional Stability)*.

Το μοντέλο Big Five (που ονομάζεται επίσης μοντέλο OCEAN) είναι το πιο ευρέως αποδεκτό μοντέλο προσωπικότητας στην ψυχολογία σήμερα.

2.2.2 Καταλληλότητα του Big Five Μοντέλου

Σύμφωνα με τον Digman [60], πολλές ψυχολογικές μελέτες κατέληξαν ανεξάρτητα στο συμπέρασμα ότι πέντε χαρακτηριστικά γνωρίσματα επαρκούν για την περιγραφή της ανθρώπινης προσωπικότητας. “Εάν χρησιμοποιείται μεγάλος αριθμός κλιμάκων αξιολόγησης και εάν το εύρος των αριθμών αυτών είναι πολύ ευρύ, η περιγραφή της προσωπικότητας αντιστοιχεί σχεδόν εξ ολοκλήρου σε πέντε ισχυρούς παράγοντες”, όπως υποστηρίζεται από το [1]. Ωστόσο, δεν υπάρχει πλήρης συμφωνία για την έννοια του κάθε χαρακτηριστικού, δεδομένου ότι θεωρούνται αόριστες. Για παράδειγμα υπάρχει κάποια διαφωνία όσον αφορά το χαρακτηριστικό Open to experience. Κάποιοι προτιμούν τη συμβολική λέξη intellect αντί να λένε Openness to experience.

Επιπλέον, το ίδιο μοντέλο χρησιμοποιείται σε διάφορες γλώσσες και πολιτισμούς, όπως τον κινεζικό [62] και τον ινδικό [63]. Σύμφωνα με τους ερευνητές, όπως [64, 65], το Openness to experience χαρακτηριστικό δεν υποστηρίζεται ιδιαίτερα στους ασιατικούς πολιτισμούς όπως στην Κίνα και την Ιαπωνία. Σε αυτήν την περίπτωση ένας διαφορετικός πέμπτος παράγοντας προσδιορίζεται μερικές φορές προκειμένου να υποκατασταθεί αυτός που χρησιμοποιείται από τον αγγλικό πολιτισμό. Επίσης, η σχέση μεταξύ της γλώσσας και της προσωπικότητας έχει μελετηθεί (βλ. [66] για μια έρευνα), αν και οι περισσότερες έρευνες επικεντρώνονται στην αγγλική γλώσσα.

Παρά τα προβλήματα και τις επικρίσεις, το μοντέλο Big Five θεωρείται κατάλληλο για την περιγραφή της προσωπικότητας του ατόμου. Αυτό το μοντέλο έχει καθιερωθεί, δεδομένου ότι είναι βολικό για υπολογιστικές και μαθησιακές προσεγγίσεις και, επιπρόσθετα, είναι αρκετά γενικό ώστε να εφαρμόζεται σε πολλές γλώσσες και πολιτισμούς, με εξαίρεση το χαρακτηριστικό του Openness to experience που δεν μπορεί να χρησιμοποιηθεί άμεσα στους ανατολικούς πολιτισμούς.

Αν και το Big Five μοντέλο περιγράφει επαρκώς τις επίμονες αντιδράσεις της ανθρώπινης συμπεριφοράς σε ευρείες κατηγορίες περιβαλλοντικών ερεθισμάτων, χαρακτηρίζοντας με τον τρόπο αυτό ένα μοναδικό άτομο [67], η προσωπικότητα αλλάζει με την πάροδο του χρόνου και προσαρμόζεται στο περιβάλλον. Για παράδειγμα, όπως ο DeYoung [68] επεσήμανε, οι στόχοι, τα κίνητρα και το πλαίσιο επηρεάζουν τον τρόπο που οι άνθρωποι εμφανίζουν την προσωπικότητά τους. Οι άνθρωποι μπορεί επίσης να προσποιούνται ότι έχουν διαφορετικά χαρακτηριστικά προσωπικότητας ανάλογα με τις προθέσεις τους στο πλαίσιο μιας συγκεκριμένης κατάστασης. Η γενική θέση των ψυχολόγων σχετικά με αυτά τα προβλήματα (αυτό είναι επίσης στη βάση της εργασίας του Adelman [69]) είναι ότι τα άτομα έχουν κάποια μάλλον σταθερά χαρακτηριστικά προσωπικότητας πυρήνα και άλλα πιο μεταβλητά περιφερειακά χαρακτηριστικά. Ωστόσο, ένας τέτοιος διαχωρισμός δεν έχει εφαρμοστεί στο Big Five μοντέλο και ως εκ τούτου, τα χαρακτηριστικά επηρεάζονται από εξωτερικούς παράγοντες, όπως το πλαίσιο μιας δεδομένης κατάστασης.

2.3 Εκμάθηση Μηχανών για Επεξεργασία Φυσικής Γλώσσας

2.3.1 Επεξεργασία Φυσικής Γλώσσας

Η επεξεργασία φυσικής γλώσσας (Natural Language Processing - NLP) είναι ένας τομέας της επιστήμης των υπολογιστών, της τεχνητής νοημοσύνης (επίσης αποκαλούμενης μηχανικής μάθησης) και της γλωσσολογίας που ασχολείται με τις αλληλεπιδράσεις μεταξύ ηλεκτρονικών υπολογιστών και ανθρώπων (φυσικών) γλωσσών. Είναι η διαδικασία ενός υπολογιστή που εξάγει σημαντικές πληροφορίες από την εισαγωγή φυσικής γλώσσας ή / και παράγει έξοδο φυσικής γλώσσας. Είναι η ανάλυση

της ανθρώπινης γλώσσας με βάση τη σημασιολογία και τις διάφορες τεχνικές ανάλυσης [70]. Ο στόχος του NLP είναι να προσδιορίσει τον υπολογιστικό μηχανισμό που απαιτείται για να παρουσιάζει διάφορες μορφές γλωσσικής συμπεριφοράς (δηλ. Επιστημονικό στόχο). Επίσης, σχεδιάζει, εφαρμόζει και δοκιμάζει συστήματα που επεξεργάζονται φυσικές γλώσσες για πρακτικές εφαρμογές (δηλ. Μηχανικός Στόχος).

Το NLP είναι μια πειθαρχία μεταξύ της γλωσσολογίας και της επιστήμης των υπολογιστών που ασχολείται με τις υπολογιστικές πτυχές της διδασκαλίας της ανθρώπινης γλώσσας. Το κύριο καθήκον της είναι να κατασκευάσει προγράμματα για να επεξεργαστεί λέξεις και κείμενα σε φυσική γλώσσα. Οι κύριες πτυχές των NLP είναι:

- Απόκτηση Πληροφορίας (Information Retrieval - IR): Αφορά την αποθήκευση, την αναζήτηση και την ανάκτηση πληροφοριών από έγγραφα κειμένου. Είναι ένας τομέας της επιστήμης των υπολογιστών πιο κοντά στις βάσεις δεδομένων και βασίζεται σε μερικές από τις μεθόδους NLP.
- Μηχανική Μετάφραση (Machine Translation - MT): Σχετίζεται με την αυτόματη μετάφραση από τη μια ανθρώπινη γλώσσα στην άλλη [71].
- Γλωσσολογική Ανάλυση (Language Analysis): Αφορά την ανάλυση μιας πρότασης εισόδου για την κατασκευή συντακτικού δένδρου και την περαιτέρω ανάλυση του συναισθήματος που γίνεται για την εύρεση σημαντικών λέξεων σε μια πρόταση.

Γλωσσολογία είναι η επιστήμη της γλώσσας. Η μελέτη περιλαμβάνει τους ήχους (φωνολογία), τον σχηματισμό του λόγου (μορφολογία), τη δομή της φράσης (σύνταξη), τη σημασία (σημασιολογία) και την κατανόηση (πραγματισμός).

- Φωνολογική Ανάλυση (Phonological Analysis): συσχετίζει τους ήχους με τις λέξεις που αναγνωρίζουμε. Το φωνή είναι η μικρότερη μονάδα ήχου και τα τηλέφωνα συγκεντρώνονται σε ήχους λέξεων.
- Μορφολογική Ανάλυση (Morphological Analysis): Η μορφολογία είναι μια υπο-πειθαρχία της γλωσσολογίας που μελετά τη δομή των λέξεων. Αφορά την απόκτηση νέων λέξεων από τις υπάρχουσες. Στο NLP, οι λέξεις είναι γνωστές ως λεξικά και ένα σύνολο λέξεων αποτελεί ένα λεξικό. Το λεξικό είναι μια ενότητα που λέει ποιες λέξεις υπάρχουν και ποιες ιδιότητες έχουν.
- Συντακτική Ανάλυση (Syntactic Analysis): είναι η ανάλυση των λέξεων σε μια πρόταση για να γνωρίζουμε τη γραμματική δομή μιας φράσης και αυτές οι λέξεις μετατρέπονται σε δομές που δείχνουν πώς οι λέξεις σχετίζονται μεταξύ τους.
- Σημασιολογική Ανάλυση (Semantic Analysis): Αφορά την έννοια της γλώσσας. Το πρώτο βήμα στο σύστημα σημασιολογικής επεξεργασίας είναι να αναζητήσουμε τις μεμονωμένες λέξεις σε ένα λεξικό (ή λεξικό) και να εξαγάγουμε τις έννοιές τους [72].
- Πραγματολογική Ανάλυση (Pragmatic Analysis): Στόχος είναι να ερμηνεύσει αυτό που ειπώθηκε σε αυτό που πραγματικά σήμαινε. Πρόκειται για το πώς χρησιμοποιούνται οι προτάσεις σε διαφορετικές καταστάσεις και πώς η χρήση επηρεάζει την ερμηνεία της φράσης.

Οι περισσότερες εφαρμογές NLP, όπως η εξαγωγή πληροφοριών, η μηχανική μετάφραση, η ανάλυση συναισθημάτων και η απάντηση σε ερωτήσεις, απαιτούν τόσο συντακτική όσο και σημασιολογική ανάλυση σε διάφορα επίπεδα.

- IR & Web Search: είναι η επιστήμη της αναζήτησης εγγράφων, πληροφοριών εντός εγγράφων και μεταδεδομένων σχετικά με τα έγγραφα, καθώς και της αναζήτησης βάσεων δεδομένων και του Παγκόσμιου Ιστού.

- Εξαγωγή Πληροφοριών (Information Extraction - IE): είναι ένας τύπος ανάκτησης πληροφοριών, ο στόχος του οποίου είναι η αυτόματη εξαγωγή δομημένων πληροφοριών, δηλαδή κατηγοριοποιημένων και συμφραζομένων και σημασιολογικά καλά καθορισμένων δεδομένων από συγκεκριμένο τομέα, από μη δομημένα μηχανικά αναγνώσιμα έγγραφα.
- Ερώτηση-Απάντηση (Question Answering - QA): είναι η απάντηση από τα έγγραφα στην εξεχθείσα ή παραγόμενη απάντηση.
- Περίληψη Κειμένου (Text Summarization): Η περίληψη κειμένου είναι η διαδικασία απόσταξης των σημαντικότερων πληροφοριών από μια πηγή, προκειμένου να παραχθεί μια συνοπτική έκδοση.
- MT: είναι η χρήση λογισμικού ηλεκτρονικών υπολογιστών για τη μετάφραση κειμένου ή ομιλίας από τη μια φυσική γλώσσα στην άλλη.
- Αναγνώριση Φωνής και Σύνθεση (Speech Recognition & Synthesis): είναι η εξαγωγή της κειμενικής αναπαράστασης μιας προφορικής ομιλίας
- Κατανόηση και Παραγωγή Φυσικής Γλώσσας (Natural Language Understanding and Generation - NLU, NLG): Το NLG σύστημα είναι σαν ένας μεταφραστής που μετατρέπει μια εκπροσώπηση μέσω υπολογιστή σε μια φυσική αντιπροσώπευση γλώσσας.
- Συζήτηση Ανθρώπου-Υπολογιστή (Human-Computer Conversation): είναι ο διάλογος μεταξύ ανθρώπου και υπολογιστή που χρησιμοποιεί τη φυσική γλώσσα.
- Παραγωγή Κειμένου (Text Generation): Μια μέθοδος για τη δημιουργία προτάσεων από "λέξεις-κλειδιά" ή "λέξεις-κλειδιά".
- Αναγνώριση Χειρόγραφων Εγγράφων (Hand writing recognition): Δυνατότητα ενός υπολογιστή να λαμβάνει και να ερμηνεύει έξυπνες χειρόγραφες εισόδους από πηγές όπως έγγραφα, φωτογραφίες, οθόνες αφής και άλλες συσκευές.

2.3.2 Ορισμός της Μάθησης Μηχανών

Η Μηχανική Μάθηση (Machine Learning - ML) είναι υποπεδίο Τεχνητής Νοημοσύνης (Artificial Intelligence - AI). Ο στόχος του ML είναι γενικά η κατανόηση της δομής των δεδομένων και η προσαρμογή των δεδομένων σε μοντέλα που μπορούν να κατανοηθούν και να χρησιμοποιηθούν από τους ανθρώπους. Παρόλο που το ML είναι ένα πεδίο στην επιστήμη των υπολογιστών, διαφέρει από τις παραδοσιακές υπολογιστικές προσεγγίσεις. Στην παραδοσιακή υπολογιστική, οι αλγόριθμοι είναι σύνολα ρητά προγραμματισμένων οδηγιών που χρησιμοποιούνται από τους υπολογιστές για τον υπολογισμό ή την επίλυση προβλημάτων. Οι αλγόριθμοι ML επιτρέπουν στους υπολογιστές να εκπαιδεύονται στις εισόδους δεδομένων και να χρησιμοποιούν στατιστική ανάλυση για να εξάγουν τιμές που εμπίπτουν σε ένα συγκεκριμένο εύρος. Εξαιτίας αυτού, το ML διευκολύνει τους υπολογιστές στη δημιουργία μοντέλων από δείγματα δεδομένων, προκειμένου να αυτοματοποιήσει τις διαδικασίες λήψης αποφάσεων με βάση τις εισροές δεδομένων.

Στο ML, τα καθήκοντα ταξινομούνται γενικά σε ευρείες κατηγορίες. Αυτές οι κατηγορίες βασίζονται στον τρόπο με τον οποίο λαμβάνεται η μάθηση ή στον τρόπο με τον οποίο δίδεται ανάδραση στην εκμάθηση στο αναπτυγμένο σύστημα. Δύο από τις πιο ευρέως υιοθετημένες μεθόδους ML είναι η **επιβλεπόμενη μάθηση (supervised learning)** η οποία εκπαιδεύει τους αλγόριθμους που βασίζονται σε δεδομένα εισόδου και εξόδου παραδείγματος που επισημαίνονται από τον άνθρωπο και η **μάθηση χωρίς επίβλεψη (unsupervised learning)** που παρέχει τον αλγόριθμο χωρίς επισημασμένα δεδομένα προκειμένου να του επιτρέψει να βρει δομή μέσα στα δεδομένα εισόδου του.

2.3.3 Ορισμός Διαφορετικών Μεθόδων ML

Επιβλεπόμενη Μάθηση (Supervised Learning)

Στην επιβλεπόμενη μάθηση, υπάρχουν μεταβλητές εισόδου (x) και μεταβλητή εξόδου (Y) και ο στόχος είναι η εκμάθηση της συνάρτησης χαρτογράφησης από την είσοδο στην έξοδο μέσω ενός αλγορίθμου.

$$Y = f(X) \quad (2.1)$$

Ο στόχος είναι να προσεγγίσουμε τη λειτουργία χαρτογράφησης τόσο καλά όσο όταν εισάγονται στο μοντέλο νέα δεδομένα εισόδου (x), οι αντίστοιχες μεταβλητές εξόδου (Y) θα μπορούσαν να προβλεφθούν με επιτυχία.

Ονομάζεται επιβλεπόμενη μάθηση επειδή η διαδικασία ενός αλγορίθμου που μαθαίνει από το σύνολο δεδομένων κατάρτισης μπορεί να θεωρηθεί ως δάσκαλος που επιβλέπει τη διαδικασία μάθησης. Γνωρίζουμε τις σωστές απαντήσεις, ο αλγόριθμος κάνει επανειλημμένα προβλέψεις για τα δεδομένα εκπαίδευσης και διορθώνεται από τον δάσκαλο. Η μάθηση σταματά όταν ο αλγόριθμος επιτυγχάνει ένα αποδεκτό επίπεδο απόδοσης.

Τα επιβλεπόμενα προβλήματα μάθησης μπορούν να ομαδοποιηθούν περαιτέρω σε προβλήματα **παλινδρόμηση (regression)** και **ταξινόμηση (classification)**.

- Η παλινδρόμηση είναι το πρόβλημα της εκτίμησης ή της πρόβλεψης μιας συνεχούς ποσότητας. Ποια θα είναι η αξία του S&P 500 ένα μήνα από σήμερα; Πόσο ψηλά θα είναι ένα παιδί ως ενήλικας; Πόσοι από τους πελάτες μας θα φύγουν για έναν ανταγωνιστή φέτος; Αυτά είναι παραδείγματα ερωτήσεων που θα πέσουν κάτω από την οπισθοδρόμηση. Για την επίλυση αυτών των προβλημάτων σε ένα εποπτευόμενο πλαίσιο μηχανικής μάθησης, θα συγκεντρώναμε προηγούμενα ζεύγη εισόδου / εξόδου "σωστής απάντησης" που αντιμετωπίζουν το ίδιο πρόβλημα. Για τις εισροές, θα εντοπίζαμε χαρακτηριστικά που πιστεύουμε ότι θα ήταν προβλέψιμα για τα αποτελέσματα που θέλουμε να προβλέψουμε.
- Η ταξινόμηση αφορά την ανάθεση παρατηρήσεων σε διακριτές κατηγορίες, παρά την εκτίμηση των συνεχών ποσοτήτων. Στην απλούστερη περίπτωση, υπάρχουν δύο πιθανές κατηγορίες. αυτή η περίπτωση είναι γνωστή ως δυαδική ταξινόμηση. Πολλά σημαντικά ερωτήματα μπορούν να πλασιασθούν από την άποψη της δυαδικής ταξινόμησης. Θα μας αφήσει ένας συγκεκριμένος πελάτης για έναν ανταγωνιστή; Ένας συγκεκριμένος ασθενής έχει καρκίνο; Μία δεδομένη εικόνα περιέχει ένα hot dog; Οι αλγόριθμοι για την εκτέλεση της δυαδικής ταξινόμησης είναι ιδιαίτερα σημαντικοί επειδή πολλοί από τους αλγόριθμους για την εκτέλεση του γενικότερου είδους ταξινόμησης όπου υπάρχουν αυθαίρετες ετικέτες είναι απλά μια δέσμη δυαδικών ταξινομητών που δουλεύουν μαζί.

Μάθηση Χωρίς Επίβλεψη (Unsupervised Learning)

Από την άλλη πλευρά, υπάρχει μια εντελώς διαφορετική τάξη καθηκόντων που αναφέρεται ως μάθηση χωρίς επίβλεψη. Οι επιβλεπόμενες μαθησιακές εργασίες βρίσκουν μοτίβα όπου έχουμε ένα σύνολο δεδομένων "σωστών απαντήσεων" για να μάθουμε από. Οι μαθησιακές εργασίες χωρίς επίβλεψη βρίσκουν μοτίβα όπου δεν το κάνουμε. Αυτό μπορεί να οφείλεται στο γεγονός ότι οι "σωστές απαντήσεις" δεν είναι παρατηρήσιμες ή δεν μπορούν να ληφθούν, ή ίσως για ένα δεδομένο πρόβλημα, δεν υπάρχει μία "σωστή απάντηση".

Μια μεγάλη υποκατηγορία εργασιών χωρίς επίβλεψη είναι το πρόβλημα της ομαδοποίησης. Η ομαδοποίηση αναφέρεται στην ομαδοποίηση των παρατηρήσεων μαζί με τέτοιο τρόπο ώστε τα μέλη

μιας κοινής ομάδας να είναι παρόμοια μεταξύ τους και διαφορετικά από τα μέλη άλλων ομάδων. Μια κοινή εφαρμογή εδώ είναι στο μάρκετινγκ, όπου θέλουμε να εντοπίσουμε τμήματα πελατών ή προοπτικές με παρόμοιες προτιμήσεις ή αγοραστικές συνήθειες. Μία σημαντική πρόκληση στη συγκέντρωση είναι ότι συχνά είναι δύσκολο ή αδύνατο να γνωρίζουμε πόσα συμπλέγματα θα πρέπει να υπάρχουν ή πώς πρέπει να φαίνονται τα σμήνη.

Μια πολύ ενδιαφέρουσα κατηγορία προβλημάτων χωρίς επίβλεψη είναι η γενετική μοντελοποίηση. Τα γενετικά μοντέλα είναι μοντέλα που μιμούνται τη διαδικασία που παράγει τα δεδομένα εκπαίδευσης. Ένα καλό γενετικό μοντέλο θα ήταν σε θέση να παράγει νέα δεδομένα που μοιάζουν με τα δεδομένα κατάρτισης με κάποια έννοια. Αυτός ο τύπος μάθησης δεν εποπτεύεται επειδή η διαδικασία που παράγει τα δεδομένα δεν είναι άμεσα παρατηρήσιμη - μόνο τα ίδια τα δεδομένα είναι παρατηρήσιμα.

Ενισχυμένη Μάθηση (Reinforcement Learning)

Ένας νεότερος τύπος μαθησιακού προβλήματος που έχει αποκτήσει πρόσφατα μεγάλη έλξη ονομάζεται ενισχυμένη μάθηση. Στην ενισχυμένη μάθηση, δεν παρέχουμε στο μηχανήμα παραδείγματα σωστών ζευγών εισόδου / εξόδου, αλλά παρέχουμε μια μέθοδο για τη μηχανή να ποσοτικοποιήσει την απόδοσή της με τη μορφή ενός σήματος ανταμοιβής. Οι μέθοδοι ενισχυμένης μάθησης μοιάζουν με τον τρόπο με τον οποίο μαθαίνουν οι άνθρωποι και τα ζώα: το μηχανήμα δοκιμάζει ένα σωρό διαφορετικά πράγματα και ανταμείβεται όταν κάνει κάτι καλά.

Η ενισχυμένη μάθηση είναι χρήσιμη σε περιπτώσεις όπου ο χώρος της λύσης είναι τεράστιος ή άπειρος και ισχύει συνήθως σε περιπτώσεις όπου η μηχανή μπορεί να θεωρηθεί ως παράγοντας που αλληλεπιδρά με το περιβάλλον της. Μια από τις πρώτες μεγάλες επιτυχίες αυτού του τύπου μοντέλου ήταν από μια μικρή ομάδα που εκπαιδεύτηκε ένα μοντέλο μάθησης οπλισμού για να παίξει βιντεοπαιχνίδια Atari χρησιμοποιώντας μόνο την έξοδο pixel από το παιχνίδι ως είσοδο [73]. Το μοντέλο τελικά ήταν σε θέση να ξεπεράσει τους ανθρώπινους παίκτες σε τρία παιχνίδια.

2.3.4 Χαρακτηριστικά

Η μηχανική των χαρακτηριστικών (feature engineering) είναι η διαδικασία μετατροπής των ακατέργαστων δεδομένων σε χαρακτηριστικά που αντιπροσωπεύουν καλύτερα το υποκείμενο πρόβλημα στα μοντέλα πρόβλεψης, με αποτέλεσμα τη βελτιωμένη ακρίβεια μοντέλων σε δεδομένα που δεν βλέπουν. Η μηχανική των χαρακτηριστικών περιλαμβάνει διάφορα βήματα όπως παρουσιάζονται στη συνέχεια.

Προεπεξεργασία

Η προεπεξεργασία αναλύει τις απόψεις από τη συντακτική άποψη και η αρχική σύνταξη της φράσης δεν διαταράσσεται. Οι διάφορες τεχνικές, όπως η προσθήκη ετικετών με το μέρος των ομιλιών (Part-Of-Speech (POS) tagging), η απομάκρυνση λέξεων και η διακοπή λέξεων, εφαρμόζονται σε ένα σύνολο δεδομένων για τη μείωση του θορύβου και διευκολύνουν την εξαγωγή χαρακτηριστικών.

POS tagging. Το POS tagging είναι μια γλωσσική τεχνική που χρησιμοποιείται από το 1960 [74, 75, 76, 77, 78, 79, 80, 81]. Το POS tagging [82] εκχωρεί μια ετικέτα σε κάθε λέξη σε ένα κείμενο και ταξινομεί μια λέξη σε μια συγκεκριμένη μορφολογική κατηγορία όπως το ουσιαστικό, το ρήμα, το επίθετο κτλ. Το POS tagging είναι αποτελεσματικό για την εξαγωγή ρητών χαρακτηριστικών από την άποψη της ακρίβειας που επιτεύχθηκαν στο [74, 77, 79, 83]. Τα κρυμμένα μοντέλα Markov (Hidden Markov Models) χρησιμοποιούνται ευρέως για την ανάπτυξη ετικετών POS λόγω ακρίβειας σε σύγκριση με άλλες τεχνικές όπως η βασισμένη σε κανόνες, η στατιστική και η μηχανική

μάθηση [84]. Διαφορετικά εργαλεία POS tagging της αγγλικής γλώσσας, όπως ο γλωσσικός αναλυτής επεξεργαστών NL, το POS marker Stanford, το Gate ANNIE POS Tagger και το Markers POS tagger χρησιμοποιούνται για το σκοπό αυτό. Το πακέτο εργαλείων NLTK [85] με βάση την γλώσσα προγραμματισμού Python διαθέτει μια πλούσια συλλογή όλων των ενοτήτων, συμπεριλαμβανομένων των POS, που χρειάζονται οι ερευνητές του NLP.

Stemming and Lemmatization. Τα Stemming και Lemmatization είναι δύο βασικές μορφολογικές διεργασίες της διαδικασίας προεπεξεργασίας κατά τη διάρκεια της εξαγωγής χαρακτηριστικών [74, 76, 77, 78, 79, 81, 86, 81]. Η διαδικασία που προέρχεται από το stemming μετατρέπει όλες τις λέξεις που εμφανίζονται στο κείμενο στην αντίστοιχη ρίζα τους [82]. Για παράδειγμα, οι λέξεις "αυτόματη" και "αυτοματοποίηση" μετατρέπονται και οι δύο σε "αυτοματ". Το stemming έχει καλύτερη και ταχύτερη απόδοση σε εφαρμογές όπου η ακρίβεια δεν είναι σημαντικό ζήτημα. Ο πρώτος stemmer δημοσιεύθηκε από τον Julie Beth Lovins το 1968. Ο Martin Porter σχεδίασε και δημοσίευσε τον stemmer του τον Ιούλιο του 1980. Οι Porter και Lancaster είναι οι αλγόριθμοι που προέρχονται από την Python NLTK. Το RSLP Stemmer1, το ISRI Stemmer2 και το SnowballStemmer3 είναι μη αγγλικά plugins. Το λήμμα μιας λέξης περιλαμβάνει τη βασική του μορφή [87]. Για παράδειγμα, οι λέξεις "παίζει", "έπαιξε" και "παίζουν" έχουν "παίζει" ως το λήμμα τους. Το lemmatization συγκεντρώνει διάφορες μορφές λέξεων σε μια μόνο. Το stemming αφαιρεί τις λεγόμενες λέξεις, ενώ το lemmatization αντικαθιστά λέξεις με τη μορφή βάσης τους. Σε σχέση με το stemming, το lemmatization θεωρείται πιο ακριβής διαδικασία. Σε αντίθεση με την προέλευση, το lemmatization απαιτεί πρόσθετη υποστήριξη λεξικών για αναζήτηση και ευρετηρίαση, γεγονός που αυξάνει την ακρίβειά του στις εφαρμογές εξαγωγής χαρακτηριστικών, αλλά υποβαθμίζει την ταχύτητα του lemmatizer. Το Word Net Lemmatizer με τη βάση δεδομένων του Word Net χρησιμοποιείται για την αναζήτηση των λημμάτων.

Stop Word Removal. Η έννοια των stop words εισήχθη για πρώτη φορά από τον Hans Luhn, H.P [88]. Οι stop words είναι κοινές λέξεις υψηλής συχνότητας όπως "το", "του", "και", "ένα". Υπάρχουν διαφορετικές διαθέσιμες μέθοδοι για την κατάργηση των stop words [82] ώστε να ενισχύσουν τελικά την απόδοση του αλγόριθμου εξαγωγής χαρακτηριστικών [74, 82, 77]. Η αφαίρεση των stop words μειώνει τη διαστασιολόγηση των συνόλων δεδομένων και έτσι οι λέξεις-"κλειδιά" που απομένουν στο κείμενο μπορούν να αναγνωριστούν ευκολότερα από τις τεχνικές αυτόματης εξαγωγής χαρακτηριστικών. Οι λέξεις που πρέπει να αφαιρεθούν λαμβάνονται από μια ευρέως διαθέσιμη λίστα από stop words. Ο Savoy [89] είχε δώσει μια τεράστια συλλογή από λίστες με stop words. Σε απλούστερο επίπεδο, οι stop words επαναλαμβάνονται στην επιλεγμένη λίστα λέξεων και αφαιρούνται από το κείμενο.

Επιλογή Χαρακτηριστικών

Τα τελευταία τριάντα χρόνια, οι διαστάσεις των δεδομένων που εμπλέκονται στο ML και σε προβλήματα data mining έχουν αυξηθεί σημαντικά. Τα δεδομένα με εξαιρετικά υψηλό αριθμό διαστάσεων παρουσιάζουν σοβαρές προκλήσεις στις υπάρχουσες μεθόδους μάθησης [90], εισάγοντας το πρόβλημα του "curse of dimensionality" [91]. Με την παρουσία ενός μεγάλου αριθμού χαρακτηριστικών, ένα μαθησιακό μοντέλο τείνει να παρουσιάσει το πρόβλημα του "overfitting", με αποτέλεσμα να εκφυλίζονται οι επιδόσεις τους. Για να αντιμετωπίσουμε το πρόβλημα του "curse of dimensionality", έχουν μελετηθεί τεχνικές μείωσης των διαστάσεων, ένας σημαντικός υποκλάδος στον τομέα έρευνας ML και data mining. Η επιλογή χαρακτηριστικών είναι ευρέως αποδεκτή τεχνική για τη μείωση των διαστάσεων. Επιδίδει να επιλέξει ένα μικρό υποσύνολο των σχετικών χαρακτηριστικών από τα αρχικά, σύμφωνα με κάποια σχετικά κριτήρια αξιολόγησης, οδηγώντας συνήθως σε καλύτερες μαθησιακές επιδόσεις (π.χ. ανώτερη μάθηση ακρίβεια για την ταξινόμηση), χαμηλότερο υπολογιστικό κόστος και καλύτερη ερμηνεία μοντέλου.

Ανάλογα με το αν το σετ δεδομένων είναι ετικετοποιημένο ή όχι, οι αλγόριθμοι επιλογής χαρακτηριστικών μπορούν να κατηγοριοποιηθούν σε επιβλεπόμενη [92, 93], χωρίς επίβλεψη [94, 95] και ημι-επιβλεπόμενη επιλογή χαρακτηριστικών [96, 97]. Οι επιβλεπόμενες μέθοδοι επιλογής χαρακτηριστικών μπορούν να κατηγοριοποιηθούν ευρύτερα σε μοντέλα φίλτρων, μοντέλα περιτυλίγματος και ενσωματωμένα μοντέλα. Το μοντέλο φίλτρου χωρίζει την επιλογή χαρακτηριστικών από την εκμάθηση ταξινομητή ώστε να μην αλληλεπιδράσει η προκατάληψη ενός αλγορίθμου μάθησης με την προκατάληψη ενός αλγορίθμου επιλογής χαρακτηριστικών. Βασίζεται σε μέτρα γενικών χαρακτηριστικών των δεδομένων εκπαίδευσης, όπως η απόσταση, η συνέπεια, η εξάρτηση, η πληροφορία και ο συσχετισμός. Η ανακούφιση [98], η βαθμολογία Fisher [99] και οι μέθοδοι που βασίζονται στις πληροφορίες [100] είναι μεταξύ των πλέον αντιπροσωπευτικών αλγορίθμων του μοντέλου φίλτρου. Το μοντέλο περιτυλίγματος χρησιμοποιεί την προβλεπτική ακρίβεια ενός προκαθορισμένου αλγορίθμου μάθησης για τον προσδιορισμό της ποιότητας των επιλεγμένων χαρακτηριστικών. Αυτές οι μέθοδοι είναι απαγορευτικά δαπανηρές για την εκτέλεση δεδομένων με μεγάλο αριθμό χαρακτηριστικών.

Λόγω αυτών των αδυναμιών σε κάθε μοντέλο, το ενσωματωμένο μοντέλο, προτάθηκε ώστε να γεφυρωθεί η διαφορά μεταξύ του φίλτρου και των μοντέλων περιτυλίγματος. Πρώτον, ενσωματώνει τα στατιστικά κριτήρια, όπως το μοντέλο φίλτρου, για να επιλέξει διάφορα υποψηφία χαρακτηριστικά υποψηφίων με δεδομένη καρδιανότητα. Δεύτερον, επιλέγει το υποσύνολο με την υψηλότερη ακρίβεια ταξινόμησης [101]. Έτσι, το ενσωματωμένο μοντέλο συνήθως επιτυγχάνει τόσο συγκρίσιμη ακρίβεια με το μοντέλο περιτυλίγματος όσο και συγκρίσιμη απόδοση στο μοντέλο φίλτρου. Το ενσωματωμένο μοντέλο εκτελεί επιλογή χαρακτηριστικών στο χρόνο εκμάθησης. Με άλλα λόγια, επιτυγχάνει την τοποθέτηση μοντέλου και την επιλογή χαρακτηριστικών ταυτόχρονα [102, 103].

Πολλοί ερευνητές έδωσαν επίσης προσοχή στην ανάπτυξη επιλογής χωρίς επίβλεψη. Η επιλογή χαρακτηριστικών χωρίς επίβλεψη είναι ένα λιγότερο περιορισμένο πρόβλημα αναζήτησης χωρίς τις ετικέτες κλάσης και μπορεί να προκαλέσει πολλά εξίσου έγκυρα υποσύνολα χαρακτηριστικών. Με δεδομένα μεγάλης διαστάσεως, είναι απίθανο να ανακτήσει τα σχετικά χαρακτηριστικά χωρίς να λάβει υπόψη πρόσθετους περιορισμούς. Μια άλλη βασική δυσκολία είναι η αντικειμενική μέτρηση των αποτελεσμάτων της επιλογής χαρακτηριστικών [104]. Μια περιεκτική επισκόπηση σχετικά με την επιλογή χαρακτηριστικών χωρίς επίβλεψη μπορεί να βρεθεί στο [105].

Η επιβλεπόμενη επιλογή χαρακτηριστικών αξιολογεί τη συνάφεια των χαρακτηριστικών που επιλέχθηκαν καθοδηγούμενη από τις ετικέτες επισήμανσης. Ωστόσο ένας καλός επιλογέας χρειάζεται επαρκώς επισημασμένα δεδομένα, των οποίων η διαδικασία επισήμανσης είναι χρονοβόρα. Αν και η επιλογή χωρίς επίβλεψη λειτουργεί με μη επισημασμένα δεδομένα, είναι δύσκολο να αξιολογηθεί η συνάφεια των χαρακτηριστικών. Είναι κοινό να έχουμε ένα σύνολο δεδομένων με τεράστιο αριθμό διαστάσεων αλλά μικρό μέγεθος δειγμάτων. Τα δεδομένα μεγάλης διαστάσεως με μικρό πλήθος επισημασμένων δειγμάτων επιτρέπουν επίσης ένα μεγάλο χώρο υποθέσεων, αλλά με πολύ λίγους περιορισμούς (ετικέτες). Ο συνδυασμός των δύο χαρακτηριστικών δεδομένων δηλώνει μια νέα ερευνητική πρόκληση. Σύμφωνα με την παραδοχή ότι τα επισημασμένα και μη επισημασμένα δεδομένα λαμβάνονται από τον ίδιο πληθυσμό που παράγεται από την έννοια του στόχου, η ημι-επιβλεπόμενη επιλογή χαρακτηριστικών χρησιμοποιεί τόσο τα επισημασμένα όσο και τα μη επισημασμένα δεδομένα για να εκτιμήσει την συνάφεια των επιλεγόμενων χαρακτηριστικών [97].

Η στάθμιση χαρακτηριστικών θεωρείται ως γενίκευση της επιλογής χαρακτηριστικών. Στην επιλογή χαρακτηριστικών, σε κάθε ένα χαρακτηριστικό έχει εκχωρηθεί ένα δυαδικό βάρος, όπου 1 σημαίνει ότι το χαρακτηριστικό έχει επιλεγεί και 0 σε διαφορετική περίπτωση. Ωστόσο, η στάθμιση χαρακτηριστικών αποδίδει μια τιμή, συνήθως στο διάστημα $[0, 1]$ ή $[-1, 1]$, σε κάθε λειτουργία. Όσο μεγαλύτερη είναι αυτή η τιμή, τόσο πιο σημαντικό θα είναι το χαρακτηριστικό. Στα περισσότερα από αυτά οι αλγόριθμοι βάρους χαρακτηριστικών αντιστοιχούν σε ένα ενιαίο (παγκόσμιο) βάρος για κάθε χαρακτηριστικό σε όλες τις περιπτώσεις. Ωστόσο, η σχετική σημασία, η συνάφεια και ο θόρυβος στις

διάφορες διαστάσεις μπορεί να διαφέρουν σημαντικά ανάλογα με την τοποθεσία δεδομένων. Υπάρχουν τοπικοί αλγόριθμοι επιλογής χαρακτηριστικών όπου η επιλογή των χαρακτηριστικών γίνεται ειδικά για μια δοκιμαστική περίπτωση (τοπική), η οποία είναι συνηθισμένη σε απλούς αλγορίθμους όπως ο k-Nearest Neighbor (kNN) [106, 107]. Η ιδέα είναι ότι η επιλογή χαρακτηριστικών ή η στάθμιση γίνεται στο χρόνο ταξινόμησης (και όχι σε χρόνο εκπαίδευσης), διότι η γνώση ενός δοκιμαστικού δεδομένου οξύνει τη δυνατότητα επιλογής χαρακτηριστικών.

Τυπικά, μια μέθοδος επιλογής χαρακτηριστικών αποτελείται από τέσσερα βασικά βήματα, δηλαδή από την παραγωγή υποσυνόλων, την αξιολόγηση των υποσυνόλων, το κριτήριο τερματισμού και την επικύρωση των αποτελεσμάτων. Στο πρώτο βήμα, ένα υποσύνολο υποψηφίων χαρακτηριστικών θα επιλεγεί με βάση μια δεδομένη στρατηγική αναζήτησης, η οποία αποστέλλεται, στο δεύτερο στάδιο, να αξιολογηθεί σύμφωνα με κάποιο κριτήριο αξιολόγησης. Το υποσύνολο που ταιριάζει καλύτερα με το κριτήριο αξιολόγησης, θα επιλεγεί από όλους τους υποψηφίους που έχουν αξιολογηθεί μετά την εκπλήρωση του κριτηρίου τερματισμού. Στο τελικό βήμα, το επιλεγμένο υποσύνολο θα είναι επικυρωμένο χρησιμοποιώντας γνώση τομέα ή ένα σύνολο επικύρωσης.

2.3.5 Παραδοσιακές Προσεγγίσεις ML

Naive Bayes

Ο Naive Bayes είναι ένας δημοφιλής και απλός αλγόριθμος για δυαδικές και πολυτάξιες εργασίες ταξινόμησης. Αρχικά εφαρμόστηκε για το φιλτράρισμα ανεπιθύμητων ηλεκτρονικών μηνυμάτων στο Διαδίκτυο [108]. Ονομάζεται Naive Bayes επειδή ο υπολογισμός των πιθανοτήτων για κάθε υπόθεση είναι απλοποιημένος για να καταστήσει τον υπολογισμό τους εφικτό. Επίσης, όπως προκύπτει από το όνομα, βασίζεται στο Θεώρημα Bayes, σύμφωνα με το οποίο:

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)} \quad (2.2)$$

- $P(h|d)$: πιθανότητα της υπόθεσης h δεδομένης της πληροφορίας d . Αυτή είναι η posterior πιθανότητα.
- $P(d|h)$: πιθανότητα της πληροφορίας d δεδομένου ότι η υπόθεση h είναι αληθής.
- $P(h)$: πιθανότητα να είναι η υπόθεση h αληθής (ανεξαρτήτως των δεδομένων). Αυτή είναι η prior πιθανότητα της υπόθεσης h .
- $P(d)$: πιθανότητα της πληροφορίας d (ανεξαρτήτως της υπόθεσης).

Όταν χρησιμοποιούμε τον ταξινομητή Naive Bayes μας ενδιαφέρει να υπολογίσουμε την οπίσθια πιθανότητα $P(h|d)$ από την προηγούμενη πιθανότητα $P(h)$ με $P(D)$ και $P(d|h)$. Μετά τον υπολογισμό της οπίσθιας πιθανότητας για μια σειρά διαφορετικών υποθέσεων, μπορούμε να επιλέξουμε την υπόθεση με την υψηλότερη πιθανότητα. Αυτή είναι η μέγιστη πιθανή υπόθεση και τυπικά ονομάζεται Maximum A Posteriori (MAP). Αυτό μπορεί να γραφτεί ως εξής:

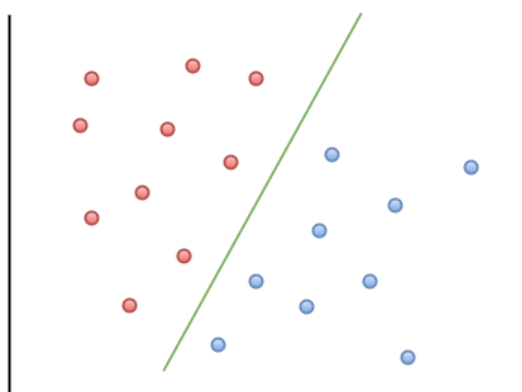
$$MAP(h) = \max(P(h|d)) = \max\left(\frac{P(d|h)P(h)}{P(d)}\right) = \max(P(d|h)P(h)) \quad (2.3)$$

Το $P(d)$ είναι ένας όρος ομαλοποίησης που μας επιτρέπει να υπολογίσουμε την πιθανότητα. Μπορούμε να το απαλείψουμε όταν μας ενδιαφέρει η πιο πιθανή υπόθεση, καθώς είναι σταθερός και χρησιμοποιείται μόνο για την ομαλοποίηση. Εστιάζοντας στην εργασία ταξινόμησης, αν έχουμε αθροιστή αριθμό παρουσιών σε κάθε κατηγορία στα δεδομένα εκπαίδευσης, τότε η πιθανότητα κάθε κατηγορίας (π.χ. $P(h)$) θα είναι ίση.

Support Vector Machine

Οι Support Vector Machines (SVM) είναι ίσως ένας από τους πιο δημοφιλείς αλγόριθμους μηχανικής μάθησης. Ήταν εξαιρετικά δημοφιλής όταν αναπτύχθηκε κατά τη δεκαετία του 1990 και εξακολουθεί να είναι η μέθοδος go-to για έναν αλγόριθμο υψηλής απόδοσης με ελάχιστο συντονισμό.

Τα SVM βασίζονται στην έννοια των επιπέδων απόφασης που καθορίζουν τα όρια απόφασης. Μια επιφάνεια απόφασης είναι εκείνο που χωρίζει μεταξύ ενός συνόλου αντικειμένων που έχουν διαφορετικές ταξικές ιδιότητες. Ένα σχηματικό παράδειγμα φαίνεται στο σχήμα 2.4. Σε αυτό το παράδειγμα, τα αντικείμενα ανήκουν είτε στην κλάση "μπλε" είτε "κόκκινο". Η διαχωριστική γραμμή ορίζει ένα όριο στη δεξιά πλευρά του οποίου όλα τα αντικείμενα είναι "μπλε" και στα αριστερά της οποίας όλα τα αντικείμενα είναι "κόκκινα". Κάθε νέο αντικείμενο (λευκός κύκλος) που εμπίπτει στα δεξιά φέρει ετικέτα, δηλ. ταξινομείται ως "μπλε" (ή ταξινομείται ως "κόκκινο" αν πέσει στα αριστερά της διαχωριστικής γραμμής).



Σχήμα 2.4: Γραμμική ταξινόμηση χρησιμοποιώντας SVM

Τα παραπάνω είναι ένα κλασικό παράδειγμα ενός γραμμικού ταξινομητή, δηλ. ενός ταξινομητή που χωρίζει ένα σύνολο αντικειμένων στις αντίστοιχες ομάδες τους («μπλε» και «κόκκινο» σε αυτή την περίπτωση) με μια γραμμή. Ωστόσο, οι περισσότερες εργασίες ταξινόμησης δεν είναι τόσο απλές και απαιτούνται συχνά πιο σύνθετες δομές προκειμένου να γίνει ο βέλτιστος διαχωρισμός, δηλαδή να ταξινομηθούν σωστά τα νέα αντικείμενα (περιπτώσεις δοκιμών) βάσει των διαθέσιμων παραδειγμάτων (περιπτώσεις τρένων). Τα προβλήματα ταξινόμησης που απαιτούν διαχωρισμό γραμμών διαχωρισμού για τη διάκριση μεταξύ αντικειμένων διαφορετικών κατηγοριών είναι γνωστά ως ταξινομητές υπερ-επιπέδου. Τα SVM είναι ιδιαίτερα κατάλληλα για την αντιμετώπιση τέτοιων εργασιών. Ένα μοντέλο SVM χαρτογραφεί τα αρχικά αντικείμενα σε μια αναδιάταξη χρησιμοποιώντας ένα σύνολο μαθηματικών λειτουργιών, γνωστών ως πυρήνες. Η διαδικασία της αναδιάταξης των αντικειμένων είναι γνωστή ως μετασχηματισμός. Σημειώστε ότι σε αυτή τη νέα ρύθμιση, τα αντικείμενα που έχουν αντιστοιχιστεί είναι γραμμικά διαχωρίσιμα και έτσι, αντί να κατασκευάσουμε την αρχική περίπλοκη καμπύλη, το μόνο που πρέπει να κάνουμε είναι να βρούμε μια βέλτιστη γραμμή που μπορεί να διαχωρίσει τα «μπλε» και τα «κόκκινα» αντικείμενα.

Random Forest

Ο Random Forest είναι μια μέθοδος μάθησης για τα προβλήματα ταξινόμησης, παλινδρόμησης και άλλα αντίστοιχα που λειτουργούν κατασκευάζοντας ένα πλήθος δέντρων αποφάσεων στο χρόνο εκπαίδευσης και εξάγοντας την κλάση που είναι η λειτουργία των τάξεων (ταξινόμηση) ή η μέση πρόβλεψη (παλινδρόμηση) τα μεμονωμένα δέντρα. Τα Random Forest δεν παρουσιάζουν την τάση για overfitting σε σύγκριση με τα δέντρα αποφάσεων.

Η έννοια του δέντρου αποφάσεων είναι στενότερα συνδεδεμένο με το σύστημα βασισμένο σε κανόνες (rule-based). Δεδομένου του συνόλου επισημασμένων δεδομένων εκπαίδευσης και τα εξαγόμενα χαρακτηριστικά, ο random forest θα παρουσιάσει ορισμένους κανόνες ρύθμισης. Οι ίδιοι κανόνες ρύθμισης μπορούν να χρησιμοποιηθούν για την εκτέλεση της πρόβλεψης στο σύνολο δεδομένων δοκιμής. Ο ψευδοκώδικας που χρησιμοποιείται στα δέντρα απόφασης μπορεί να συνοψιστεί ως εξής:

1. Τοποθετήστε το καλύτερο χαρακτηριστικό του συνόλου δεδομένων στη ρίζα του δέντρου.
2. Διαχωρίστε το σύνολο εκπαίδευσης σε υποσύνολα. Τα υποσύνολα πρέπει να δημιουργούνται με τέτοιο τρόπο ώστε κάθε υποσύνολο να περιέχει δεδομένα με την ίδια τιμή για ένα χαρακτηριστικό.
3. Επαναλάβετε το βήμα 1 και το βήμα 2 σε κάθε υποσύνολο μέχρι να βρείτε κόμβους φύλλων σε όλους τους κλάδους του δέντρου.

Στα δέντρα απόφασης, για την πρόβλεψη μιας ετικέτας κλάσης για ένα δεδομένο ξεκινάμε από τη ρίζα του δέντρου. Συγκρίνουμε τις τιμές της ιδιότητας ρίζας με το χαρακτηριστικό της εγγραφής. Με βάση τη σύγκριση, ακολουθούμε τον κλάδο που αντιστοιχεί στην τιμή και μεταβαίνουμε στον επόμενο κόμβο. Συνεχίζουμε να συγκρίνουμε τις τιμές χαρακτηριστικών μας με άλλους εσωτερικούς κόμβους του δέντρου μέχρι να φτάσουμε σε έναν κόμβο φύλλων με την προβλεπόμενη ετικέτα κλάσης.

Ένας Random Forest αλγόριθμος έχει όλα τα υπερπαραμετρικά στοιχεία ενός decision-tree ταξινομητή και επίσης όλα τα υπερπαραμετρικά στοιχεία ενός bagging ταξινομητή για τον έλεγχο του ensembling. Ο Random Forest επιφέρει επιπλέον τυχαιότητα στο μοντέλο, κατά την ανάπτυξη των δέντρων. Αντί να ψάχνει για τον καλύτερο δυνατό διαχωρισμό ενός κόμβου, αναζητά την καλύτερη δυνατή δημιουργία ενός τυχαίου υποσυνόλου χαρακτηριστικών. Αυτή η διαδικασία δημιουργεί μια μεγάλη ποικιλία, η οποία γενικά οδηγεί σε ένα καλύτερο μοντέλο. Επομένως, όταν αναπτύσσεται ένα δέντρο στον Random Forest, μόνο ένα τυχαίο υποσύνολο των χαρακτηριστικών εξετάζεται για τον διαχωρισμό ενός κόμβου. Τα δέντρα μπορούν να γίνουν ακόμα πιο τυχαία, χρησιμοποιώντας τυχαία κατώτατα όρια πάνω από αυτά, για κάθε χαρακτηριστικό, αντί να γίνεται αναζήτηση για τα καλύτερα δυνατά όρια (όπως συμβαίνει σε ένα κανονικό δέντρο αποφάσεων). Ο ψευδοκώδικας για τον Random Forest παρουσιάζεται παρακάτω:

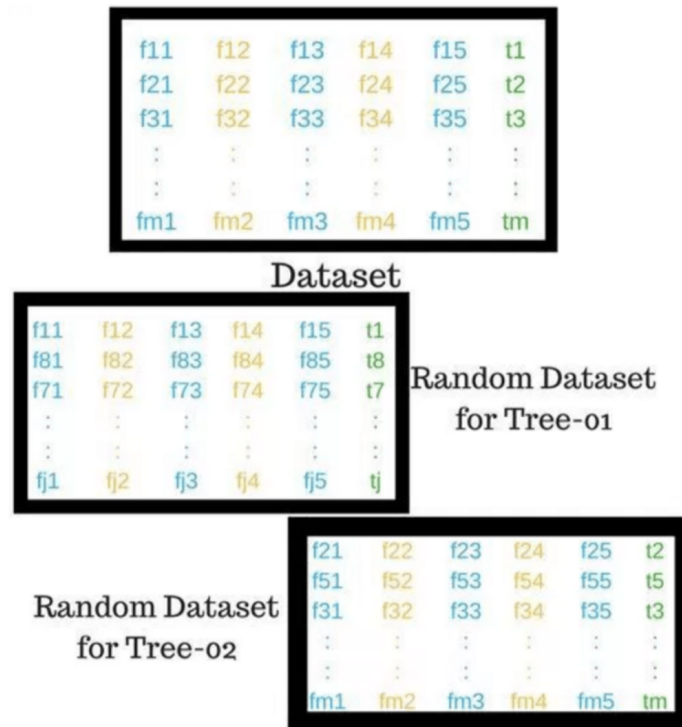
1. Τυχαία επίλεξε k χαρακτηριστικά από τα m συνολικά χαρακτηριστικά, όπου $k \ll m$.
2. Ανάμεσα στα k χαρακτηριστικά, υπολόγισε τον κόμβο d χρησιμοποιώντας το καλύτερο σημείο διαχωρισμού.
3. Χώρισε τον κόμβο σε κόμβους-παιδιά χρησιμοποιώντας τον βέλτιστο διαχωρισμό.
4. Επανάλαβε τα βήματα 1 μέχρι 3 μέχρι l αριθμός από κόμβους να έχουν δημιουργηθεί.
5. Κατασκεύασε το "δάσος" επαναλαμβάνοντας τα βήματα 1 έως 4 για n φορές για την δημιουργία n δέντρων.

Η διαδικασία εκπαίδευσης του ταξινομητή Random Forest απεικονίζεται επίσης στο σχήμα 2.5¹.

Μετά την εκπαίδευση του ταξινομητή όπως περιγράφεται παραπάνω, ακολουθεί η διαδικασία πρόβλεψης:

1. Πάρε τα χαρακτηριστικά των δεδομένων δοκιμής και χρησιμοποίησε τους κανόνες του κάθε τυχαία κατασκευασμένου decision-tree για την πρόβλεψη του αποτελέσματος και αποθήκευσε την πρόβλεψη αυτή (target).

¹ <http://dataaspirant.com>



Σχήμα 2.5: Διαδικασία εκπαίδευσης του τυχαίου δασικού ταξινομητή

2. Υπολόγισε τις ψήφους για τις προβλέψεις του κάθε target.
3. Λάβε το target που έχει προβλεφθεί με τις περισσότερες ψήφους ως την τελική πρόβλεψη του ταξινομητή random forest.

2.3.6 Διανύσματα Λέξεων (Word Embeddings)

Τα διανύσματα λέξεων (word embeddings) χρησιμοποιούνται συνήθως σε πολλές εργασίες NLP επειδή είναι χρήσιμες αναπαραστάσεις λέξεων και συχνά οδηγούν σε καλύτερη απόδοση στις διάφορες εργασίες που εκτελούνται. Οι ενσωματώσεις λέξεων επιτρέπουν την αναπαράσταση λέξεων από μια σειρά αριθμών.

Word2vec

Το Word2vec [109] είναι ένα νευρωνικό δίκτυο δύο επιπέδων που επεξεργάζεται το κείμενο. Η είσοδος του είναι ένα corpus κειμένου και η έξοδος του είναι ένα σύνολο διανυσμάτων: διανύσματα χαρακτηριστικών για λέξεις σε αυτό το corpus. Ενώ το Word2vec δεν είναι ένα βαθύ νευρωνικό δίκτυο (βλ. Ενότητα 2.4), μετατρέπει το κείμενο σε αριθμητική μορφή που μπορούν να καταλάβουν τα βαθιά δίκτυα.

Ο σκοπός και η χρησιμότητα του Word2vec είναι η ομαδοποίηση των φορέων παρόμοιων λέξεων μαζί στο διανυσματικό χώρο. Δηλαδή ανιχνεύει μαθηματικά ομοιότητες μεταξύ λέξεων. Το Word2vec δημιουργεί διανύσματα που είναι κατανομημένες αριθμητικές αναπαραστάσεις χαρακτηριστικών λέξεων, χαρακτηριστικά όπως το πλαίσιο μεμονωμένων λέξεων. Το κάνει χωρίς ανθρώπινη παρέμβαση.

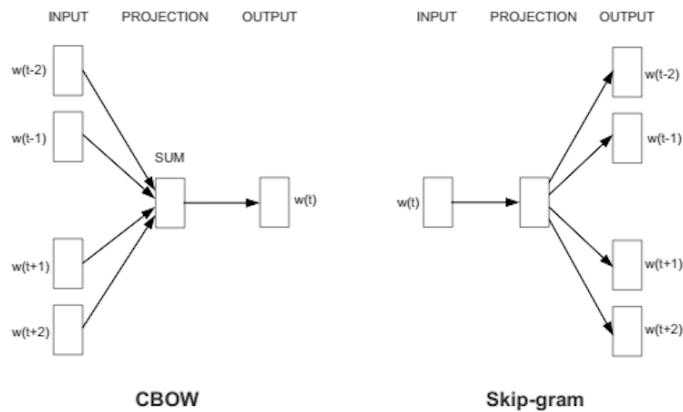
Λαμβάνοντας υπόψη αρκετά δεδομένα, τη χρήση και τα συμφοραζόμενα, το Word2vec μπορεί να κάνει πολύ ακριβείς εικασίες για το νόημα μιας λέξης με βάση προηγούμενες εμφανίσεις. Αυτές οι εικασίες μπορούν να χρησιμοποιηθούν για να εδραιωθεί η συσχέτιση μιας λέξης με άλλες λέξεις (π.χ. "άνθρωπος" είναι ένα "αγόρι", μια "γυναίκα" και ένα "κορίτσι") ή έγγραφα συμπλέγματος και

να τα ταξινομήσει κατά θέμα. Αυτά τα συμπλέγματα μπορούν να αποτελέσουν τη βάση αναζήτησης, ανάλυσης συναισθημάτων και συστάσεων σε διάφορους τομείς όπως η επιστημονική έρευνα, η νομική ανακάλυψη, το ηλεκτρονικό εμπόριο και η διαχείριση πελατειακών σχέσεων.

Η έξοδος του νευρωνικού δικτύου Word2vec είναι ένα λεξιλόγιο στο οποίο κάθε στοιχείο έχει ένα διάνυσμα συνδεδεμένο με αυτό, το οποίο μπορεί να τροφοδοτηθεί σε ένα βαθύ δίκτυο μάθησης ή απλά να χρησιμοποιηθεί για να ανιχνεύσει τις σχέσεις μεταξύ των λέξεων.

Με τη μέτρηση της ομοιότητας του συνημίτονου, καμία ομοιότητα δεν εκφράζεται ως γωνία 90 μοιρών, ενώ η συνολική ομοιότητα του 1 είναι μια γωνία 0 μοιρών, η πλήρης επικάλυψη.

Οι φορείς που χρησιμοποιούμε για να αναπαριστούμε λέξεις ονομάζονται neural word embeddings. Το Word2vec είναι παρόμοιο με έναν αυτόματο κωδικοποιητή, που κωδικοποιεί κάθε λέξη σε ένα διάνυσμα, αλλά αντί να εκπαιδεύει σε σχέση με τις λέξεις εισόδου μέσω της ανασυγκρότησης, όπως κάνει μια περιορισμένη μηχανή Boltzmann, το word2vec εκπαιδεύει λέξεις σε σχέση με άλλες λέξεις που τις γειτονεύουν στο σώμα εισόδου. Αυτό γίνεται με έναν από τους δύο τρόπους, είτε χρησιμοποιώντας το πλαίσιο για να προβλεφθεί μια λέξη στόχου (μια μέθοδος γνωστή ως Continuous Bag-Of-Words, ή CBOW), είτε χρησιμοποιώντας μια λέξη για να προβλεφθεί ένα πλαίσιο στόχου, το οποίο καλείται skip-gram. Αυτές οι δύο μέθοδοι παρουσιάζονται στο σχήμα 2.6.



Σχήμα 2.6: Μέθοδοι CBOW και skip-gram για την κατάρτιση word2vec embeddings

Όντας πιθανολογικοί στη φύση τους, αυτές υποτίθεται ότι επιτελούν ανώτερες από τις προσδιοριστικές μεθόδους (γενικά). Παρόλο που το CBOW (πρόβλεψη της λέξης-στόχου από το πλαίσιο) και το skip-gram (πρόβλεψη των λέξεων-πλαισίων από την λέξη-στόχο) είναι απλώς ανεστραμμένες μέθοδοι, η καθεμιά έχει τα πλεονεκτήματα / μειονεκτήματα της. Δεδομένου ότι το CBOW μπορεί να χρησιμοποιήσει πολλές λέξεις-πλαίσια για να προβλέψει τη λέξη-στόχο, μπορεί ουσιαστικά να εξομαλύνει την κατανομή. Αυτό είναι ουσιαστικά σαν νομιμοποίηση και προσφέρει πολύ καλές επιδόσεις όταν τα δεδομένα εισόδου μας δεν είναι τόσο μεγάλα. Ωστόσο, το μοντέλο skip-gram είναι πιο fine-grained, ώστε να μπορούμε να εξάγουμε περισσότερες πληροφορίες και ουσιαστικά να έχουμε ακριβέστερες αναπαραστάσεις των λέξεων όταν έχουμε ένα μεγάλο σύνολο δεδομένων (τα μεγάλα δεδομένα είναι πάντα ο καλύτερος ρυθμιστής). Το Skip-gram με αρνητική υπο-δειγματοληψία ξεπερνά κάθε άλλη μέθοδο γενικά.

GloVe

Σε αντίθεση με το word2vec, το GloVe [110] επιδιώκει να καταστήσει σαφές αυτό που δηλώνει το word2vec εμμέσως: Η κωδικοποίηση της έννοιας ως μετατόπισης διανυσμάτων σε ένα χώρο ενσωμάτωσης - φαινομενικά μόνο ένα δευτερεύον προϊόν του word2vec - είναι ο καθορισμένος στόχος

του GloVe.

Συγκεκριμένα, το GloVe απεικονίζει ότι ο λόγος των πιθανών συν-εμφανίσεων δύο λέξεων (και όχι οι ίδιες οι πιθανότητες συνυπάρχουσας εμφάνισης) είναι αυτό που περιέχει πληροφορίες και έτσι κοιτάζει να κωδικοποιήσει αυτές τις πληροφορίες ως διανυσματικές διαφορές.

Για να επιτευχθεί αυτό, η μέθοδος GloVe αποσκοπεί άμεσα στη μείωση της διαφοράς μεταξύ του εσωτερικού γινομένου των διανυσμάτων δύο λέξεων και του λογαρίθμου του αριθμού των συν-εμφανίσεών τους. Καθώς οι μετρήσεις συν-εμφάνισης μπορούν να κωδικοποιηθούν άμεσα σε μια μήτρα συν-εμφάνισης λέξεων-πλαισίων, το GloVe παίρνει ένα τέτοιο πλέγμα παρά ολόκληρο το σώμα ως είσοδο.

2.4 Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks)

2.4.1 Εισαγωγή

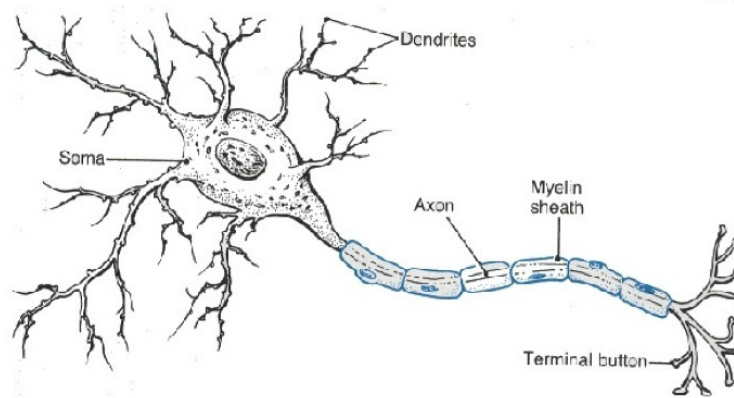
Τα τελευταία 10 χρόνια, τα καλύτερα συστήματα τεχνητής νοημοσύνης - όπως τα αναγνωριστικά ομιλίας σε smartphones ή ο τελευταίος αυτόματος μεταφραστής της Google - προέκυψαν από μια τεχνική που ονομάζεται "βαθιά μάθηση". Η βαθιά εκμάθηση είναι στην πραγματικότητα ένα νέο όνομα για μια προσέγγιση της τεχνητής νοημοσύνης που ονομάζεται νευρωνικά δίκτυα, τα οποία εισέρχονται και εξέρχονται από τη μόδα για περισσότερα από 70 χρόνια. Τα νευρωνικά δίκτυα προτάθηκαν για πρώτη φορά το 1944 από τους Warren McCulloch και Walter Pitts, δυο ερευνητές του Πανεπιστημίου του Σικάγο που μετακόμισαν στο MIT το 1952 ως ιδρυτικά μέλη αυτού που καλείται μερικές φορές το πρώτο γνωστικό τμήμα της επιστήμης. Τα νευρωνικά δίκτυα ήταν ένας σημαντικός τομέας έρευνας τόσο στην επιστήμη των νευροεπιστημών όσο και στην επιστήμη των υπολογιστών έως το 1969, όταν, σύμφωνα με την επιστήμη των υπολογιστών, σκοτώθηκαν από τους μαθητές του MIT Marvin Minsky και Seymour Papert, οι οποίοι αργότερα θα γίνονταν συν-διευθυντές στο νέο εργαστήριο τεχνητής νοημοσύνης του MIT. Η τεχνική παρουσίασε μια αναζωπύρωση στη δεκαετία του '80, έπεσε ξανά στην έκλειψη κατά την πρώτη δεκαετία του νέου αιώνα και επέστρεψε σαν gangbusters στη δεύτερη, τροφοδοτούμενη σε μεγάλο βαθμό από την αυξημένη ικανότητα επεξεργασίας των γραφικών chip.

2.4.2 Τεχνητά Νευρωνικά Δίκτυα

Ένα Δίκτυο Τεχνητών Νευρώνων (Artificial Neural Network - ANN), ευρέως γνωστό ως Νευρωνικό Δίκτυο, είναι ένα υπολογιστικό μοντέλο βασισμένο στη δομή και τις λειτουργίες των βιολογικών νευρωνικών δικτύων. Είναι εμπνευσμένο από τον ανθρώπινο εγκέφαλο και στοχεύει στη λήψη, επεξεργασία και μετάδοση πληροφοριών από την άποψη της Πληροφορικής.

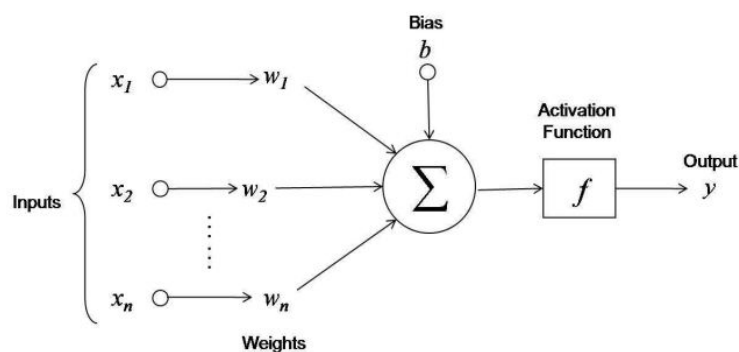
Η ακριβής λειτουργία του ανθρώπινου εγκεφάλου είναι ακόμα ένα μυστήριο. Ωστόσο, ορισμένες πτυχές αυτού του καταπληκτικού επεξεργαστή είναι γνωστές. Συγκεκριμένα, το πιο βασικό στοιχείο του ανθρώπινου εγκεφάλου είναι ένας συγκεκριμένος τύπος κυττάρου ο οποίος, σε αντίθεση με το υπόλοιπο σώμα, δεν φαίνεται να αναγεννάται. Επειδή αυτός ο τύπος κυττάρου είναι το μόνο μέρος του σώματος που δεν αντικαθίσταται αργά, θεωρείται ότι αυτά τα κύτταρα μας παρέχουν τις ικανότητες μας να θυμόμαστε, να σκεφτόμαστε και να εφαρμόζουμε προηγούμενες εμπειρίες στην κάθε δράση μας. Αυτά τα κύτταρα, και τα 100 δισεκατομμύρια από αυτά, είναι γνωστά ως νευρώνες. Κάθε ένας από αυτούς τους νευρώνες μπορεί να συνδεθεί με έως και 200.000 άλλους νευρώνες, αν και τυπικά είναι 1.000 έως 10.000. Η δύναμη του ανθρώπινου νου προέρχεται από τους απλούς αριθμούς αυτών των βασικών συνιστωσών και τις πολλαπλές συνδέσεις μεταξύ τους. Προέρχεται επίσης από τον γενετικό προγραμματισμό και τη μάθηση.

Οι μεμονωμένοι νευρώνες είναι περίπλοκοι. Ένα παράδειγμα της δομής ενός βιολογικού νευρώνα απεικονίζεται στο σχήμα 2.7² Έχουν μια μυριάδα από μέρη, υποσυστήματα και μηχανισμούς ελέγχου. Μεταφέρουν πληροφορίες μέσω πλήθους ηλεκτροχημικών οδών. Υπάρχουν πάνω από εκατό διαφορετικές κατηγορίες νευρώνων, ανάλογα με τη μέθοδο ταξινόμησης που χρησιμοποιείται. Μαζί αυτοί οι νευρώνες και οι συνδέσεις τους σχηματίζουν μια διαδικασία η οποία δεν είναι δυαδική, ούτε σταθερή και ούτε σύγχρονη. Εν ολίγοις, δεν είναι τίποτα όπως οι ηλεκτρονικοί υπολογιστές που διατίθενται σήμερα, ή ακόμα και τα ANNs.



Σχήμα 2.7: Επισκόπηση ενός βιολογικού νευρώνα

Αυτά τα ANN προσπαθούν να αναπαράγουν μόνο τα πιο βασικά στοιχεία αυτού του περίπλοκου, ευπροσάρμοστου και ισχυρού οργανισμού. Συγκεκριμένα, ο στόχος των ANNs δεν είναι η μεγαλοπρεπής ανακατασκευή του εγκεφάλου. Αντίθετα, οι ερευνητές που ασχολούνται με τα ANNs επιζητούν την κατανόηση των δυνατοτήτων της φύσης για την οποία οι άνθρωποι μπορούν να σχεδιάσουν λύσεις σε προβλήματα που δεν έχουν επιλυθεί από την παραδοσιακή υπολογιστική. Για να γίνει αυτό, η βασική μονάδα των νευρωνικών δικτύων, οι τεχνητοί νευρώνες, προσομοιάζουν τις τέσσερις βασικές λειτουργίες φυσικών νευρώνων. Το σχήμα 2.8 δείχνει μια θεμελιώδη αναπαράσταση ενός τεχνητού νευρώνα.



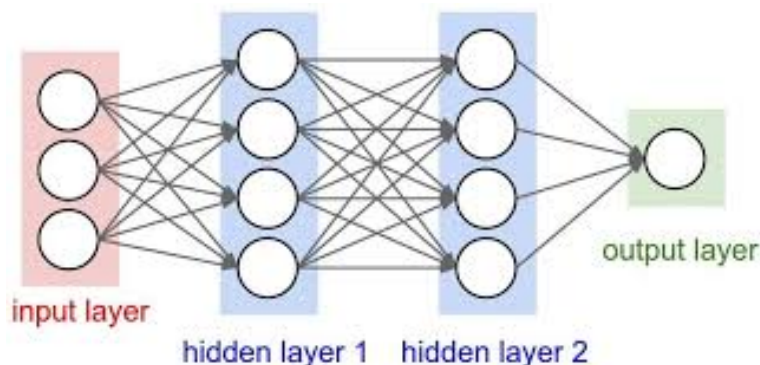
Σχήμα 2.8: Επισκόπηση των λειτουργιών ενός τεχνητού νευρώνα

Όπως παρουσιάζεται στο Σχήμα 2.8, οι διάφορες εισόδους στο δίκτυο αντιπροσωπεύονται από το μαθηματικό σύμβολο, x_n . Κάθε μία από αυτές τις εισόδους πολλαπλασιάζεται με ένα βάρος σύνδεσης. Αυτά τα βάρη αντιπροσωπεύονται από w_n . Στην απλούστερη περίπτωση, τα προϊόντα αυτά αθροίζονται απλά, τροφοδοτούνται μέσω μίας συνάρτησης μεταφοράς f για να παράγουν ένα απο-

² <http://www.neuropsychologysketches.com/Neurons.html>

τέλεσμα και στη συνέχεια παράγουν την έξοδο y . Στην πιο γενική περίπτωση, υπάρχουν διάφορες λειτουργίες μεταφοράς (ενεργοποίησης) από την άποψη ενός τεχνητού νευρώνα.

Για να μάθουν πολύπλοκες μη γραμμικές λειτουργίες, μπορούν να σχεδιαστούν και να υλοποιηθούν αρχιτεκτονικές που συνδυάζουν διάφορους τεχνητούς νευρώνες. Τέτοιες αρχιτεκτονικές ονομάζονται πολυεπίπεδα αντίληπτρα (Multi-Layer Perceptrons - MLPs). Αντί για MLPs, μπορούν να υλοποιηθούν και τα FFNNs, όπου κάθε νευρώνας συνδέεται με όλους τους νευρώνες του προηγούμενου στρώματος και δεν υπάρχουν συνδέσεις μεταξύ των νευρώνων του ίδιου στρώματος.



Σχήμα 2.9: Επισκόπηση της αρχιτεκτονικής ενός ANN

Η δομή ενός ANN απεικονίζεται στο σχήμα 2.9³. Όπως παρατηρήθηκε, το δίκτυο αποτελείται από τα ακόλουθα στρώματα:

- Στρώμα Εσόδου (Input layer): Όλες οι εισοδοί δίνονται στο δίκτυο μέσω αυτού του στρώματος.
- Κρυφά Επίπεδα (Hidden layer(s)): Μέσω αυτού του στρώματος, η είσοδος υπόκειται σε επεξεργασία και εξάγονται τα χαρακτηριστικά από το δίκτυο. Όσο κινούμαστε προς ανώτερα κρυμμένα στρώματα, εξάγονται χαρακτηριστικά ανώτερου σημασιολογικού περιεχομένου.
- Στρώμα Εξόδου (Output layer): Μετά την επεξεργασία των δεδομένων, η απόφαση λαμβάνεται από το δίκτυο σε αυτό το στρώμα.

Εάν ένα νευρωνικό δίκτυο αποτελείται από περισσότερα από ένα κρυμμένα στρώματα τότε αναφέρεται ως "Βαθύ Νευρωνικό Δίκτυο" ("Deep Neural Network" (DNN)). Όταν αυξάνεται το βάθος ενός νευρικού δικτύου και καταργούνται οι μη γραμμικές λειτουργίες μεταφοράς, τότε το δίκτυο μπορεί να μάθει πιο σύνθετες λειτουργίες.

Εκπαίδευση

Ένα βασικό χαρακτηριστικό της διαδικασίας εκπαίδευσης των νευρωνικών δικτύων είναι μια επαναληπτική μαθησιακή διαδικασία στην οποία οι περιπτώσεις δεδομένων παρουσιάζονται στο δίκτυο μία φορά και τα βάρη που σχετίζονται με τις τιμές εισόδου ρυθμίζονται κάθε φορά. Αφού παρουσιαστούν όλες οι περιπτώσεις, η διαδικασία ξεκινά ξανά από την αρχή. Κατά τη διάρκεια αυτής της φάσης μάθησης, το δίκτυο μαθαίνει προσαρμόζοντας τα βάρη έτσι ώστε να είναι σε θέση να προβλέψει τη σωστή ετικέτα κλάσης των δειγμάτων εισόδου. Σε αυτή την ενότητα θα παρουσιαστούν και θα περιγραφούν οι βασικές διαδικασίες που απαιτούνται για την εκπαίδευση.

Συνάρτηση ενεργοποίησης (Activation function). Οι συναρτήσεις ενεργοποίησης είναι σημαντικές για ένα ANN ώστε να μάθει και να κατανοήσει τα περίπλοκα πρότυπα. Η κύρια λειτουργία

³ <https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/>

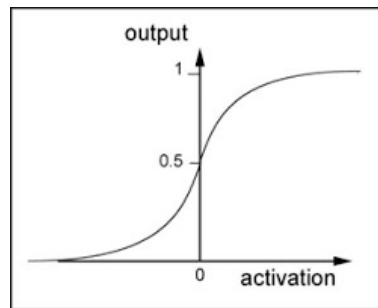
της είναι η εισαγωγή μη γραμμικών ιδιοτήτων στο δίκτυο. Αυτό που κάνει είναι ότι υπολογίζει το "σταθμισμένο άθροισμα", προσθέτει κατεύθυνση και αποφασίζει εάν θα "ενεργοποιήσει" ένα συγκεκριμένο νευρώνα ή όχι. Η λειτουργία μη γραμμικής ενεργοποίησης βοηθάει το μοντέλο να κατανοεί την πολυπλοκότητα και να δίνει ακριβή αποτελέσματα.

Οι συναρτήσεις που χρησιμοποιούνται συνήθως ως συναρτήσεις ενεργοποίησης σε ένα ANN είναι:

- *Σιγμοειδής συνάρτηση (Sigmoid function)*: Η σιγμοειδής συνάρτηση είναι μια λειτουργία ενεργοποίησης όπου ζυγίζει τις τιμές μεταξύ 0 και 1 εφαρμόζοντας ένα κατώφλι και μπορεί να περιγραφεί από την ακόλουθη εξίσωση:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.4)$$

Τη απεικόνιση αυτής της λειτουργίας παρουσιάζεται στο Σχήμα 2.10. Όταν το σταθμισμένο ποσό εφαρμόζεται στη θέση x , οι τιμές κλιμακώνονται μεταξύ 0 και 1. Η τιμή δεν φθάνει ποτέ στο μηδέν ούτε υπερβαίνει το 1 στην παραπάνω εξίσωση. Οι μεγάλοι αρνητικοί αριθμοί κλιμακώνονται προς το 0 και οι μεγάλοι θετικοί αριθμοί κλιμακώνονται προς 1.



Σχήμα 2.10: Sigmoid λειτουργία

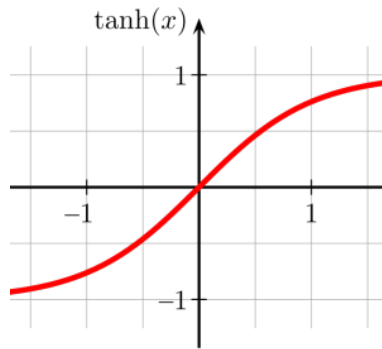
- *Υπερβολική εφαπτομένη (Hyperbolic tangent)*: Η συνάρτηση \tanh είναι μια συνάρτηση ενεργοποίησης που επανακαθορίζει τις τιμές μεταξύ -1 και 1 εφαρμόζοντας ένα κατώφλι ακριβώς όπως μια λειτουργία sigmoid. Το πλεονέκτημα αυτής της συνάρτησης είναι ότι οι τιμές της \tanh είναι κεντραρισμένες στο μηδέν που βοηθούν τον επόμενο νευρώνα κατά τη διάρκεια της διάδοσης. Αυτή η λειτουργία περιγράφεται ως εξής:

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.5)$$

Η απεικόνιση αυτής της λειτουργίας παρουσιάζεται στο σχήμα 2.11. Όταν το σταθμισμένο άθροισμα των εισροών εφαρμόζεται στο $\tanh(x)$, αναδιπλούμε τις τιμές μεταξύ -1 και 1. Οι μεγάλοι αρνητικοί αριθμοί κλιμακώνονται προς -1 και οι μεγάλοι θετικοί αριθμοί κλιμακώνονται προς το 1.

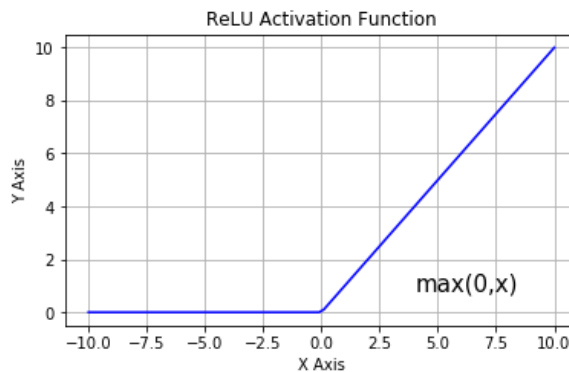
- *Rectified Linear Unit (ReLU)*: Αυτή είναι μια από τις πιο ευρέως χρησιμοποιούμενες συναρτήσεις ενεργοποίησης. Τα οφέλη της ReLU είναι η χαμηλή πυκνότητα (sparsity) και το γεγονός ότι επιτρέπει μόνο τιμές που είναι θετικές, ενώ οι αρνητικές τιμές που δεν θα περάσουν που θα επιταχύνουν τη διαδικασία και θα αναιρούν ή θα μειώνουν την πιθανότητα εμφάνισης νεκρού νευρώνα. Η μαθηματική περιγραφή αυτής της λειτουργίας παρουσιάζεται παρακάτω:

$$f(x) = (0, \max) \quad (2.6)$$



Σχήμα 2.11: Υπερβολική λειτουργία εφαπτομένης

Αυτή η λειτουργία θα επιτρέψει μόνο τις μέγιστες τιμές που θα περάσουν κατά τη διάρκεια της εμπρόσθιας διάδοσης, όπως φαίνεται στην εικόνα 2.12. Τα μειονεκτήματα της ReLU είναι ότι όταν η κλίση τείνει στο μηδέν για τις αρνητικές τιμές, δεν υπάρχει σύγκλιση προς το τοπικό ελάχιστο οδηγώντας έτσι σε ένα νεκρό νευρώνα κατά την οπίσθια διάδοση (back propagation).



Σχήμα 2.12: Λειτουργία ReLU

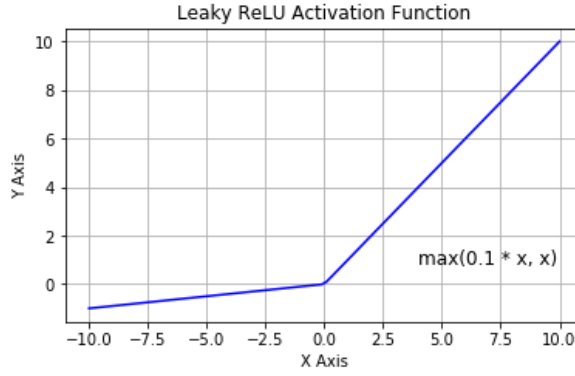
- *Leaky ReLU*: Αυτή η συνάρτηση ξεπερνά το μειονέκτημα της ReLU επιτρέποντας μια μικρή αρνητική τιμή κατά τη διάρκεια της οπίσθιας διάδοσης σε περίπτωση ενός νεκρού προβλήματος ReLU (dead ReLU problem). Αυτό τελικά θα ενεργοποιήσει τον νευρώνα και θα το μειώσει. Αυτή η λειτουργία περιγράφεται από:

$$f(x) = 1(x < 0)(\alpha x) + 1(x \geq 0)(x) \quad (2.7)$$

όπου α είναι μια μικρή σταθερά

Η απεικόνιση αυτής της συνάρτησης παρουσιάζεται στο Σχήμα 2.13. Αυτή η συνάρτηση ενεργοποίησης έχει επίσης μειονεκτήματα, κατά την εμπρόσθια διάδοση αν ο ρυθμός εκμάθησης είναι πολύ ψηλός, θα υπερβεί ένα όριο με συνέπεια τη θανάτωση του νευρώνα. Αυτό θα συμβεί όταν ο ρυθμός εκμάθησης δεν έχει ρυθμιστεί στο βέλτιστο επίπεδο.

Συνάρτηση κόστους (Cost function). Η συνάρτηση κόστους αξιολογεί την απόδοση του μοντέλου. Συγκεκριμένα, υπολογίζει πόσο καλά ένα σύνολο παραμέτρων W μπορεί να υπολογίσει τις πραγματικές τιμές σε ένα σύνολο δεδομένων. Το είδος της χρησιμοποιούμενης συνάρτησης κόστους εξαρτάται από τη φύση του προβλήματος που πρέπει να επιλυθεί.



Σχήμα 2.13: Λειτουργία ReLU διαρροής

Στα ANNs, η συνάρτηση ενεργοποίησης *softmax* εφαρμόζεται συνήθως στην έξοδο του τελευταίου στρώματος. Η συνάρτηση softmax περιγράφεται από την ακόλουθη εξίσωση:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad (2.8)$$

Αυτή η συνάρτηση συμπιέζει τις τιμές ενός K -διάστατου διανύσματος z στο διάστημα μεταξύ 0 και 1 που στοχεύει στο άθροισμα όλων των τιμών στο 1. Εφαρμόζοντας αυτή τη συνάρτηση στις τιμές ενός διανύσματος, οι τιμές όλων των κατηγοριών ενός προβλήματος κανονικοποιούνται προκειμένου για να ανταποκριθούν στις αντίστοιχες πιθανότητες.

Η πιο συχνά χρησιμοποιούμενη συνάρτηση κόστους είναι η απώλεια διατροφικής εντροπίας (cross entropy loss), που παρουσιάζεται στην Εξίσωση 2.9. Αυτή η συνάρτηση χρησιμοποιείται σε ένα ANN όπως περιγράφεται στην Εξίσωση 2.10

$$L_i = -\log\left(\frac{e^{y_i}}{\sum_j e^{f_j}}\right) \quad (2.9)$$

$$L(W) = -\frac{1}{n} \sum_{i=1}^N \sum_{j=1}^K y_{ij} \log(p_{ij}) \quad (2.10)$$

όπου W είναι τα βάρη του δικτύου που χρειάζονται εκτίμηση, i είναι οι παρατηρήσεις που τροφοδοτούνται στο δίκτυο, N είναι ο συνολικός αριθμός παρατηρήσεων που περιέχονται στο σύνολο δεδομένων, j είναι οι κλάσεις του προβλήματος ταξινόμησης, K είναι ο συνολικός αριθμός των κατηγοριών και y είναι η εκτιμώμενη τιμή για την παρατήρηση i .

Αυτή η λειτουργία υπολογίζει την απόσταση μεταξύ δύο κατανομών πιθανοτήτων. Όταν χρησιμοποιούμε ένα ANN για να λύσουμε ένα πρόβλημα ταξινόμησης, η έξοδος του δικτύου για μια ενιαία παρατήρηση είναι μια κατανομή πιθανότητας σε όλες τις κατηγορίες του προβλήματος. Αυτή η κατανομή στη συνέχεια συγκρίνεται με το one-hot διάνυσμα που αντιπροσωπεύει την πραγματική κατηγορία της παρατήρησης.

Κανονικοποίηση (Regularization). Η κανονικοποίηση αποτελεί βασικό στοιχείο για την αποτροπή του overfitting. Επίσης, ορισμένες τεχνικές ρύθμισης μπορούν να χρησιμοποιηθούν για τη μείωση της χωρητικότητας του μοντέλου (model capacity) διατηρώντας παράλληλα την ακρίβεια, για παράδειγμα, για να οδηγήσουν μερικές από τις παραμέτρους στο μηδέν. Αυτό μπορεί να είναι επιθυμητό για να μειωθεί το μέγεθος του μοντέλου ή να μειωθεί το κόστος αξιολόγησης σε κινητό περιβάλλον όπου η ισχύς του επεξεργαστή περιορίζεται.

Οι συνήθεις μέθοδοι κανονικοποίησης που χρησιμοποιούνται είναι:

- *Έγκαιρη διακοπή (Early-stopping)*: Αντιμετωπίζει το πρόβλημα του overfitting το οποίο διακόπτει την διαδικασία εκπαίδευσης, όταν χειροτερέψει η απόδοση του μοντέλου σε ένα σετ επικύρωσης (validation set). Ένα σετ επικύρωσης είναι ένα σύνολο παραδειγμάτων που δεν χρησιμοποιούμε ποτέ για οπίσθια διάδοση, αλλά δεν είναι επίσης μέρος του δοκιμαστικού συνόλου (test set). Τα παραδείγματα επικύρωσης θεωρούνται αντιπροσωπευτικά των μελλοντικών παραδειγμάτων δοκιμών. Η έγκαιρη διακοπή ρυθμίζει αποτελεσματικά τον αριθμό υπερπαραμέτρων εποχών / βημάτων. Διαισθητικά, καθώς το μοντέλο βλέπει περισσότερα δεδομένα και μαθαίνει πρότυπα και συσχετίσεις, τόσο η εκπαίδευση όσο και το σφάλμα δοκιμής μειώνονται. Μετά από αρκετά πέρασματα πάνω από τα δεδομένα εκπαίδευσης, το μοντέλο μπορεί να ξεκινήσει να παρουσιάζει overfitting και να μαθαίνει τον θόρυβο των δεδομένων του σετ εκπαίδευσης. Σε αυτή την περίπτωση, το σφάλμα της εκπαίδευσης θα συνεχίσει να μειώνεται ενώ το σφάλμα δοκιμής (πόσο καλά γενικεύουμε) θα επιδεινωθεί. Η έγκαιρη διακοπή αφορά την εύρεση της σωστής στιγμής με το ελάχιστο σφάλμα δοκιμής.
- *Dropout* [111]: Σε κάθε επανάληψη εκπαίδευσης ένα στρώμα dropout αφαιρεί τυχαία ορισμένους κόμβους στο δίκτυο μαζί με όλες τις εισερχόμενες και εξερχόμενες συνδέσεις. Η απόρριψη νευρώνων μπορεί να εφαρμοστεί σε κρυφό ή σε στρώμα εισόδου. Το dropout θεωρείται ως καλή μέθοδος κανονικοποίησης για τους ακόλουθους λόγους. Οι κόμβοι γίνονται πιο ευαίσθητοι στα βάρη των άλλων κόμβων (συν-προσαρμοστικό) και επομένως το μοντέλο είναι πιο ανθεκτικό. Εάν μια κρυφή μονάδα πρέπει να λειτουργεί καλά με διαφορετικούς συνδυασμούς άλλων κρυφών μονάδων, είναι πιο πιθανό να κάνει κάτι μεμονωμένα χρήσιμο. Επιπλέον, το dropout μπορεί να θεωρηθεί ως μια μορφή ensembling πολλαπλών μοντέλων ("σύνολο"), η οποία δείχνει καλύτερη απόδοση στις περισσότερες εργασίες εκμάθησης μηχανών (π.χ. κατάρτιση σε σύνολα είναι η διαίσθηση πίσω από τυχαία δάση ή δέντρα απόφασης κλίσης αύξησης κλίσης).
- *Ποινές βάρους (Weight penalty)*: Είναι συνήθης τρόπος για την κανονικοποίηση. Βασίζεται έντονα στην έμμεση υπόθεση ότι ένα μοντέλο με μικρά βάρη είναι κατά κάποιο τρόπο απλούστερο από ένα δίκτυο με μεγάλα βάρη. Οι ποινές προσπαθούν να κρατήσουν τα βάρη μικρά ή ανύπαρκτα (μηδέν) εκτός αν υπάρχουν μεγάλες κλίσεις για να το εξουδετερώσουν, γεγονός που καθιστά τα μοντέλα επίσης πιο ερμηνευτικά. Μια εναλλακτική ονομασία στη βιβλιογραφία για τις ποινές βάρους είναι η "weight decay" καθώς αναγκάζει τα βάρη να αποσυντεθούν προς το μηδέν.

- *L2 norm*: Η L2 νόρμα τιμωρεί την τετραγωνική τιμή του βάρους. Αυτή η τεχνική κανονικοποίησης τείνει να οδηγεί όλα τα βάρη σε μικρότερες τιμές.

$$L2 = \sum_{i=0}^N w_i^2 \quad (2.11)$$

- *L1 norm*: τιμωρεί την απόλυτη τιμή του βάρους (είναι μια λειτουργία σχήματος V). Αυτή η τεχνική τείνει να οδηγήσει μερικά βάρη ακριβώς στο μηδέν (εισάγοντας χαμηλή πυκνότητα (sparsity) στο μοντέλο), επιτρέποντας παράλληλα κάποια βάρη να είναι μεγάλα.

$$L1 = \sum_{i=0}^N |w_i| \quad (2.12)$$

Βελτιστοποίηση (Optimization). Οι αλγόριθμοι βελτιστοποίησης μας βοηθούν να ελαχιστοποιήσουμε (ή να μεγιστοποιήσουμε) μια συνάρτηση κόστους. Ο Gradient Descent (GD) είναι ένας από

τους πιο δημοφιλείς αλγόριθμους για την εκτέλεση βελτιστοποίησης και μακράν ο πιο συνηθισμένος τρόπος βελτιστοποίησης των νευρωνικών δικτύων. Ο GD είναι ένας τρόπος ελαχιστοποίησης μιας αντικειμενικής συνάρτησης $J(\theta)$. Το $J(\theta)$ παραμετροποιείται από τις παραμέτρους του μοντέλου με την ενημέρωση των παραμέτρων προς την αντίθετη κατεύθυνση της κλίσης της αντικειμενικής συνάρτησης $\nabla_{\theta} J(\theta)$ σε σχέση με τις παραμέτρους. Ο ρυθμός εκμάθησης η καθορίζει το μέγεθος των βημάτων που παίρνουμε για να φθάσουμε σε ένα (τοπικό) ελάχιστο. Με άλλα λόγια, ακολουθούμε την κατεύθυνση της κλίσης της επιφάνειας που δημιουργείται από την αντικειμενική συνάρτηση προς τα κάτω μέχρι να φτάσουμε στην κοιλάδα. Ο απλός αλγόριθμος GD υπολογίζει την κλίση της συνάρτησης κόστους σε σχέση με τις παραμέτρους θ για ολόκληρο το σύνολο δεδομένων εκπαίδευσης:

$$\theta = \theta - \nabla_{\theta} J(\theta) \quad (2.13)$$

Αντίθετα, ο Stochastic Gradient Descent (SGD) εκτελεί μια ενημέρωση παραμέτρων για κάθε παράδειγμα εκπαίδευσης x_i και ετικέτα y_i :

$$\theta = \theta - \nabla_{\theta} J(\theta; x_i; y_i) \quad (2.14)$$

Ο GD εκτελεί πλεονάζοντες υπολογισμούς για μεγάλα σύνολα δεδομένων, καθώς επαναπροσδιορίζει κλίσεις για παρόμοια παραδείγματα πριν από κάθε ενημέρωση παραμέτρων. Ο SGD απομακρύνει αυτή την πλεοναστικότητα εκτελώντας μια ενημέρωση τη φορά. Είναι επομένως συνήθως πολύ ταχύτερος και μπορεί επίσης να χρησιμοποιηθεί για να μάθει online. Εκτός από τον SGD υπάρχουν και άλλοι αλγόριθμοι για βελτιστοποίηση όπως η μέθοδος Nesterov Momentum [112]. Αυτή τη στιγμή χρησιμοποιούνται μέθοδοι βελτιστοποίησης με αυτόματη ρύθμιση του ρυθμού εκμάθησης, όπως Adagrad [113], Adadelta [114] και Adam [115], η οποία είναι η ευρέως χρησιμοποιούμενη τεχνική βελτιστοποίησης.

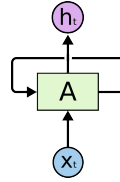
Οπίσθια διάδοση (Backpropagation). Ένα ANN μπορεί να εκπροσωπείται από έναν κατευθυνόμενο γράφο, όπου κάθε κόμβος αντιστοιχεί σε ένα συγκεκριμένο βάρος του δικτύου. Για να ενημερωθούν τα βάρη του δικτύου μετά από κάθε υπολογισμό της συνάρτησης κόστους χρησιμοποιείται ο αλγόριθμος οπίσθιας διάδοσης (backpropagation) [116]. Το βασικό στοιχείο του backpropagation είναι μια έκφραση για την μερικό παράγωγο $\frac{\partial C}{\partial w}$ της συνάρτησης κόστους C σε σχέση με οποιοδήποτε βάρος w (ή μεροληψία b) στο δίκτυο. Η έκφραση υποδεικνύει πόσο γρήγορα αλλάζει το κόστος όταν αλλάζουν τα βάρη και τα biases. Τα πλεονεκτήματα του backpropagation δεν περιορίζονται στον γρήγορο υπολογισμό, αλλά ο αλγόριθμος αυτός παρέχει επίσης μια διαισθητική ερμηνεία για το πώς η αλλαγή των βαρών και των biases μεταβάλλει τη συνολική συμπεριφορά του δικτύου.

2.4.3 Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks - RNNs)

Ένα RNN είναι μια ειδική περίπτωση των ANNs όπου οι συνδέσεις μεταξύ των μονάδων σχηματίζουν ένα κατευθυνόμενο γράφο κατά μήκος μιας ακολουθίας. Αυτό του επιτρέπει να παρουσιάζει δυναμική χρονική συμπεριφορά για μια χρονική ακολουθία. Σε αντίθεση με τα FFNNs, τα RNNs μπορούν να χρησιμοποιήσουν την εσωτερική τους κατάσταση (μνήμη) για να επεξεργαστούν ακολουθίες εισόδων. Για να το επιτύχει αυτό, το RNN δημιουργεί βρόχους στο δίκτυο, γεγονός που του επιτρέπει να διατηρεί τις πληροφορίες. Αυτή η δομή βρόχου επιτρέπει στο νευρωνικό δίκτυο να πάρει την ακολουθία εισόδου, όπως παρουσιάζεται στο Σχήμα 2.14⁴.

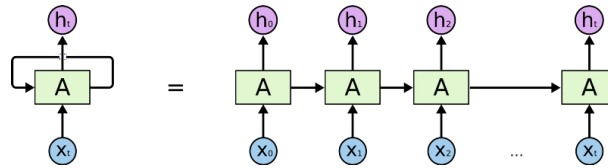
Η λειτουργία ενός RNN μπορεί να γίνει καλύτερα κατανοητή στο Σχήμα 2.15. Πρώτον, παίρνει το x_0 από την ακολουθία εισόδου και στη συνέχεια εξάγει h_0 (κρυφή κατάσταση του RNN) που μαζί

⁴ <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>



Σχήμα 2.14: Βρόχοι σε RNNs

με το x_1 είναι η είσοδος για το επόμενο βήμα. Έτσι, τα h_0 και x_1 είναι η είσοδος για το επόμενο βήμα. Ομοίως, το h_1 από το επόμενο είναι η είσοδος μαζί με το x_2 για το επόμενο βήμα και ούτω καθεξής. Με αυτό τον τρόπο, θυμάται τα συμφοραζόμενα κατά την εκπαίδευση.



Σχήμα 2.15: Ξετυλιγμένη έκδοση ενός RNN

Επομένως, για κάθε βήμα χρόνου t οι εξισώσεις που περιγράφουν τη λειτουργία του RNN είναι:

$$h_t = f_h(W_h x_t + U_h h_{t-1} + b_h) \quad (2.15)$$

$$y_t = f_y(W_y h_t + b_y) \quad (2.16)$$

όπου

- h_t : η κρυφή κατάσταση την χρονική στιγμή t
- x_t : το διάνυσμα της ακολουθίας εισόδου την χρονική στιγμή t
- y_t : η έξοδος του RNN την χρονική στιγμή t
- W_x, U_h, W_y : τα βάρη του δικτύου για h, x, y , αντίστοιχα
- b_h : bias για το h
- f_x, f_h : συναρτήσεις ενεργοποίησης για τα x και h , αντίστοιχα

Αμφίδρομο RNN

Ένα αμφίδρομο RNN (Bidirectional RNN - BiRNN) αποτελείται από το συνδυασμό δύο διαφορετικών RNNs, όπου κάθε ένα επεξεργάζεται την ακολουθία εισόδου με διαφορετική κατεύθυνση. Το κίνητρο για τη χρήση των BiRNNs είναι να παραχθεί μια πιο ακριβής αναπαράσταση της ακολουθίας εισόδου. Για το σκοπό αυτό, υπάρχει ένα δεξιά κατευθυνόμενο RNN το οποίο επεξεργάζεται προοδευτικά την ακολουθία ξεκινώντας από το διάνυσμα x_0 της εισόδου και ένα αριστερά κατευθυνόμενο RNN το οποίο ξεκινά την επεξεργασία της ακολουθίας από το διάνυσμα x_T , όπου T είναι το μήκος της ακολουθίας εισόδου. Επομένως, σε κάθε βήμα t , οι εξισώσεις που περιγράφουν ένα BiRNN είναι:

$$h_i = \vec{h}_i || \bar{h}_i, h_i \in \mathcal{R}^{2N} \quad (2.17)$$

όπου $||$ αντιπροσωπεύει τη συνένωση των δύο διανυσμάτων και N είναι ο αριθμός των διαστάσεων κάθε RNN.

Μηχανισμός προσοχής (Attention mechanism)

Όλοι οι φορείς που συνιστούν την ακολουθία εισόδου συμβάλλουν εξίσου στην έννοια που εκφράζεται στη συνολική είσοδο. Για το λόγο αυτό, μπορεί να χρησιμοποιηθεί ένας μηχανισμός προσοχής (attention mechanism) [117, 118] για να βρεθεί η σχετική συμβολή (σημασία) κάθε διανύσματος εισόδου μιας ακολουθίας. Ο μηχανισμός προσοχής αποδίδει ένα βάρος a_i σε κάθε επισημασμένο διάνυσμα h_i . Στην συνέχεια, υπολογίζεται η σταθερή αναπαράσταση r ολόκληρης της εισόδου ως το σταθμισμένο άθροισμα όλων των επισημασμένων λέξεων.

$$e_i = \tanh(W_h h_i + b_h), \quad e_i \in [-1, 1] \quad (2.18)$$

$$a_i = \frac{\exp(e_i)}{\sum_{t=1}^T \exp(e_t)}, \quad \sum_{i=1}^T a_i = 1 \quad (2.19)$$

$$r = \sum_{i=1}^T a_i h_i, \quad r \in R^{2L} \quad (2.20)$$

όπου τα W_h και b_h είναι τα βάρη του στρώματος προσοχής.

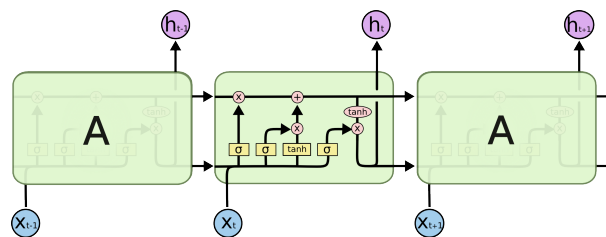
Δίκτυα μακράς βραχυπρόθεσμης μνήμης (Long Short-Term Memory (LSTM))

Υπάρχουν δύο μεγάλα εμπόδια που τα RNN έχουν ή πρέπει να αντιμετωπίσουν:

- *Exploding Gradients*: Το πρόβλημα όταν ο αλγόριθμος εκχωρεί πολύ μεγάλα βάρη στο δίκτυο.
- *Vanishing Gradients*: Το πρόβλημα όταν οι τιμές μιας κλίσης είναι πολύ μικρές και το μοντέλο σταματά να μαθαίνει ή αργεί πάρα πολύ η διαδικασία εκπαίδευσης εξαιτίας αυτού.

Τα LSTMs που προτάθηκαν από [119] μπορούν να ξεπεράσουν αυτά τα προβλήματα. Το LSTM είναι μια παραλλαγή του RNN που έχει το πλεονέκτημα της διατήρησης των συνδέσεων μεγάλων αποστάσεων μεταξύ των λέξεων και την απόσταση ασήμαντων λέξεων από την πύλη του κυττάρου (cell gate) μέσω του στρώματος της πύλης λήθης (forget gate layer).

Συγκεκριμένα, σε ένα RNN υπάρχουν τρεις πύλες: είσοδος (input gate) i , λήθη (forget gate) f και έξοδος (output gate) o . Αυτές οι πύλες καθορίζουν αν πρέπει να αφήσουν ή όχι νέα είσοδο (input gate), να διαγράψουν τις πληροφορίες επειδή δεν είναι σημαντικές (forget gate) ή να την αφήσουν να επηρεάσει την έξοδο στο τρέχον βήμα (output gate). Οι πύλες σε ένα LSTM είναι αναλογικές, με τη μορφή sigmoids, που σημαίνει ότι κυμαίνονται από 0 έως 1. Το γεγονός ότι είναι αναλογικές, τους επιτρέπει να κάνουν backpropagation με αυτό. Τα προβληματικά ζητήματα των vanishing gradients λύνονται μέσω του LSTM επειδή διατηρεί τις κλίσεις αρκετά απότομες και συνεπώς η εκπαίδευση είναι σχετικά μικρή και η ακρίβεια υψηλή. Η δομή ενός κυττάρου LSTM απεικονίζεται στο σχήμα 2.16.



Σχήμα 2.16: Δομή ενός κυττάρου LSTM

Συγκεκριμένα, δεδομένης μιας ακολουθίας $x_1, x_2, \dots, x_t, \dots, x_l$ των διανυσμάτων για μια ακολουθία εισόδου μήκους l , για το t^{th} διάνυσμα x_t , με τις εισόδους h_{t-1} και c_{t-1} , h_t και c_t υπολογίζονται ως εξής:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \quad (2.21)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \quad (2.22)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \quad (2.23)$$

$$u_t = \tanh(W_u x_t + U_u h_{t-1} + b_u), \quad (2.24)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot u_t, \quad (2.25)$$

$$h_t = o_t \odot \tanh(c_t), \quad (2.26)$$

που

- $W_j \in \mathbb{R}^{d \times d}, U_j \in \mathbb{R}^{d \times m}$ for $j \in \{i, f, o, u\}$ είναι οι μήτρες βαρών
- $b_j \in \mathbb{R}^d$ είναι τα bias διανύσματα
- $\sigma(\cdot)$ είναι η σιγμοειδής συνάρτηση
- $\tanh(\cdot)$ είναι η συνάρτηση υπερβολικής εφαπτομένης
- \odot είναι το εσωτερικό γινόμενο

2.4.4 Μεταφορά Μάθησης (Transfer Learning - TL)

Το TL έχει φέρει επανάσταση στο Computer Vision (CV), αλλά οι υπάρχουσες προσεγγίσεις στο NLP απαιτούν ακόμη αλλαγές και εκπαίδευση από την αρχή. Το TL παίζει καθοριστικό ρόλο όταν ένα συγκεκριμένο σύνολο δεδομένων έχει ανεπαρκή επισημασμένα παραδείγματα για την εκπαίδευση ενός ακριβούς μοντέλου. Σε τέτοια σενάρια, οι γνώσεις που συσσωρεύονται στο πλαίσιο ενός προ-εκπαιδευμένου μοντέλου σε ένα σύνολο δεδομένων πηγής (source task) μπορούν να μεταφερθούν σε ένα σύνολο δεδομένων στόχων (target tasks), με αποτέλεσμα τη βελτίωση του μοντέλου στόχου.

Συγκεκριμένα, υπάρχουν δύο διαφορετικά σενάρια όπου το TL μπορεί να εφαρμοστεί επιτυχώς:

- Μεταφορά γνώσης σε ένα σημασιολογικά παρόμοιο / ίδιο πρόβλημα αλλά με διαφορετικό σύνολο δεδομένων.
 - Source task (S) - Ένα μεγάλο σύνολο δεδομένων για ένα συγκεκριμένο πρόβλημα
 - Target task (T) - Ένα μικρό σύνολο δεδομένων για το ίδιο πρόβλημα
- Η μεταφορά γνώσης σε ένα πρόβλημα που είναι σημασιολογικά διαφορετικό αλλά μοιράζεται την ίδια αρχιτεκτονική νευρωνικού δικτύου έτσι ώστε να μπορούν να μεταφερθούν οι νευρωνικές παράμετροι.
 - Source task (S) - Ένα μεγάλο σύνολο δεδομένων για ένα συγκεκριμένο πρόβλημα

- Target task (T) - Ένα μικρό σύνολο δεδομένων για ένα σημασιολογικά διαφορετικό πρόβλημα

Παραδοσιακά, μόνο το πρώτο στρώμα οποιουδήποτε βαθιού νευρωνικού δικτύου είναι προ-εκπαιδευμένο (βλ. Ενότητα 2.3.6). Η λεπτομερή ρύθμιση (fine-tuning) των προ-εκπαιδευμένων word embeddings [120], μια απλή τεχνική TL που στοχεύει μόνο το πρώτο στρώμα ενός νευρωνικού δικτύου, έχει μεγάλη επίδραση στην πράξη.

Πρόσφατα προτάθηκαν μέθοδοι που υπερβαίνουν τη μεταφορά λέξεων. Υπάρχουν τρεις διαφορετικές προσεγγίσεις που χρησιμοποιούνται από την άποψη του TL:

- **Προεκπαιδευμένα χαρακτηριστικά (Pre-trained features):** Τα τελικά πλήρως συνδεδεμένα στρώματα των βαθιών νευρωνικών δικτύων γενικά υποτίθεται ότι συλλαμβάνουν πληροφορίες σχετικές με την επίλυση του προβλήματος (εδώ στο source task). Για ένα νέο πρόβλημα, μπορούμε έτσι να χρησιμοποιήσουμε απλώς τα έτοιμα χαρακτηριστικά που εξάγονται από το source task και να τα συνενώσουμε με τα χαρακτηριστικά του τελικού πλήρως συνδεδεμένου στρώματος για το target task. Σε αυτήν την περίπτωση, τόσο τα χαρακτηριστικά από την πηγή όσο και το σύνολο δεδομένων στόχου θα συμβάλλουν στην τελική πρόβλεψη στο σύνολο δεδομένων στόχων.
- **Αρχικοποίηση παραμέτρων (Parameter Initialization):** Η προσέγγιση INIT (Initialization of Parameter Initialization) αρχίζει να εκπαιδεύει το δίκτυο στο source task και στη συνέχεια χρησιμοποιεί απευθείας τις συντονισμένες παραμέτρους για την προετοιμασία του δικτύου για το target task. Μετά τη μεταφορά, μπορούμε να διορθώσουμε τις παραμέτρους στον τομέα-στόχο, δηλαδή να ρυθμίσουμε με ακρίβεια τις παραμέτρους του target task.
- **Μάθηση πολλαπλών εργασιών (Multi-task learning):** Η μάθηση πολλαπλών εργασιών (MULTI-Task (MULT)), από την άλλη πλευρά, ταυτόχρονα εκπαιδεύει δείγματα και στους δύο τομείς.

Κεφάλαιο 3

Αναπαράσταση και Ταξινόμηση Διαλογικών Πράξεων Χρησιμοποιώντας Βαθιά Νευρωνικά Δίκτυα

3.1 Εισαγωγή

Η αναγνώριση της πρόθεσης ενός ομιλητή στο πλαίσιο του διαλόγου είναι σχετικά κοντά στην έννοια της αναγνώρισης της διαλογικής πράξης (DA). Οι DAs μπορούν να ερμηνευτούν ως συγκεκριμένες πράξεις ομιλίας με μια συγκεκριμένη επικοινωνιακή λειτουργία και σημασιολογικό περιεχόμενο [121]. Το σημασιολογικό περιεχόμενο προσδιορίζει τα αντικείμενα, τις προτάσεις, τα γεγονότα κ.λπ. που αφορούν την DA. Η επικοινωνιακή λειτουργία προσδιορίζει τον τρόπο με τον οποίο ο παραλήπτης πρέπει να χρησιμοποιήσει το σημασιολογικό περιεχόμενο για να ενημερώσει την κατάσταση πληροφοριών του.

Η έννοια της “πρόθεσης (intent)” θεωρείται ως μια διάσταση των DAs. Με άλλα λόγια, μια πρόθεση (διάσταση) είναι μια τάξη διαλογικών πράξεων που ασχολούνται με μια συγκεκριμένη πτυχή της επικοινωνίας που μια DA μπορεί να αντιμετωπίσει ανεξάρτητα από άλλες διαστάσεις [122].

Η ταξινόμηση των DAs αποτελεί ένα σημαντικό βήμα επεξεργασίας στα Συστήματα Ομιλούμενου Διαλόγου (Spoken Dialogue Systems - SDS) που βοηθούν στην κατανόηση της εισόδου του χρήστη. Συνήθως, αυτό εφαρμόζεται μέσω της εκχώρησης ετικετών στις φράσεις των χρηστών που περιγράφουν (λεξικολογικά) τις αντίστοιχες πράξεις. Οι DAs μπορούν να θεωρηθούν ως οι ελάχιστες μονάδες γλωσσικής επικοινωνίας που συνδέονται άμεσα με τις επικοινωνιακές προθέσεις του ομιλητή [123]. Η έξοδος της ταξινόμησης DA μπορεί να αξιοποιηθεί από άλλα συστήματα SDS, συμπεριλαμβανομένων των συστημάτων κατανόησης της φυσικής γλώσσας και διαχείρισης του διαλόγου.

Διάφορες προσεγγίσεις έχουν χρησιμοποιηθεί για ταξινόμηση DA συμπεριλαμβανομένων των Bayesian Networks (BN), HMMs [124], Feed-Forward Neural Networks (FFNN) [125], Δένδρων Απόφασης (decision trees) [43] και SVMs [126]. Η πλειοψηφία αυτών των προσεγγίσεων εξέτασε τόσο την σημασία της φράσης του ομιλητή όσο και την σειρά των διαδοχικών φράσεων στο πλαίσιο του διαλόγου. Πρόσφατα, τα Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks - DNNs) έχουν χρησιμοποιηθεί για την ταξινόμηση DA [127, 128, 129, 125] παρέχοντας μια σημαντική αύξηση στην ακρίβεια ταξινόμησης του προβλήματος ανεξαρτήτως της φύσεως των συνομιλιών.

Μια πρόκληση στον τομέα της ταξινόμησης DA είναι η κατασκευή μοντέλων που να είναι ανεξάρτητα του εξεταζόμενου τομέα και να παρουσιάζουν καλή απόδοση σε διαφορετικές ταξινομίες (coarse- vs. fine-grained) ετικετών DA. Στις πρόσφατες προσεγγίσεις βαθιάς μάθησης (π.χ. [127, 129, 128]) Τα DNNs βασίζονται σε word embeddings που είναι γενικά ή τυχαία ορισμένα, αγνοώντας τη συγκεκριμένη τοπολογία στο εξεταζόμενο πρόβλημα. Στο [128], διερευνήθηκε η απόδοση των συστημάτων DA χρησιμοποιώντας διαφορετικά συστήματα γενικών word embeddings και αποδείχθηκε ότι η απόδοση εξαρτάται από την ευκρίνεια των ετικετών DA.

Στο πλαίσιο αυτής της διατριβής, εξετάζουμε την ενσωμάτωση της σημασιολογίας των DAs στο πλαίσιο των RNNs. Συγκεκριμένα, προτείνουμε ένα νέο σχήμα για την αυτόματη κωδικοποίηση της σημασιολογίας των DAs μέσω της εξαγωγής ενός συνόλου σημασιολογικά χαρακτηριστικών λέξεων-

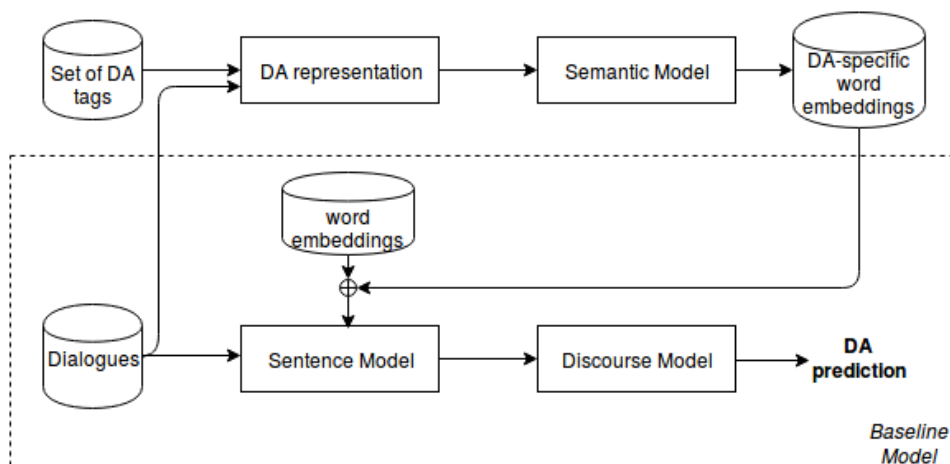
κλειδιών. Αυτές οι λέξεις-κλειδιά μπορούν να θεωρηθούν ως μέλη των σημασιολογικών υπο-χώρων που αντιστοιχούν στην αντίστοιχη ετικέτα DA. Η σημασία αυτών των λέξεων-κλειδιών είναι σχετική με κάθε DA και υπολογίζεται από ένα μοντέλο παλινδρόμησης που εκμεταλλεύεται τα γενικά word embeddings. Η ταξινόμηση μιας άγνωστης φράσης βασίζεται στον υπολογισμό της σημασιολογικής ομοιότητας μεταξύ των επιμέρους λέξεων και των προαναφερθέντων υποπεριοχών DA, τα οποία δίδονται ως χαρακτηριστικά στο χρησιμοποιούμενο DNN εκτός από τα τυπικά word embeddings.

3.2 Επισκόπηση Βιβλιογραφίας

Οι πρώιμες προσεγγίσεις της ταξινόμησης DA εκμεταλλεύτηκαν τις λεξικές πληροφορίες, τη σύνταξη, τη σημασιολογία, την προσοδία και το ιστορικό του διαλόγου με χειροκίνητη εξαγωγή των χαρακτηριστικών [130, 124, 131, 132, 133, 134]. Οι [130] κατασκεύασαν ταξινομητές DA σε δεδομένα μηνυμάτων με χρήση λεξικών, συντακτικών και σημασιολογικών χαρακτηριστικών, δημιουργώντας σταθερά λεξικά με συγκεκριμένες θεματικές ενότητες με λέξεις-κλειδιά. Οι [124] εκμεταλλεύτηκαν λεξικολογικά και προσοδικά χαρακτηριστικά, εξαγόμενα από τους διαλόγους, σε συνδυασμό με πληροφορίες σχετικές με την ακολουθία των DAs. Το μοντέλο που προτάθηκε είναι ένα Hidden Markov Model (HMM), όπου κάθε κατάσταση του αντιστοιχεί σε μια διαδοχική DA, επιτυγχάνοντας ακρίβεια ταξινόμησης 71.0% όταν εφαρμόζεται στο corpus του Switchboard-DAMSL [42]. Οι [134] εξέτασαν το ρόλο της συναισθηματικής ανάλυσης μέσω των συναισθηματικών λεξικών στην αναγνώριση των DA. Όσον αφορά την ανάλυση συναισθηματικών κειμένων, τα σημασιολογικά χαρακτηριστικά εξάγονται με βάση τα καταναμημένα σημασιολογικά μοντέλα που κατασκευάστηκαν από τους [135].

Πρόσφατα, η εξέλιξη της βαθιάς μάθησης επέτρεψε την εφαρμογή διαφορετικών μοντέλων DNN στα προβλήματα NLP, συμπεριλαμβανομένης της ταξινόμησης DA. Ο [127] χρησιμοποίησε ένα μείγμα από Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks - CNNs) ως προτεινόμενο μοντέλο για την εξαγωγή χαρακτηριστικών από κάθε φράση και Αναδρομικών Νευρωνικών Δικτύων (Recurrent Neural Networks - RNNs) ως προτεινόμενο μοντέλο για την εξαγωγή πληροφοριών σχετικά με την ακολουθία των DAs. Το μοντέλο αυτό βελτίωσε την απόδοση στην ταξινόμηση DA στην βάση δεδομένων Switchboard-DAMSL, φτάνοντας σε ακρίβεια 73.9%. Ο [128] δημιούργησε ένα μοντέλο βασισμένο στα RNN και CNN το οποίο ενσωματώνει τις προηγούμενες δηλώσεις μέσω ενός ANN με δύο κρυμμένα στρώματα για την εξαγωγή πληροφοριών σχετικών με την ιστορία του διαλόγου. Ο [125] πρότεινε μια υβριδική αρχιτεκτονική που συνδυάζει ένα μοντέλο RNN για την εξαγωγή πληροφοριών σχετικών με μια συγκεκριμένη φράση με πληροφορίες για τη σχέση μεταξύ δύο διαδοχικών φράσεων υπό τη μορφή λανθάνουσας μεταβλητής. Όταν μεγιστοποιείται η πιθανότητα των σχέσεων της διάρθρωσης του διαλόγου, που προέρχονται από το μοντέλο, αντιμετωπίζοντας το μοντέλο το σχετικό με την συγκεκριμένη φράση ως παράπλευρο παράγοντα στην ταξινόμηση DA, επιτυγχάνεται ακρίβεια 77.0%. Ο [129] χρησιμοποίησε μια δομή LSTM [119] με προεκπαιδευμένα word embeddings και ανέφερε ακρίβεια ταξινόμησης 80.1% που υπερβαίνει το state-of-the-art.

Για τη δοκιμή των διαφόρων μοντέλων που προτείνονται για την ταξινόμηση DA, χρησιμοποιήθηκαν διάφορα σχήματα επισήμανσης καθώς και σύνολα δεδομένων [42, 43, 136, 137]. Οι [42] παρείχαν ένα σύνολο δεδομένων επισήμασμένο με 42 ετικέτες DA, σύμφωνα με το σχήμα επισήμανσης του Dialog Act Markup in Several Layers (DAMSL) [41]. Οι [43] πρότειναν ένα σχήμα επισήμανσης πέντε κατηγοριών με βάση την βάση δεδομένων MRDA. Ωστόσο, καταβάλλονται προσπάθειες για την ανάπτυξη ενός σχήματος επισήμανσης DA που να είναι ανεξάρτητο από τα επιμέρους προβλήματα και να μπορεί να χρησιμοποιηθεί με μεθόδους αυτόματης επισήμανσης [138, 139, 140]. Παρ'όλα αυτά, υπάρχουν ακόμα περιορισμένα δεδομένα επισήμασμένα με βάση τις αρχές αυτών των σχημάτων, όπως το πρότυπο ISO 24617-2 και το DIT ++ [138, 139].



Σχήμα 3.1: Επισκόπηση του προτεινόμενου μοντέλου.

3.3 Προτεινόμενο Μοντέλο

Τα δύο μέρη που αποτελούν το προτεινόμενο μοντέλο απεικονίζονται στο Σχήμα 3.1. Το πρώτο μέρος (μοντέλο πρότασης) δημιουργεί ένα διάνυσμα αναπαράστασης της φράσης με βάση τη δομή LSTM που προτείνεται από το [141] και χρησιμοποιείται επίσης από το [129]. Το μοντέλο της πρότασης χρησιμοποιεί word embeddings για τον υπολογισμό ομοιότητας μεταξύ των συστατικών λέξεων των εκφράσεων και των ετικετών DA. Αυτό το μοντέλο περιγράφεται λεπτομερώς στην Ενότητα 3.3.1. Το δεύτερο μέρος είναι ένα μοντέλο διάρθρωσης του διαλόγου που ταξινομεί την τρέχουσα φράση με βάση το διάνυσμα αναπαράστασης του καθώς και των διανυσμάτων αναπαράστασης των προηγούμενων φράσεων όπως προτείνεται από το [128]. Το μοντέλο διάρθρωσης του διαλόγου περιγράφεται λεπτομερώς στην Ενότητα 3.3.2. Στο baseline μοντέλο προσθέτουμε τη σημασιολογική αναπαράσταση των ετικετών DA.

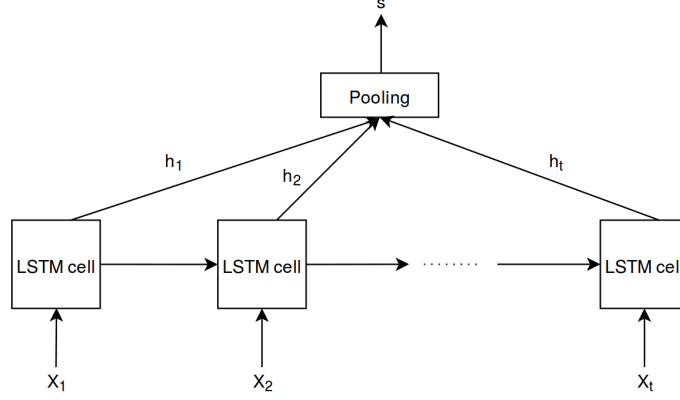
3.3.1 Μοντέλο Πρότασης

Το προτεινόμενο μοντέλο πρότασης είναι μια επέκταση του baseline μοντέλου πρότασης με σημασιολογικά χαρακτηριστικά σχετικά με τις ετικέτες DA όπως απεικονίζεται στο Σχήμα 3.1. Το baseline μοντέλο και η προτεινόμενη προσέγγιση της εξαγωγής σημασιολογικών χαρακτηριστικών περιγράφονται στη συνέχεια.

Baseline Μοντέλο Πρότασης

Το baseline μοντέλο απεικονίζεται στο Σχήμα 3.2. Δεδομένης μιας έκφρασης που περιέχει λέξεις l , το μοντέλο τη μετατρέπει σε ακολουθία l διανυσμάτων d διαστάσεων X_1, X_2, \dots, X_l . Αυτή η ακολουθία δίνεται ως είσοδος στο δίκτυο LSTM που παράγει μια αναπαράσταση s διανύσματος m διαστάσεων της έκφρασης. Το LSTM είναι μια παραλλαγή του RNN που έχει το πλεονέκτημα της διατήρησης μνήμης μέσω συνδέσεων μεταξύ των λέξεων και την απόσταση ασήμαντων λέξεων από την πύλη κυττάρου (cell gate) μέσω του στρώματος της πύλης λήθης (forget gate layer). Συγκεκριμένα, δεδομένης της ακολουθίας $X_1, X_2, \dots, X_t, \dots, X_l$ λέξεων, το διάνυσμα αναπαράστασης της t^{th} λέξης X_t , με τις εισόδους h_{t-1} και c_{t-1} , h_t και c_t υπολογίζεται ως εξής [119]:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \quad (3.1)$$



Σχήμα 3.2: Επισκόπηση του baseline μοντέλου για την δημιουργία του διανύσματος αναπαράστασης της φράσης s .

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \quad (3.2)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \quad (3.3)$$

$$u_t = \tanh(W_u x_t + U_u h_{t-1} + b_u), \quad (3.4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot u_t, \quad (3.5)$$

$$h_t = o_t \odot \tanh(c_t), \quad (3.6)$$

όπου $W_j \in \mathbb{R}^{d \times d}$, $U_j \in \mathbb{R}^{d \times m}$ για τα $j \in \{i, f, o, u\}$ είναι οι μήτρες βάρους, $b_j \in \mathbb{R}^d$ είναι διανύσματα πόλωσης και $\sigma(\cdot)$ είναι η σιγμοειδής συνάρτηση, $\tanh(\cdot)$ είναι η συνάρτηση υπερβολικής εφαπτομένης και \odot είναι το εσωτερικό γινόμενο.

Στο στρώμα συγκέντρωσης (pooling layer), όλα τα διανύσματα h_1, h_2, \dots, h_t που έχουν υπολογιστεί συνδυάζονται για τη δημιουργία ενός μόνο διανύσματος που αντιπροσωπεύει τη φράση. Ο συνδυασμός των διανυσμάτων h μπορεί να παραχθεί εφαρμόζοντας οποιοδήποτε από τα ακόλουθα σχήματα: max-pooling, mean-pooling και last-pooling. Το max-pooling διατηρεί το μέγιστο των διανυσμάτων h , το mean-pooling το μέσο των διανυσμάτων h και το last-pooling διατηρεί το τελευταίο διάνυσμα του h , δηλαδή το διάνυσμα h_t . Προκειμένου να αποκτηθούν μεγαλύτερες εξαρτήσεις μεταξύ των λέξεων της φράσης, δύο LSTM κύτταρα στοιβάζονται όπως προτείνεται από τους [142] και [143]. Επομένως, το προτεινόμενο μοντέλο έχει δύο κρυμμένα στρώματα.

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & d(k_1, w_1) \bar{s}(k_1, t_i) & \dots & d(k_N, w_1) \bar{s}(k_N, t_i) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & d(k_1, w_K) \bar{s}(k_1, t_i) & \dots & d(k_N, w_K) \bar{s}(k_N, t_i) \end{bmatrix} \cdot \begin{bmatrix} a_{i0} \\ a_{i1} \\ \vdots \\ a_{iN} \end{bmatrix} = \begin{bmatrix} 1 \\ \bar{s}(w_1, t_i) \\ \vdots \\ \bar{s}(w_K, t_i) \end{bmatrix} \quad (3.7)$$

Αναπαράσταση των DAs (DA Representation)

Τα τυπικά word embeddings που συνθέτουν την είσοδο του μοντέλου πρότασης, δεν μοντελοποιούν άμεσα τις σημασιολογικές πληροφορίες σχετικά με τη σχέση μεταξύ κάθε λέξης w και κάθε ετικέτας

DA. Εδώ παρουσιάζουμε ένα σημασιολογικό μοντέλο που εξάγει αυτόματα τη σημασιολογία του w . Συγκεκριμένα, το σημασιολογικό μοντέλο υπολογίζει τη σημασιολογική ομοιότητα μεταξύ της w και κάθε DA. Το πρώτο βήμα προς τον υπολογισμό της σημασιολογικής ομοιότητας μεταξύ w και κάθε μιας από τις ετικέτες DA είναι η επιλογή λέξεων-κλειδιών που είναι αντιπροσωπευτικές του πλαισίου των ετικετών DA όπως περιγράφεται στην επόμενη παράγραφο.

Επιλογή Λέξεων-Κλειδιών ια να προσδιορίσουμε αυτόματα τις λέξεις-κλειδιά που είναι αντιπροσωπευτικές των DA, χρησιμοποιούμε τις παρακάτω μετρήσεις:

1. Το saliency της λέξης w , που μετρά το περιεχόμενο πληροφοριών της w σε σχέση με ένα συγκεκριμένο πρόβλημα (ταξινόμηση DA στην περίπτωση αυτή), όπως προτείνεται από το [144]:

$$L(w) = \sum_{i=1}^T p(t_i|w) \log \frac{p(t_i|w)}{p(t_i)}, \quad (3.8)$$

όπου $L(w)$ είναι το saliency της w , T είναι ο αριθμός των ετικετών DA, $p(t_i|w)$ είναι η πιθανότητα της i^{th} ετικέτας DA t_i δεδομένης της w , και $p(t_i)$ είναι η πιθανότητα της i^{th} ετικέτας DA t_i ,

2. Η συχνότητα της λέξης w , επισημασμένη ως $f(w)$,
3. Η μέγιστη πιθανότητα μιας ετικέτας DA δεδομένης της λέξης w ($\max_{i=1}^T p(t_i|w)$), όπου t_i είναι η i^{th} ετικέτα DA.

Η εξαγωγή λέξεων-κλειδιών στη συνέχεια βασίζεται σε κατώτατα όρια (βλέπε Ενότητα 3.5.1) που εφαρμόζονται στο αποτέλεσμα του saliency της w και της συχνότητας ($S(w)f(w)$) και στη μέγιστη πιθανότητα των DAs δεδομένης της w ($\max_{i=1}^T p(t_i|w)$).

Σημασιολογικό Μοντέλο (Semantic Model). Μετά τον προσδιορισμό των λέξεων-κλειδιών, η σημασιολογική ομοιότητα μεταξύ της λέξης w και κάθε ετικέτας DA υπολογίζεται ως εξής:

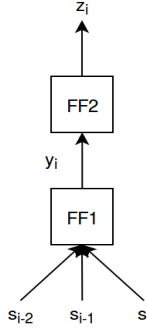
$$s(w, t_i) = \sum_{j=1}^N a_{ij} \frac{p(t_i|k_j)p(k_j)}{p(t_i)} d(k_j, w), \quad (3.9)$$

όπου $s(w, t_i)$ είναι η σημασιολογική ομοιότητα μεταξύ w και της i^{th} ετικέτας DA t_i κανονικοποιημένη στην περιοχή 0 έως 1, N είναι ο συνολικός αριθμός λέξεων-κλειδιών και a_{ij} είναι τα βάρη που αντιστοιχούν σε κάθε λέξη-κλειδί k_j για κάθε DA t_i που υπολογίζονται στο (3.7) για κάθε $i \in [1, T]$. $p(t_i|k_j)$ είναι η πιθανότητα της i^{th} ετικέτας DA t_i δεδομένης της λέξης-κλειδιού k_j , $p(t_i)$ είναι η πιθανότητα της i^{th} ετικέτας DA t_i , $p(k_j)$ είναι η πιθανότητα της λέξης-κλειδιού k_j , $\frac{p(t_i|k_j)p(k_j)}{p(t_i)} = p(k_j|t_i)$ είναι η πιθανότητα να είναι η λέξη-κλειδί k_j αντιπροσωπευτική της i^{th} ετικέτας DA t_i , κανονικοποιημένη στην περιοχή 0 έως 1 και $d(k_j, w)$ είναι η ομοιότητα μεταξύ των διανυσμάτων του w και της λέξης-κλειδιού k_j .

Στην (3.7) όπου υπολογίζονται τα βάρη a , το K είναι το μέγεθος του λεξιλογίου διαλόγου και $\bar{s}(w_k, t_i)$ είναι η εκτιμώμενη σημασιολογική ομοιότητα μεταξύ της w_k και της i^{th} ετικέτας DA t_i . Το $\bar{s}(w_k, t_i)$ υπολογίζεται εφαρμόζοντας την Εξίσωση 3.9 και θέτοντας τα βάρη a ίσα με 1.

3.3.2 Μοντέλο Διάρθρωσης του Διαλόγου (Discourse Model)

Το μοντέλο διάρθρωσης του διαλόγου απεικονίζεται στο Σχήμα 3.3. Έστω ότι το s_i είναι η διανυσματική αναπαράσταση της i^{th} φράσης του διαλόγου που υπολογίζεται από το μοντέλο πρότασης. Η ακολουθία s_{i-2}, s_{i-1}, s_i χρησιμοποιείται ως είσοδος σε ένα ANN με δύο στρώματα. Ο στόχος του



Σχήμα 3.3: Επισκόπηση του μοντέλου διάρθρωσης του διαλόγου που προβλέπει την ετικέτα DA z_i της φράσης s_i .

μοντέλου διάρθρωσης του διαλόγου είναι να προβλέψουμε τη DA της i^{th} φράσης ($z_i \in \mathbb{R}^T$). Η έξοδος του πρώτου στρώματος του ANN υπολογίζεται ως εξής:

$$y_i = \tanh\left(\sum_{d=0}^2 W_{-d}s_{i-d} + b_1\right), \quad (3.10)$$

όπου $W_0, W_{-1}, W_{-2} \in \mathbb{R}^{T \times m}$ είναι οι μήτρες βάρους, $b_1 \in \mathbb{R}^T$ είναι η διανυσματική πόλωση, $y_i \in \mathbb{R}^T$ είναι η αναπαράσταση DA της έκφρασης s_i και T είναι ο αριθμός των DA.

Στη συνέχεια, η είσοδος του δεύτερου στρώματος του ANN είναι η αναπαράσταση του φορέα y_i που παρέχεται από το πρώτο στρώμα. Η τελική έξοδος του δικτύου είναι η πρόβλεψη της DA για τη φράση s_i που υπολογίζεται ως εξής:

$$z_i = \text{softmax}(U_0 y_i + b_2), \quad (3.11)$$

όπου $U_0 \in \mathbb{R}^{T \times T}$ και $b_2 \in \mathbb{R}^T$ είναι οι πίνακες βάρους και διανύσματος πόλωσης, αντίστοιχα. Για το μοντέλο διάρθρωσης του διαλόγου χρησιμοποιείται για το πρώτο στρώμα το μέγεθος του ιστορικού δύο προηγούμενων δηλώσεων και δεν λαμβάνεται υπόψη το ιστορικό για το δεύτερο στρώμα, όπως συνιστάται από τους [128].

3.4 Πειραματικό Σύνολο Δεδομένων

Το σύνολο δεδομένων που χρησιμοποιείται είναι το σύνολο δεδομένων Switchboard-DAMSL [42], το οποίο επισημαίνεται με τις 42 DAMSL ετικέτες. Το σύνολο δεδομένων Switchboard χρησιμοποιήθηκε αρχικά για την εκπαίδευση και τη δοκιμή διάφορων αλγορίθμων επεξεργασίας ομιλίας. Επίσης, έχει χρησιμοποιηθεί για άλλα προβλήματα όπως η Αυτόματη Αναγνώριση Ομιλίας (Automatic Speech Recognition - ASR) [145] και η προσαρμογή ακουστικού μοντέλου [146], συμπεριλαμβανομένης της μοντελοποίησης των DAs [131]. Αυτό το σύνολο δεδομένων χωρίζεται σε υποσύνολα εκπαίδευσης και δοκιμής όπως προτείνεται από τον [124]. Το σετ εκπαίδευσης (train set) περιλαμβάνει 1.155 διαλόγους (199.050 δηλώσεις) και το σετ δοκιμής (test set) 19 διαλόγους (3.927 διαλέξεις) που συγκεντρώθηκαν μέσω τηλεφώνου από 500 διαφορετικούς ομιλητές. Οι γραπτοί διάλογοι παρέχονται επίσης. Το θέμα της συζήτησης μεταξύ δύο ηχείων εισάγεται από έναν ηλεκτρονικό ρομπότ με γνώμονα τον υπολογιστή και καταγράφεται η συνομιλία που ακολουθεί. Περίπου 70 περιστασιακά θέματα παρουσιάστηκαν. Στον Πίνακα 3.1 παρουσιάζεται η διάρκεια των διαλόγων (από την άποψη του αριθμού των διατυπώσεων) που περιλαμβάνονται στο σύνολο δεδομένων. Ένα σύνολο επικύρωσης (validation set) δημιουργήθηκε με τυχαία επιλογή 115 διαλόγων (13.192 προφορικές δηλώσεις) από το σετ εκπαίδευσης.

# of Utterances per dialogue	Train set	Test set
min value	92	187
max value	954	679
mean value	334.6	410.0

Πίνακας 3.1: Σύνολο δεδομένων Switchboard-DAMSL.

Στον Πίνακα 3.2 παρουσιάζονται αντιπροσωπευτικά παραδείγματα των οκτώ πιο συχνών DAs. Επιπλέον, η κατανομή των DAs πάνω στο σύνολο δεδομένων αναφέρεται στον Πίνακα 3.3. Όπως φαίνεται σε αυτόν τον πίνακα, η πιο συχνή DA είναι η “Statement-non-opinion”.

Καμία προεπεξεργασία, συμπεριλαμβανομένων εργαλείων για την απογύμνωση των σημείων στίξης και την αλλαγή της κεφαλαιοποίησης, δεν εφαρμόζεται στο σύνολο δεδομένων. Για τα πειράματα που ακολουθούν η ακρίβεια (accuracy) ταξινόμησης χρησιμοποιείται ως μέτρηση αξιολόγησης.

DA tag	Example
Statement-non-opinion	There’s no one else that works there.
Acknowledge (Backchannel)	Sure.
Statement-opinion	but I think its relevance is pretty limited.
Agree/Accept	That’s right.
Abandoned or Turn-Exit	Do you,-
Appreciation	Well good.
Yes-No-Question	So do you have a family too?
Non-verbal	<Laughter>.

Πίνακας 3.2: Παραδείγματα των πιο συχνών DAs.

3.5 Ρύθμιση Παραμέτρων

Σε αυτό το σημείο, περιγράφουμε τη διαδικασία επιλογής των λέξεων-κλειδιών του σημασιολογικού μοντέλου (βλέπε Ενότητα 3.5.1) και τη ρύθμιση των υπερπαραμέτρων του μοντέλου βάσης LSTM (βλ. Ενότητα 3.5.2). Για τον συντονισμό χρησιμοποιήσαμε το σύνολο ανάπτυξης που αναφέρεται στο τμήμα 4.5.

3.5.1 Επιλογή Λέξεων-κλειδιών

Για την επιλογή των λέξεων-κλειδιών, η ακρίβεια ταξινόμησης υπολογίζεται όταν εφαρμόζονται διαφορετικά όρια στις μετρήσεις που περιγράφονται στην Ενότητα 3.3.1. Η καλύτερη απόδοση επιτυγχάνεται όταν έχουν επιλεγεί 323 λέξεις-κλειδιά (για $S(w)f(w) = 200$ και $\max_{i=1}^T p(t_i|w) = 0.5$).

DA tag	Train set (%)	Test set (%)
Statement-non-opinion	36.9	31.5
Acknowledge (Backchannel)	18.8	18.2
Statement-opinion	12.7	17.1
Agree/Accept	7.6	8.6
Abandoned or Turn-Exit	5.5	5.0
Appreciation	2.3	2.2
Yes-No-Question	2.3	2.0
Non-verbal	1.7	1.9
Remaining DAs	12.2	13.5

Πίνακας 3.3: Σχετική συχνότητα (%) των DAs.

Ενδεικτικά παραδείγματα των επιλεγμένων λέξεων-κλειδιών για τις πιο συχνές DAs παρουσιάζονται στον Πίνακα 3.4.

3.5.2 Παράμετροι του LSTM

Για την εφαρμογή του baseline μοντέλου (βλέπε Ενότητα 3.3.1) χρησιμοποιήθηκαν τα πακέτα NN που παρείχαν τα [141] και [147]. Μία υπερπαράμετρος κάθε φορά ρυθμίζεται διατηρώντας τις υπόλοιπες σταθερές για να προσδιοριστεί η καλύτερη διαμόρφωση. Με βάση τα ευρήματα που ελήφθησαν από τη βιβλιογραφία [129], αρχικοποιούμε τις παραμέτρους με τις ακόλουθες τιμές: word embeddings = 200 διαστάσεων διανύσματα με GloVe [110], ποσοστό απόσβεσης = 0.7, dropout = 0.3, pooling-mechanism = mean-pooling.

Word Embeddings. Διατηρώντας σταθερά τα υπερπαραμετρικά στοιχεία του δικτύου LSTM, δοκιμάζονται διαφορετικές τεχνικές από λέξη σε διάνυσμα και αριθμοί διαστάσεων των διανυσμάτων λέξεων, που αποτελούν μέρος της εισόδου στο δίκτυο. Τα διανύσματα λέξεων εκπαιδεύονται είτε με τη μέθοδο word2vec [120, 109] στο σύνολο δεδομένων του GoogleNews ή με τη μέθοδο GloVe [110] στο σύνολο δεδομένων του CommonCrawl. Όσον αφορά τις διαστάσεις των ενσωματωμένων λέξεων, χρησιμοποιούμε τα αναφερόμενα στο [128]¹. Τα word embeddings συνδέονται στη συνέχεια με τα χαρακτηριστικά που εξάγονται από το σημασιολογικό μοντέλο. Η απόδοση για διάφορες διαστάσεις παρουσιάζεται στον Πίνακα 3.5. Όπως φαίνεται σε αυτόν τον πίνακα, η καλύτερη απόδοση (75.6%) επιτυγχάνεται όταν χρησιμοποιούνται word embeddings 200 διαστάσεων. Επομένως, τα πειράματα που ακολουθούν χρησιμοποιούν αυτή τη ρύθμιση.

Ρυθμός πτώσης (Decay Rate). Ο ρυθμός πτώσης είναι ένας παράγοντας κανονικοποίησης της ενημέρωσης των βαρών σύνδεσης δικτύου, προκειμένου να αποφευχθεί το πρόβλημα του overfitting. Τυπικά, η τιμή ρυθμού πτώσης βρίσκεται μεταξύ 0 και 1. Εδώ εξετάζονται οι ρυθμοί πτώσης που συνιστώνται στη βιβλιογραφία [128, 129], όπως φαίνεται στον Πίνακα 3.5. Η καλύτερη απόδοση

¹ Η μέθοδος word2vec αποδίδει χαμηλότερη ακρίβεια ταξινόμησης (κατά 0.2%) σε σύγκριση με το GloVe και δεν αναφέρεται στον Πίνακα 3.5.

DA tag	Selected keywords
Statement-non-opinion	want, can't, work, mine, decided, always, remember
Acknowledge (Backchannel)	huh-uh, huh, yeah, yep, what?, huh?
Statement-opinion	seem, think, scary, ought, worse, difficult
Agree/Accept	true, agree, yes
Abandoned or Turn-Exit	-, -/, -
Appreciation	gosh, dear, wow, kidding
Yes-No-Question	mean?, there?, then?, all?
Non-verbal	<Laughter>, <Noise>, <Clicking>., <sniffing>

Πίνακας 3.4: Παραδείγματα αυτόματα επιλεγμένων λέξεων-κλειδιών (εμφανίζονται για τις πιο συχνές DAs).

(ακρίβεια 75.6 %) επιτυγχάνεται με ρυθμό πτώσης ίσο με 0.7 και αυτή η ρύθμιση χρησιμοποιείται για τα υπόλοιπα πειράματα.

Απόρριψη (Dropout). Για τα περισσότερα DNNs, το dropout [148] χρησιμοποιείται ως τεχνική κανονικοποίησης ενάντια στο πρόβλημα του overfitting. Στον Πίνακα 3.5, η επίδραση της τιμής του dropout στην ακρίβεια ταξινόμησης παρουσιάζεται για τιμές που κυμαίνονται μεταξύ 0.0 και 0.5 όπως προτείνεται στη βιβλιογραφία [128, 129]. Η καλύτερη απόδοση (ακρίβεια 75.6%) επιτυγχάνεται όταν ο ρυθμός εγκατάλειψης είναι ίσος με 0.3 και αυτή η ρύθμιση χρησιμοποιείται για τα πειράματα που ακολουθούν.

Μηχανισμός Συγκέντρωσης (Pooling mechanism). Οι διάφοροι μηχανισμοί που μπορούν να χρησιμοποιηθούν στο στρώμα pooling (max-, mean-, and last-pooling) όπως περιγράφεται στην ενότητα 3.3.1, δοκιμάζονται. Η απόδοση (ακρίβεια ταξινόμησης) για τα διάφορα σχήματα pooling αναφέρεται στον Πίνακα 3.5. Η υψηλότερη ακρίβεια ταξινόμησης (75,6%) προκύπτει από το μέσο όρο, όπου και υιοθετείται.

Άλλες Υπερπαραμέτροι. Εδώ, αναφέρουμε εν συντομία τις ρυθμίσεις για μια σειρά άλλων παραμέτρων σύμφωνα με τα ευρήματα της βιβλιογραφίας [129]. Η τιμή του l_2 -regularization ορίζεται σε $1e - 5$ και η λειτουργία \tanh χρησιμοποιείται ως συνάρτηση ενεργοποίησης στο LSTM. Επιπλέον, όπως αναφέρθηκε από [129], οι αλλαγές στον ρυθμό εκμάθησης δεν επηρεάζουν την απόδοση του μοντέλου. Ως εκ τούτου, ο ρυθμός εκμάθησης ορίζεται σε $1e - 3$.

3.5.3 Αποτελέσματα και Αξιολόγηση

Στον Πίνακα 3.6 αναφέρεται η ακρίβεια ταξινόμησης τόσο για το baseline όσο και για το προτεινόμενο μοντέλο. Η υψηλότερη ακρίβεια (75.6%) επιτυγχάνεται με το προτεινόμενο μοντέλο να ξεπερνά το baseline κατά 3.8% όταν χρησιμοποιείται τόσο η φράση όσο και η διάρθρωση του διαλόγου. Όσον αφορά την ανάλυση σε επίπεδο προτάσεων, η διαφορά μεταξύ του προτεινόμενου μοντέλου και του baseline είναι ακόμη μεγαλύτερη (4.3%). Στον Πίνακα 3.6 παρουσιάζεται επίσης η απόδοση του baseline μοντέλου, κατά την εφαρμογή της προεπεξεργασίας του συνόλου δεδομένων. Σε αυτή

Word embeddings	Decay rate	Dropout	Pooling mechanism	Classification Accuracy(%)
50				74.7
150	0.7	0.3	mean	75.4
200				75.6
300				75.2
	0.3			74.3
200	0.5	0.3	mean	75.1
	0.9			74.1
		0.0		75.4
		0.1		75.4
200	0.7	0.2	mean	75.5
		0.4		75.4
		0.5		75.2
200	0.7	0.3	max	75.3
			last	75.2

Πίνακας 3.5: Απόδοση υπερπαραμέτρων LSTM με βάση πειραματισμούς.

Model	Analysis Level	Preprocessing	Classification Accuracy(%)
Baseline	sentence	✗	69.5
		✓	72.8
Proposed		✗	73.8
Baseline	Sentence & discourse	✗	71.8
		✓	73.9
Proposed		✗	75.6

Πίνακας 3.6: Απόδοση του baseline και του προτεινόμενου μοντέλου.

την περίπτωση, το προτεινόμενο μοντέλο εξακολουθεί να υπερβαίνει τη baseline με ακρίβεια 1,7%.

Με βάση τα αποτελέσματα του Πίνακα 3.6, το προτεινόμενο μοντέλο ωφελείται από τις πρόσθετες σημασιολογικές πληροφορίες. Επιπλέον, αποδεικνύεται ότι το προτεινόμενο μοντέλο αποφεύγει την ανάγκη για προεπεξεργασία του συνόλου δεδομένων ².

Η απόδοση του προτεινόμενου μοντέλου είναι συγκρίσιμη με το state-of-the-art ³ ακρίβεια ταξινόμησης που ισούται με 77.0% [125]. Ένα πλεονέκτημα της προτεινόμενης μεθόδου είναι η χρήση ενός απλού μοντέλου σε σύγκριση με το [125] που απαιτεί την εξαγωγή χαρακτηριστικών επιπέδου διαρθρώσεως διαλόγου σε λανθάνουσα μορφή.

² Αυτό δικαιολογείται πειραματικά, επομένως δεν αναφέρεται η απόδοση του προτεινόμενου μοντέλου κατά την εφαρμογή της προεπεξεργασίας δεδομένων

³ Επίσης, επαναλήφθηκαν (χρησιμοποιώντας την ίδια υλοποίηση μοντέλου και δεδομένα) τα πειράματα που προτείνονται στο [129] χωρίς να επιτευχθούν τα ίδια αποτελέσματα.

Model	Classification Accuracy(%)
<i>Majority classification baseline</i>	31.6
Proposed	75.6
HMM [124]	71.0
LSTM [128]	69.6
CNN [128]	73.1
RCNN [127]	73.9
DRLM-joint training [125]	74.0
DRLM-conditional training [125]	77.0
Tf-idf (baseline)	47.3
<i>Inter-annotator agreement</i>	84.0

Πίνακας 3.7: Απόδοση του προτεινόμενου μοντέλου και άλλων μεθόδων από την βιβλιογραφία.

3.5.4 Συμπεράσματα

Σε αυτήν την εργασία, αποδείξαμε την αποτελεσματικότητα της ενσωμάτωσης των σημασιολογικών χαρακτηριστικών που σχετίζονται με τις ετικέτες DA για την ταξινόμηση DA με χρήση των RNNs. Αυτά τα χαρακτηριστικά υπολογίστηκαν σε σχέση με ένα σύνολο σημαντικών λέξεων-κλειδιών που αντιπροσωπεύουν σημασιολογικά τις ετικέτες DA. Τα προτεινόμενα χαρακτηριστικά βρέθηκαν να δίνουν 1.7% (απόλυτη) ακρίβεια ταξινόμησης σε σχέση με την baseline μέθοδο. Επίσης, πειραματικά δείξαμε ότι η πληροφορία σχετική με την διάρθρωση του διαλόγου (δηλαδή η σχετική με τις προηγούμενες φράσεις που δηλώνουν την ιστορία του διαλόγου) βελτιώνει περαιτέρω την απόδοση του baseline μοντέλου. Σε αντίθεση με παρόμοιες προσεγγίσεις που παρουσιάζονται στη βιβλιογραφία, το προτεινόμενο μοντέλο δεν απαιτεί πρόσθετα εργαλεία για την προεπεξεργασία των γραπτών διαλόγων.

Όσον αφορά τις μελλοντικές εργασίες, σχεδιάζουμε να ερευνήσουμε την ενσωμάτωση περισσότερων χαρακτηριστικών που προέρχονται από βαθύτερη ανάλυση της διάρθρωσης του διαλόγου. Επιπλέον, στοχεύουμε στην περαιτέρω επικύρωση των πειραματικών ευρημάτων αυτού του έργου χρησιμοποιώντας σύνολα δεδομένων σε άλλες γλώσσες εκτός από την αγγλική.

Κεφάλαιο 4

Μεταφορά Μάθησης από τα Προβλήματα Αναγνώρισης Πρόθεσης και Συναισθήματος στο Πρόβλημα Αναγνώρισης Προσωπικότητας σε Επίπεδο Εγγράφου

4.1 Εισαγωγή

Σύμφωνα με τη θεωρία των χαρακτηριστικών (trait theory), τα άτομα μπορούν να χαρακτηριστούν βάσει σκέψεων, συναισθημάτων και ενεργειών που παρουσιάζουν σχετική συνέπεια σε βάθος χρόνου. Το πρόβλημα της αναγνώρισης προσωπικότητας έχει ως στόχο να κατηγοριοποιήσει τα άτομα μεταξύ ενός συνόλου ποσοτικά εκτιμημένων χαρακτηριστικών. Η αυτόματη αναγνώριση της προσωπικότητας έχει πολλές εφαρμογές στην καθημερινή ζωή. Στα συστήματα συστάσεων, προτάσεις για ένα άτομο μπορούν να αντληθούν από τις προτιμήσεις άλλων με παρόμοια προσωπικότητα [149, 150]. Στην αλληλεπίδραση ανθρώπου-μηχανής, η απόδοση των συστημάτων διαλόγου μπορεί να βελτιωθεί εξοπλίζοντας τα με την ικανότητα ανάγνωσης της προσωπικότητας των χρηστών [151, 152].

Προηγούμενες επιτυχημένες προσεγγίσεις για την αναγνώριση της προσωπικότητας από έγγραφα, αξιοποιούν τις λεξικολογικές πληροφορίες χρησιμοποιώντας παραδοσιακές και βαθιές αρχιτεκτονικές μάθησης [4, 5, 6]. Τα state-of-the-art μοντέλα αναγνώρισης προσωπικότητας ενσωματώνουν επίσης ψυχολογικά και συναισθηματικά χαρακτηριστικά, τα οποία τυπικά εξάγονται από σχετικές βάσεις γνώσεων ή λεξικά, σε επίπεδο εγγράφου. Ωστόσο, η χρήση τέτοιων σχημάτων late-fusion παρέχει μόνο coarse-grained πληροφορίες στο μοντέλο. Επιπλέον, υπάρχουν και άλλες πηγές πληροφοριών που, παρότι θεωρούνται ενδεικτικές της προσωπικότητας ενός ατόμου, δεν έχουν ακόμη αξιοποιηθεί. Συγκεκριμένα, οι προηγούμενες μελέτες υποδεικνύουν τη συσχέτιση της προσωπικότητας των ατόμων με το γενικό πλαίσιο μιας δεδομένης κατάστασης, συμπεριλαμβανομένων και των προθέσεων τους την συγκεκριμένη στιγμή [3, 2, 153, 154].

Η μεταφορά μάθησης (Transfer Learning - TL) στοχεύει στη χρήση γνώσεων από συναφή προβλήματα-πηγές (source tasks), προκειμένου να βελτιωθεί η απόδοση σε ένα πρόβλημα-στόχο (target task). Η νευρωνική μεταφορά μάθησης (Neural TL), που είναι μια κοινή πρακτική στο Computer Vision (CV) [155, 156], έχει δείξει πρόσφατα πολλά υποσχόμενα αποτελέσματα και σε NLP προβλήματα [157, 158]. Το TL προσφέρει έναν εναλλακτικό τρόπο για την ενσωμάτωση πληροφοριών σχετικών με τα συναισθήματα και της προθέσεις του ατόμου στην αυτόματη αναγνώριση προσωπικότητας. Συγκεκριμένα, οι fine-grained πληροφορίες χαμηλού επιπέδου που εξάγονται από τα βαθιά νευρικά δίκτυα μπορούν να μεταφερθούν σε μοντέλα αναγνώρισης προσωπικότητας με αυτόματα και κλιμακωτό τρόπο.

Σε αυτή την εργασία, προτείνουμε δύο μεθόδους TL για την αξιοποίηση πληροφοριών από προεκπαιδευμένα μοντέλα επιπέδου προτάσεων για αναγνώριση προσωπικότητας σε επίπεδο εγγράφου. Τα μοντέλα μας βασίζονται σε δίκτυα ιεραρχικής προσοχής (Hierarchical Attention Network - HAN) [159]. Πρώτον, εκπαιδεύουμε ένα μοντέλο στα προβλήματα που σχετίζονται με τον τομέα-πηγή (source task), δηλαδή την αναγνώριση συναισθημάτων και προθέσεων. Στη συνέχεια, χρησιμοποιούμε τον

κωδικοποιητή του προεκπαιδευμένου δικτύου για την αρχικοποίηση του κωδικοποιητή προτάσεων του HAN και την επαναπροσαρμογή (fine-tuning) του μοντέλου στην εργασία-στόχο (target task). Χρησιμοποιούμε επίσης τον ίδιο προεκπαιδευμένο κωδικοποιητή ως εξολκεία χαρακτηριστικών προκειμένου να εμπλουτίσουμε τις παραστάσεις σε επίπεδο προτάσεων του δικτύου. Περαιτέρω, όπως προτείνεται στη βιβλιογραφία [5, 160], ερευνούμε και την ενσωμάτωση των λεξικο-βασισμένων ψυχολογικών χαρακτηριστικά στο μοντέλο. Οι προτεινόμενες προσεγγίσεις TL αξιολογούνται στα σύνολα δεδομένων [161] του YouTube [162] και του stream-of-consciousness essay και επιτυγχάνουν state-of-the-art αποτελέσματα.

Οι κύριες συνεισφορές της εργασίας αυτής είναι:

- Προτείνουμε καινοτόμες προσαρμογές δύο μεθόδων TL για τη μεταφορά γνώσεων από source tasks σε επίπεδο πρότασης σε ένα target task σε επίπεδο εγγράφου, επιτυγχάνοντας state-of-the-art αποτελέσματα.
- Διεξάγουμε ανάλυση της συνεισφοράς διαφόρων τύπων πληροφοριών (δηλαδή προθέσεων, συναισθημάτων, ψυχολογισσολογίας) στο πρόβλημα αναγνώρισης της προσωπικότητας. Αυτή είναι η πρώτη φορά, από όσο γνωρίζουμε, που το πρόβλημα αναγνώρισης πρόθεσης συσχετίστηκε με το πρόβλημα της αυτόματης αναγνώρισης της προσωπικότητας.

4.2 Επισκόπηση της Βιβλιογραφίας

Αναγνώριση Προσωπικότητας. Το πιο δημοφιλές ψυχολογικό μοντέλο για την περιγραφή της προσωπικότητας ενός ατόμου είναι το μοντέλο Big Five (ή Μοντέλο Πέντε Παράγοντων - Five Factor Model (FFM)), όπως περιγράφεται στον Πίνακα 4.1. Προηγούμενες προσεγγίσεις στην αναγνώριση της προσωπικότητας επικεντρώνονται στη χρήση ψυχολογισσικών και συναισθηματικών λεξικών για την πρόβλεψη των χαρακτηριστικών της προσωπικότητας των χρηστών στα κοινωνικά μέσα [163, 164, 160]. Συγκεκριμένα, οι [67] χρησιμοποίησαν το λεξικό Linguistic Inquiry Word Count (LIWC) [165], το οποίο περιέχει ψυχολογισσικές πληροφορίες. Οι [166] εξέτασαν διαφορετικά υποσύνολα αυτών των χαρακτηριστικών προκειμένου να ενισχύσουν περαιτέρω την απόδοση των αντίστοιχων μοντέλων και οι [167] προσέθεσαν στο χώρο χαρακτηριστικών άλλα συναφή χαρακτηριστικά, όπως η φαντασία, προκειμένου να προβλέψουν τα χαρακτηριστικά της προσωπικότητας συγγραφέων. Οι [168, 4] εκμεταλλεύτηκαν πρόσθετους λεξιλογικούς πόρους, όπως τα NRC [169], MRC [170], SentiStrength¹ και SPLICE². Οι [5] εξέτασαν την απόδοση τέτοιων χαρακτηριστικών χρησιμοποιώντας διαφορετικούς αλγόριθμους σε σύνολα δεδομένων κοινωνικής δικτύωσης (twitter, facebook, youtube). Επίσης, διερεύνησαν και επωφελήθηκαν από τη συσχέτιση μεταξύ των διαστάσεων της προσωπικότητας, αναφέροντας state-of-the-art αποτελέσματα στα σύνολα δεδομένων των κοινωνικών μέσων.

Πρόσφατα, εφαρμόστηκαν επιτυχώς προσεγγίσεις βαθιάς μάθησης για την πρόβλεψη της προσωπικότητας σε μεγάλα σύνολα δεδομένων σύντομων κειμένων (π.χ. tweets, facebook statuses) [167, 171]. Οι [6] εισήγαγαν μια προσέγγιση βαθιάς μάθησης στην αναγνώριση προσωπικότητας των εγγράφων σε επίπεδο εγγράφου από συγγραφείς που βασίζονται σε ιεραρχικά συνελκτικτικά νευρωνικά δίκτυα (Convolutional Neural Networks - CNNs). Ωστόσο, λόγω περιορισμένων επισημασμένων δεδομένων, η απόδοση του ιεραρχικού CNN δεν υπερέβαινε τις επιδόσεις προηγούμενων προσεγγίσεων. Για το λόγο αυτό, οι [6] φίλτραραν το σύνολο δεδομένων βάσει του συναισθηματικού λεξικού NRC και χρησιμοποίησε το λεξικό LIWC για να κατασκευάσει επιπρόσθετα χαρακτηριστικά (Mairresse features [67]) τα οποία δίνονται στο δίκτυο στο επίπεδο του διανύσματος εγγράφου. Επιπλέον, σε

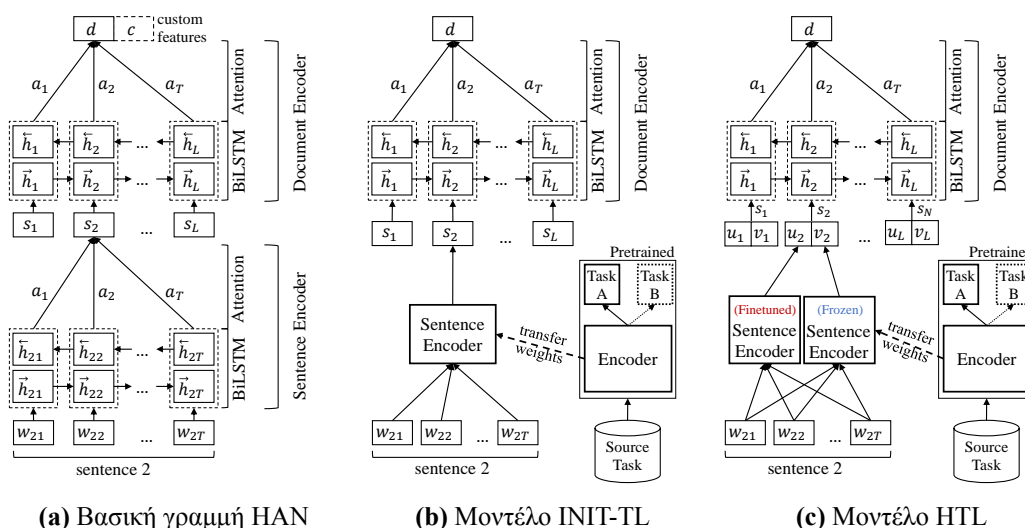
¹ <http://sentistrength.wlv.ac.uk>

² <http://splice.cmi.arizona.edu>

Big Five Trait	Description
Extraversion (Extr)	outgoing/energetic vs. solitary/reserved
Agreeableness (Agr)	friendly/compassionate vs. challenging/detached
Conscientiousness (Cons)	efficient/organized vs. easy-going/careless
Neuroticism (Neu)	sensitive/nervous vs. secure/confident
Openness (Open)	inventive/curious vs. consistent/cautious

Πίνακας 4.1: Περιγραφή μοντέλου Big Five

παρόμοια προβλήματα που σχετίζονται με τα καθήκοντα ταξινόμησης εγγράφων, όπως η εκτίμηση συναισθημάτων και η ταξινόμηση θέματος, τα HAN [159] έχουν χρησιμοποιηθεί με επιτυχία.



Σχήμα 4.1: Επισκόπηση των προτεινόμενων μοντέλων

Transfer Learning (TL). Το TL στοχεύει στη χρήση των γνώσεων από έναν source task, για τη βελτίωση της απόδοσης ενός μοντέλου σε έναν διαφορετικό, αλλά σχετικό target task. Το Neural TL έχει εφαρμοστεί με μεγάλη επιτυχία στο Computer Vision (CV) [155, 156]. Τα βαθιά νευρικά δίκτυα στην βιβλιογραφία σπάνια εκπαιδεύονται από το μηδέν και αντ' αυτού αρχικοποιούνται με προεκπαιδευμένα μοντέλα. Αξιοσημείωτα παραδείγματα περιλαμβάνουν την αναγνώριση προσώπου [172] και την οπτική απάντηση ερωτήσεων (Question Answering - QA) [173], όπου συνδυάζονται τα χαρακτηριστικά εικόνες που έχουν εκπαιδευτεί στο ImageNet [174] και τα word embeddings μέσω εκπαίδευσης χωρίς επίβλεψη. Παρόλο που η μεταφορά μάθησης έχει δει εκτεταμένη επιτυχία στην βιβλιογραφία, το neural TL πέρα από τα προεκπαιδευμένα word embeddings είναι λιγότερο διαδεδομένο στην Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing - NLP). Ορισμένες πρόσφατες προσπάθειες έχουν καταδείξει την επιτυχή χρήση των νευρωνικών γλωσσικών μοντέλων (Language Models - LM) ως γενικών εξολκέων χαρακτηριστικών [158]. Το TL έχει επίσης υιοθετηθεί σε εργασίες ταξινόμησης σε επίπεδο προτάσεων, όπως η αναγνώριση συναισθημάτων [175], η ανάλυση συναισθημάτων και η ανίχνευση σαρκασμού [157], προτρέποντας τα αντίστοιχα μοντέλα σε σημασιολογικά παρόμοια

προβλήματα. Ωστόσο, τέτοιες μέθοδοι δεν έχουν ακόμη δοκιμαστεί στον τομέα της αναγνώρισης προσωπικότητας.

4.3 Baseline Μοντέλο

Το baseline μοντέλο είναι ένα HAN, όπως απεικονίζεται στο Σχήμα 4.1a. Εκπαιδεύουμε ένα ξεχωριστό μοντέλο για κάθε διάσταση προσωπικότητας. Η είσοδος στο δίκτυο είναι ένα έγγραφο που αποτελείται από L φράσεις, όπου κάθε πρόταση περιέχει T λέξεις.

Κωδικοποιητής Πρότασης (Sentence Encoder - SE). Χρησιμοποιούμε ένα word embedding στρώμα για να προβάλλουμε τις λέξεις w_1, w_2, \dots, w_T σε ένα συνεχές χώρο διανυσμάτων R^E , όπου το E έχει το μέγεθος του embedding στρώματος. Στη συνέχεια, κωδικοποιούμε τις πληροφορίες σε κάθε πρόταση χρησιμοποιώντας δίκτυα Long Short-Term Memory (LSTM) [176]. Ένα LSTM παίρνει ως είσοδο το διάνυσμα μιας λέξης σε μια πρόταση και παράγει λέξεις h_1, h_2, \dots, h_T , όπου h_i είναι η κρυφή κατάσταση του LSTM στο χρονικό βήμα i , που συνοψίζει όλες τις πληροφορίες της φράσης μέχρι την i th λέξη. Χρησιμοποιούμε αμφίδρομο LSTM (Bidirectional LSTM - BiLSTM) για να λάβουμε επισημάνσεις λέξεων που συνοψίζουν τις πληροφορίες και από τις δύο κατευθύνσεις. Ένα BiLSTM αποτελείται από ένα forward LSTM \vec{f} που διαβάζει την πρόταση από w_1 σε w_T και ένα LSTM προς τα πίσω \overleftarrow{f} που διαβάζει την πρόταση από w_T σε w_1 . Παίρνουμε την τελική επισήμανση για μια δεδομένη λέξη w_i , συνενώνοντας τις επισημάνσεις και από τις δύο κατευθύνσεις, $h_i = \vec{h}_i \parallel \overleftarrow{h}_i$, $h_i \in R^{2S}$ όπου \parallel υποδηλώνει τη λειτουργία συνένωσης και S το μέγεθος κάθε LSTM σε επίπεδο προτάσεων.

Πάνω από το δίκτυο LSTM, προστίθεται ένας μηχανισμός προσοχής (attention mechanism) προκειμένου να εντοπιστούν οι πιο σημαντικές και αντιπροσωπευτικές λέξεις σε κάθε πρόταση. Ο μηχανισμός προσοχής αποδίδει ένα βάρος a_i σε κάθε επισήμανση λέξης h_i . Υπολογίζουμε την σταθερή αναπαράσταση s κάθε προτάσεως ως το σταθμισμένο άθροισμα όλων των σχολιασμών λέξεων.

$$e_i = \tanh(W_h h_i + b_h), \quad e_i \in [-1, 1] \quad (4.1)$$

$$a_i = \frac{\exp(e_i)}{\sum_{t=1}^T \exp(e_t)}, \quad \sum_{i=1}^T a_i = 1 \quad (4.2)$$

$$s = \sum_{i=1}^T a_i h_i, \quad s \in R^{2S} \quad (4.3)$$

όπου τα W_h και b_h είναι τα βάρη του στρώματος προσοχής.

Κωδικοποιητής Εγγράφου (Document encoder). Ο κωδικοποιητής εγγράφου χρησιμοποιεί την ίδια αρχιτεκτονική με τον κωδικοποιητή προτάσεων, εξαιρουμένου του embedding στρώματος. Διαβάζει την παραγόμενη ακολουθία αναπαραστάσεων των προτάσεων s_1, s_2, \dots, s_L και παράγει μια τελική σταθερή αναπαράσταση χαρακτηριστικών d για κάθε έγγραφο.

Στρώμα Εξόδου (Output Layer). Τροφοδοτούμε το d σε ένα τελικό στρώμα συγκεκριμένης εργασίας με έναν νευρώνα για την εκτέλεση ταξινόμησης ή παλινδρόμησης, ανά διάσταση προσωπικότητας. Κατά την προσθήκη προσαρμοσμένων χαρακτηριστικών σε επίπεδο εγγράφου, όπως απεικονίζεται στο Σχήμα 4.1a, το τελικό διάνυσμα χαρακτηριστικών υπολογίζεται ως $d \parallel c$, όπου \parallel δηλώνει τη λειτουργία συνένωσης και το c είναι το διάνυσμα προσαρμοσμένων χαρακτηριστικών.

4.4 Μεταφορά Μάθησης

Χρησιμοποιούμε τη γνώση από τα source tasks που σχετίζονται με την αναγνώριση της προσωπικότητας, προκειμένου να βελτιώσουμε τις επιδόσεις στο target task. Είναι πιθανότερο ότι τα δεδομένα

των source και target tasks προέρχονται από την ίδια κατανομή. Πιθανώς, τα χαρακτηριστικά γνωρίσματα από το source task θα είναι χρήσιμα για τη μοντελοποίηση των δειγμάτων στο target task.

Εμείς προσαρμόζουμε δύο πολύ γνωστές μεθόδους TL για τη μεταφορά γνώσεων από προβλήματα σε επίπεδο προτάσεων στο πρόβλημα σε επίπεδο εγγράφων, το οποίο διαμορφώνεται με RNN ιεραρχικής προσοχής. Συγκεκριμένα, εκπαιδεύουμε ένα SE ξεχωριστά για την αναγνώριση προθέσεων και συναισθημάτων καθώς και για τα δύο προβλήματα μέσω της μάθησης πολλαπλών εργασιών (multi-task learning). Προτείνουμε δύο διαφορετικές μεθόδους TL για την αξιοποίηση των γνώσεων που περιέχονται στο προεκπαιδευμένο SE.

4.4.1 Προεκπαίδευση του SE

Προκειμένου να προεκπαιδευτεί το SE, χρησιμοποιούμε τα προβλήματα αναγνώρισης προθέσεων και συναισθημάτων. Όπως προτείνεται στη βιβλιογραφία, οι έννοιες της δράσης (act) και συναισθημάτων συσχετίζονται σε μεγάλο βαθμό με την προσωπικότητα και ειδικότερα με συγκεκριμένα χαρακτηριστικά [153, 154]. Προσπαθούμε επίσης να συνδυάσουμε και τα δύο είδη πληροφοριών μέσω της μάθησης πολλαπλών εργασιών.

Στην περίπτωση της μάθησης πολλαπλών εργασιών, η συνάρτηση απώλειας (loss function) υπολογίζεται ως το σταθμισμένο άθροισμα των δύο μεμονωμένων απωλειών, ως εξής:

$$L = L_a + \alpha L_b \quad (4.4)$$

όπου L_a είναι η απώλεια προθέσεων, L_b είναι η απώλεια συναισθημάτων και α είναι ο πολλαπλασιαστής που χρησιμοποιείται για την άμβλυνση των ανισορροπιών απώλειας.

4.4.2 Αρχικοποίηση (INIT-TL) του SE

Ο πιο συνηθισμένος τρόπος μεταφοράς της γνώσης είναι η επαναπροσαρμογή ενός προεκπαιδευμένου μοντέλου στο target task. Δεδομένου ότι στοχεύουμε στη μεταφορά γνώσεων από ένα πρόβλημα επιπέδου προτάσεων σε ένα πρόβλημα σε επίπεδο εγγράφου, αρχικοποιούμε τα βάρη του SE με το προεκπαιδευμένο μοντέλο μας, ενώ ο κωδικοποιητής εγγράφων αρχικοποιείται τυχαία. Έπειτα, συντονίζουμε και τους δύο κωδικοποιητές για το target task³. Η προτεινόμενη αρχιτεκτονική παρουσιάζεται στο Σχήμα 4.1b.

4.4.3 Hypercolumns-TL (HTL) του SE

Ένας εναλλακτικός τρόπος μεταφοράς της γνώσης από ένα source task σε ένα target task είναι η εξαγωγή χαρακτηριστικών από ένα προεκπαιδευμένο νευρωνικό δίκτυο, γνωστό ως hypercolumns στο CV [177]. Συγκεκριμένα, χρησιμοποιούμε δύο ξεχωριστούς κωδικοποιητές προτάσεων. Ένας από αυτούς τυχαία αρχικοποιείται και προσαρμόζεται με ακρίβεια στο target task. Τα βάρη του άλλου SE αρχικοποιούνται από τον κωδικοποιητή του προεκπαιδευμένου μοντέλου και παραμένουν παγωμένα κατά τη διάρκεια της εκπαίδευσης, προκειμένου να διατηρηθεί η γνώση από το source task. Αυτό λειτουργεί επίσης ως πρόσθετο μέτρο ασφάλειας κατά του προβλήματος του catastrophic forgetting. Παίρνουμε τις εμπλουτισμένες αναπαραστάσεις προτάσεων s_i , οι οποίες τροφοδοτούνται στον κωδικοποιητή εγγράφων ως εξής:

$$s_i = u_i \parallel v_i \quad (4.5)$$

³ Πειραματικά διαπιστώσαμε ότι το προτεινόμενο μοντέλο επιτυγχάνει καλύτερα αποτελέσματα όταν οι κωδικοποιητές είναι καλά ρυθμισμένοι αντί να διατηρήσουμε το SE παγωμένο.

όπου u_i και v_i είναι οι αναπαραστάσεις των προτάσεων όπως προκύπτουν από την μοντελοποίηση συγκεκριμένων προβλημάτων και των προεκπαιδευμένων SE, αντίστοιχα. Η παραπάνω προσέγγιση παρουσιάζεται στο Σχήμα 4.1c.

4.5 Πειράματα & Αποτελέσματα

Σε αυτή την ενότητα παρουσιάζουμε το σύνολο δεδομένων που χρησιμοποιείται για την προεκπαίδευση του SE, καθώς και τα δύο σύνολα δεδομένων, στα οποία αξιολογούμε τις προτεινόμενες προσεγγίσεις μας για την αναγνώριση της προσωπικότητας. Περιγράφουμε επίσης την πειραματική μας ρύθμιση και επιδεικνύουμε τα αποτελέσματά μας, παράλληλα με την ανάλυση της συνεισφοράς κάθε πηγής πληροφοριών, δηλαδή του συναισθήματος, της πρόθεσης και της ψυχολογολογίας.

4.5.1 Πειραματικά Σύνολα Δεδομένων

Προκειμένου να αξιολογήσουμε τα προτεινόμενα μοντέλα αναγνώρισης της προσωπικότητας, χρησιμοποιούμε δύο σύνολα δεδομένων διαφορετικής φύσης. Το ένα προέρχεται από τα κοινωνικά μέσα ενημέρωσης και επισημένεται σε συνεχή κλίμακα και το άλλο αποτελείται από γραπτά δοκίμια και επισημένεται σε δυαδική κλίμακα. Τέλος, για προεκπαίδευση (βλ. Ενότητα 4.4.1) χρησιμοποιούμε ένα πολύ μεγαλύτερο σύνολο δεδομένων επισημασμένο με ετικέτες πρόθεσης και συναισθήματος.

Σύνολο Δεδομένων για Προεκπαίδευση: DailyDialog

Το σύνολο δεδομένων DailyDialog [178] αποτελείται από διαλόγους **13,118** για διάφορα θέματα, όπως η καθημερινή ζωή, η σχολική ζωή, ο πολιτισμός και η εκπαίδευση και η πολιτική. Κάθε διάλογος του συνόλου δεδομένων αποτελείται από περίπου 8 στροφές ομιλητών με 14.6 λέξεις ανά πρόταση, κατά μέσο όρο. Το σύνολο δεδομένων επισημένεται με βάση τις προθέσεις (Inform, Question, Directives, Commissive) και τα διακριτά συναισθήματα (No emotion, Anger, Disgust, Fear, Happiness, Sadness, Surprise). Στους Πίνακες 4.2 και 4.3 παρουσιάζουμε τη διανομή των ετικετών για πρόθεση και συναίσθημα στο σύνολο δεδομένων, αντίστοιχα.

Intent dimension	Absolute Value	Percentage (%)
Inform	46,532	45.2
Question	29,428	28.6
Directives	17,295	16.8
Commissive	9,724	9.4

Πίνακας 4.2: Στατιστικά ετικετών πρόθεσης στο DailyDialog

Emotion dimension	Absolute Value	Percentage (%)
Anger	1,022	0.99
Disgust	353	0.34
Fear	74	0.17
Happiness	12,885	12.51
Sadness	1,150	1.12
Surprise	1,823	1.77
No emotion	85572	83.10

Πίνακας 4.3: Στατιστικά ετικετών συναισθημάτων στο DailyDialog

Σύνολο Δεδομένων YouTube

Το σύνολο δεδομένων του YouTube [161] αποτελείται από **404** vlogs, το καθένα από διαφορετικό vlogger. Τα αποτελέσματα που αντιστοιχούν σε κάθε vlogger για κάθε ένα από τα πέντε χαρακτηριστικά της προσωπικότητας είναι αποτελέσματα που έχουν οριστεί από τους θεατές. Μετά από τον μέσο όρο των βαθμολογιών που αντιστοιχούν σε κάθε vlog, οι σχολιασμοί για κάθε χαρακτηριστικό προσωπικότητας βρίσκονται στο διάστημα [1, 7] με πραγματικές τιμές. Επιπλέον, το σύνολο δεδομένων YouTube προσφέρει ένα σύνολο εξωγενών οπτικοακουστικών χαρακτηριστικών. Στους Πίνακες ?? παρουσιάζονται οι κατανομές του αριθμού των λέξεων ανά πρόταση και ο αριθμός των προτάσεων ανά vlog αντίστοιχα.

	min	max	avg
words per sentence	1.00	137.00	14.72
sentences per vlog	2.00	141.00	37.65

Πίνακας 4.4: Στατιστικά στοιχεία για το σύνολο δεδομένων YouTube label sentences

Σύνολο Δεδομένων Stream-of-consciousness essay

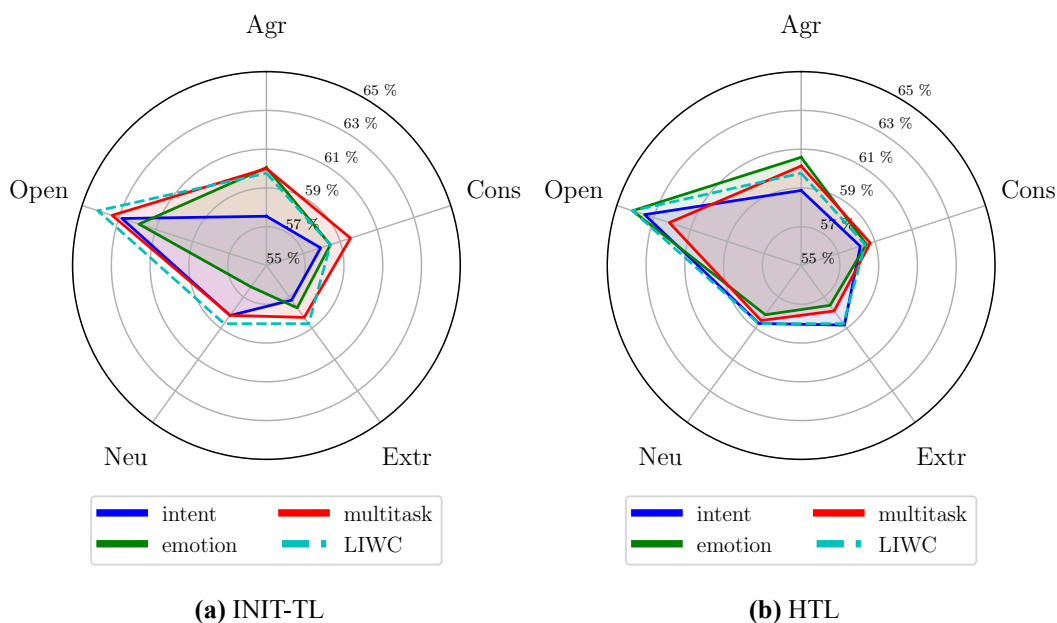
Το σύνολο δεδομένων stream-of-consciousness essay που παρέχεται από το [162] περιέχει **2,468** ανώνυμα δοκίμια επισημασμένα με τα χαρακτηριστικά προσωπικότητας των συγγραφέων. Τα stream-of-consciousness essays γράφτηκαν από τους εθελοντές σε ένα ελεγχόμενο περιβάλλον και οι συγγραφείς τους κλήθηκαν να καθορίσουν τα δικά τους χαρακτηριστικά σύμφωνα με το Big Five μοντέλο. Τα χαρακτηριστικά της προσωπικότητας για κάθε συγγραφέα επισημάνονται με μια ετικέτα "ναι" ή "όχι". Στον Πίνακα 4.5 παρουσιάζονται οι κατανομές του αριθμού των λέξεων ανά πρόταση και ο αριθμός των προτάσεων ανά δοκίμιο, αντίστοιχα.

	min	max	avg
words per sentence	1.00	892.00	13.35
sentences per essay	1.00	336.00	28.60

Πίνακας 4.5: Στατιστικά στοιχεία για το σύνολο δεδομένων stream-of-consciousness essay

Model	Extr (%)	Agr (%)	Cons (%)	Neu (%)	Open (%)
Majority	51.72	50.02	53.10	50.79	51.52
State-of-the-art (CNN, NRC & Mairesse features) [6]	58.09	59.38	56.71	57.30	62.68
HAN baseline	55.73	56.50	55.21	55.33	61.68
HAN baseline & LIWC	58.71	59.76	58.47	58.71	64.14
INIT-TL (intent) (FT)	57.21	57.54	57.94	58.19	62.84
INIT-TL (emotion) (FT)	57.70	60.05	58.47	56.36	61.91
INIT-TL (intent & emotion) (FT)	58.31	60.00	59.56	58.18	63.37
HTL (intent)	58.80	58.87	58.21	58.70	63.49
HTL (emotion)	57.54	60.58	58.54	58.14	64.14
HTL (intent & emotion)	57.90	60.13	58.75	58.51	62.15

Πίνακας 4.6: Αποτελέσματα στο σύνολο δεδομένων stream-of-consciousness essay. Οι βαθμολογίες υπολογίζονται με τη χρήση της μέτρησης ακρίβεια ταξινόμησης (*accuracy*) χρησιμοποιώντας 10-fold cross-validation.



Σχήμα 4.2: Συμβολή των πηγών πληροφόρησης για κάθε μέθοδο TL στο σύνολο δεδομένων stream-of-consciousness essay. Η μέτρηση αξιολόγησης είναι ακρίβεια ταξινόμησης (*accuracy*), πράγμα που σημαίνει ότι οι μεγαλύτερες επιφάνειες υποδηλώνουν καλύτερη απόδοση. Οι προτεινόμενες προσεγγίσεις είναι ανταγωνιστικές με το μοντέλο HAN ενισχυμένο με τα χαρακτηριστικά, που έχουν εξαχθεί βάσει του λεξικού LIWC.

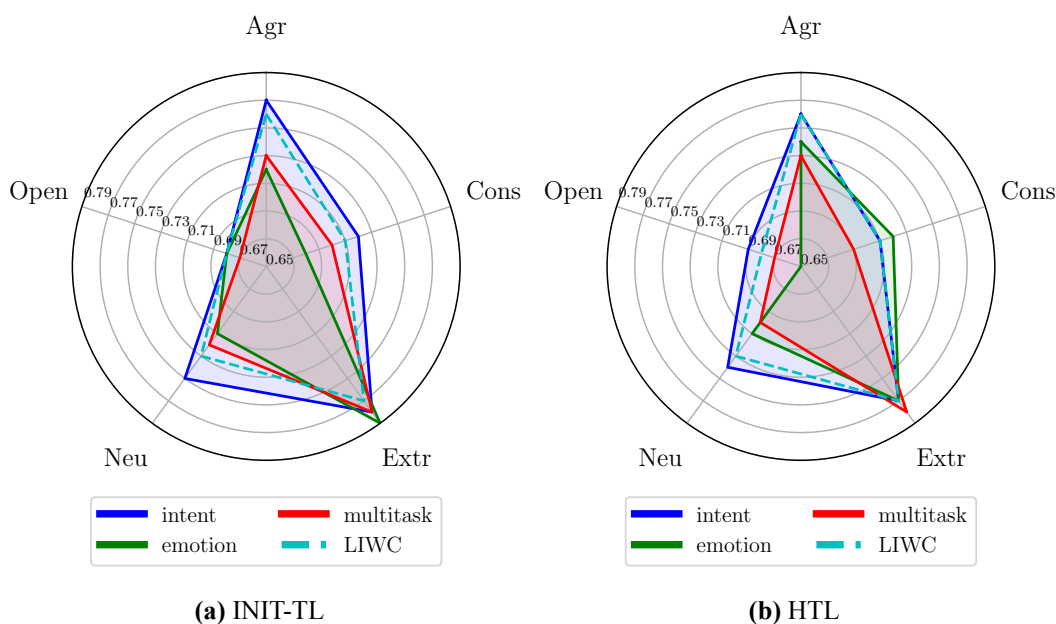
4.5.2 Πειραματική Ρύθμιση των Μοντέλων

Αρχικοποιούμε το στρώμα ενσωμάτωσης των μοντέλων μας με προεκπαιδευμένα GloVe word embeddings διάστασης 300 [110]⁴. Προσθέτουμε επίσης Gaussian θόρυβο με $\sigma = 0.1$ στο embedding

⁴ Χρησιμοποιήσαμε τα word embeddings 6B GloVe λόγω περιορισμών σχετικών με την απαιτούμενη μνήμη για την εκπαίδευση των μοντέλων.

Model	Extr (RMSE)	Agr (RMSE)	Cons (RMSE)	Neu (RMSE)	Open (RMSE)
State-of-the-art (MTSC) [5]	0.85	0.72	0.69	0.70	0.69
HAN baseline	0.80	0.87	0.74	0.78	0.74
HAN baseline & LIWC	0.77	0.76	0.71	0.73	0.68
INIT-TL (intent) (FT)	0.78	0.77	0.72	0.75	0.68
INIT-TL (emotion) (FT)	0.79	0.72	0.68	0.71	0.68
INIT-TL (intent & emotion) (FT)	0.78	0.73	0.70	0.72	0.67
HTL (intent)	0.77	0.76	0.71	0.74	0.69
HTL (emotion)	0.77	0.74	0.72	0.71	0.65
HTL (intent & emotion)	0.78	0.73	0.69	0.70	0.67

Πίνακας 4.7: Αποτελέσματα για το σύνολο δεδομένων YouTube. Οι βαθμολογίες υπολογίζονται χρησιμοποιώντας τη μέτρηση $RMSE$, με 10-fold cross-validation. Τα προσαρμοσμένα οπτικοακουστικά χαρακτηριστικά περιλαμβάνονται σε όλα τα μοντέλα (δείτε το Σχήμα 4.1a).



Σχήμα 4.3: Συμβολή των πηγών πληροφοριών για κάθε μέθοδο TL στο σύνολο δεδομένων του YouTube. Η μέτρηση αξιολόγησης είναι $RMSE$, πράγμα που σημαίνει ότι οι μικρότερες επιφάνειες υποδηλώνουν καλύτερη απόδοση. Και οι δύο προσεγγίσεις υπερβαίνουν το μοντέλο HAN ενισχυμένο με τα χαρακτηριστικά, που έχουν εξαχθεί βάσει του λεξικού LIWC για τις περισσότερες ρυθμίσεις.

στρώμα, το οποίο μπορεί να ερμηνευτεί ως μια τεχνική τυχαίας αύξησης των δεδομένων, που κάνει τα μοντέλα μας πιο ανθεκτικά στο πρόβλημα του overfitting. Επιπλέον, εφαρμόζουμε dropout [179] ίσο με 0.2 στο embedding στρώμα και χρησιμοποιούμε early stopping. Επιπλέον, ρυθμίσαμε περαιτέρω (fine-tuning) τα βάρη του embedding στρώματος κατά τη διάρκεια της εκπαίδευσης σε όλες τα προβλήματα και χρησιμοποιήσαμε τον βελτιστοποιητή Adam [115]. Τέλος, χρησιμοποιούμε ένα στρώμα BiLSTM με 150 νευρώνες (ανά κατεύθυνση), τόσο για τον κωδικοποιητή προτάσεων όσο και για τον κωδικοποιητή εγγράφων.

Χρησιμοποιούμε διαφορετική λειτουργία απώλειας για την εκπαίδευση καθενός από τα μοντέλα που περιγράφηκαν προηγουμένως. Για τα προεκπαιδευμένα μοντέλα για αναγνώριση προθέσεων και συναισθημάτων, χρησιμοποιούμε τη συνάρτηση cross-entropy loss. Στην περίπτωση της μάθησης πολλαπλών εργασιών, ορίσαμε εμπειρικά την τιμή του πολλαπλασιαστή a στην Εξίσωση 4.4 με 1.5. Για την εκπαίδευση του HAN για την αναγνώριση της προσωπικότητας στο σύνολο δεδομένων stream-of-consciousness essay χρησιμοποιούμε την συνάρτηση binary cross-entropy loss και στο σύνολο δεδομένων του YouTube τη συνάρτηση απώλειας *Mean Square Error* (MSE).

Προκειμένου να εκπαιδευτεί το SE στο σύνολο δεδομένων DailyDialog, υιοθετούμε την τυπική διαίρεση δεδομένων σε εκπαίδευσης-επικύρωσης, όπως παρέχεται από το [178]. Η αξιολόγηση των μοντέλων μας για την αναγνώριση της προσωπικότητας βασίζεται σε 10-fold cross-validation, όπως ήδη προτάθηκε στη βιβλιογραφία [5, 6] για αυτά τα σύνολα δεδομένων. Συγκεκριμένα, για το σύνολο δεδομένων YouTube, χρησιμοποιούμε τη μέτρηση αξιολόγησης *Μέσο τετραγωνικό σφάλμα* (*Root Mean Squared Error*) (RMSE), αφού αντιμετωπίζουμε ένα πρόβλημα παλινδρόμησης πολλαπλών ετικετών. Για το σύνολο δεδομένων stream-of-consciousness essay, το οποίο είναι ένα πρόβλημα δυαδικής ταξινόμησης πολλαπλών ετικετών, χρησιμοποιούμε τη μέτρηση *Ακρίβεια ταξινόμησης* (*Accuracy*) σύμφωνα με τη βιβλιογραφία.

Τέλος, σε ένα υποσύνολο των διεξαγόμενων πειραμάτων, χρησιμοποιούμε τα χαρακτηριστικά LIWC βασισμένα σε λεξικό, τα οποία χρησιμοποιούνται συνήθως στη βιβλιογραφία [5, 166, 6]. Συγκεκριμένα, εξάγουμε για κάθε έγγραφο (vlog ή δοκίμιο ανάλογα με το χρησιμοποιούμενο σύνολο δεδομένων) 81 ψυχολογικά χαρακτηριστικά σε επίπεδο εγγράφου. Αυτά τα χαρακτηριστικά σχετίζονται με διαστάσεις όπως οι κοινωνικές διαδικασίες (π.χ. “talk”, “mate”, “child”), συναισθηματικές διαδικασίες (π.χ. “happy”, “cryed” (π.χ. “worried”, “nervous”) και γνωστικές διεργασίες (π.χ. “cause”, “know”). Μια πλήρης επισκόπηση όλων των ψυχολογικών διαστάσεων που περιλαμβάνονται στο λεξικό LIWC παρέχεται από το [180]. Στην περίπτωση του συνόλου δεδομένων YouTube, χρησιμοποιούμε επίσης σε όλα τα πειράματα 26 προσαρμοσμένα οπτικοακουστικά χαρακτηριστικά που παρέχονται από το [161]. Αυτές οι προσαρμοσμένες λειτουργίες προστίθενται στα μοντέλα όπως απεικονίζονται στο Σχήμα 4.1a (υποσύνολο του προσαρμοσμένου διανύσματος χαρακτηριστικών c).

4.5.3 Αποτελέσματα

Μέθοδοι TL. Στους πίνακες 4.6 και 4.7, παρουσιάζουμε την απόδοση των δύο προτεινόμενων μεθόδων TL, δηλαδή του INIT-TL και του HTL, στα σύνολα δεδομένων stream-of-consciousness essay και YouTube αντίστοιχα. Παρέχουμε επίσης τα κορυφαία αποτελέσματα καθώς και την απόδοση του baselina μοντέλου κατά την προσθήκη των χαρακτηριστικών που έχουν εξαχθεί από το λεξικό LIWC. Για κάθε μέθοδο TL αναφέρουμε τα πειραματικά αποτελέσματα μετά την προεκπαίδευση του SE για την αναγνώριση πρόθεσης, συναισθήματος και των δύο προβλημάτων μέσω της μάθησης πολλαπλών εργασιών.

Όπως παρατηρήθηκε, και οι δύο μέθοδοι TL βελτιώνουν την απόδοση σε σχέση με το βασικό μοντέλο. Συγκεκριμένα, για το σύνολο δεδομένων stream-of-consciousness essay, η βελτίωση των αποτελεσμάτων μπορεί να φτάσει έως και 4% σε (απόλυτη) ακρίβεια (Acc), ανάλογα με το χρησιμοποιούμενο προεκπαιδευμένο μοντέλο. Στο σύνολο δεδομένων του YouTube, τα καλύτερα αποτελέσματα επιτυγχάνονται επίσης στο χαρακτηριστικό Acc , όπου η μέθοδος HTL με SE που έχει προεκπαιδευτεί από συναισθήματα μπορεί να προσφέρει βελτίωση κατά 0.15 (RMSE metric). Συνολικά, τόσο το INIT-TL όσο και το HTL παρουσιάζουν συγκρίσιμες επιδόσεις σε όλα τα γνωρίσματα και τα σύνολα δεδομένων. Ωστόσο, στις περισσότερες πειραματικές ρυθμίσεις η μέθοδος HTL επιτυγχάνει καλύτερα αποτελέσματα σε σύγκριση με το INIT-TL. Στην πραγματικότητα, για τα ειδικά χαρακτηριστικά της προσωπικότητας και τα προρυθμισμένα SE, η μέθοδος HTL ξεπερνά το INIT-TL κατά 1%

σε (απόλυτη) ακρίβεια για το σύνολο δεδομένων των δοκιμών (Πίνακας 4.6, INIT-TL έναντι HTL με πρόθεση για χαρακτηριστικό *Agr*) . Για το σύνολο δεδομένων YouTube (Πίνακας 4.7), η διαφορά απόδοσης μεταξύ των δύο μεθόδων TL είναι μικρότερη.

Και οι δύο προσεγγίσεις TL παρέχουν αποτελέσματα συγκρίσιμα με εκείνα του baseline HAN κατά την προσθήκη προσαρμοσμένων χαρακτηριστικών LIWC (state-of-the-art προσέγγιση). Συγκεκριμένα, για τις διαστάσεις προσωπικότητας των *Extr*, *Agr* και *Cons*, η μέθοδος HTL παρουσιάζει καλύτερη επίδοση από το baseline HAN ενισχυμένο με τα χαρακτηριστικά LIWC, επιτυγχάνοντας ακρίβεια ίση με **58.80%**, **60.58%** και **59.56%** (Πίνακας 4.6) αντίστοιχα. Στις υπόλοιπες διαστάσεις της προσωπικότητας, δηλαδή *Neu* και *Open*, η απόδοση των δύο προσεγγίσεων είναι συγκρίσιμη (**58.70%** και **64.14%**, αντίστοιχα όταν χρησιμοποιείται διαφορετικό προεκπαιδευμένο SE στην περίπτωση της μεθόδου HTL). Για το σύνολο δεδομένων YouTube (Πίνακας 4.7), η προσέγγισή μας ξεπερνά το baseline HAN με τα χαρακτηριστικά LIWC για όλες τις διαστάσεις, εκτός από το *Extr* όπου η απόδοση των δύο προσεγγίσεων είναι συγκρίσιμη (**0.77 RMSE**).

Πηγές Πληροφοριών. Ως περαιτέρω βήμα έρευνας, απεικονίζουμε την απόδοση των προτεινόμενων μεθόδων και αναλύουμε τη συμβολή κάθε πηγής πληροφοριών σε όλες τις μεθόδους TL και τα σύνολα δεδομένων. Συγκρίνουμε επίσης τα προτεινόμενα μοντέλα με το baseline HAN με την προσθήκη των χαρακτηριστικών LIWC (HAN-LIWC). Στο Σχήμα 4.2, παρουσιάζουμε τα αποτελέσματα του συνόλου δεδομένων stream-of-consciousness essay. Η μέτρηση αξιολόγησης είναι η ακρίβεια ταξινόμησης (accuracy), πράγμα που σημαίνει ότι οι μεγαλύτερες επιφάνειες είναι καλύτερες. Παρατηρούμε ότι και οι δύο μέθοδοι TL είναι συγκρίσιμες με το HAN-LIWC. Το Σχήμα 4.3 απεικονίζει τα αποτελέσματα στο σύνολο δεδομένων του YouTube. Δεδομένου ότι η μέτρηση αξιολόγησης είναι RMSE, μια μικρότερη επιφάνεια υποδηλώνει καλύτερη απόδοση. Σε αυτή την περίπτωση, παρατηρούμε ότι και οι δύο προσεγγίσεις είναι καλύτερες από την state-of-the-art προσέγγιση HAN-LIWC.

Μπορούμε οπτικά να επαληθεύσουμε ότι η μέθοδος HTL υπερβαίνει τη μέθοδο INIT-TL στις περισσότερες ρυθμίσεις, παρέχοντας επίσης πιο ισχυρά αποτελέσματα. Εκτός από την επικύρωση ότι τα ψυχολογικά χαρακτηριστικά (LIWC) βελτιώνουν τη δύναμη πρόβλεψης του μοντέλου μας, τεκμηριώνουμε την υπόθεση ότι τόσο οι πληροφορίες προθέσεων όσο και οι συναισθηματικές συνεισφορές συμβάλλουν στην πρόβλεψη των χαρακτηριστικών προσωπικότητας του συγγραφέα / vlogger. Όταν συγκρίνουμε την απόδοση των μοντέλων μας στα δύο σύνολα δεδομένων, παρατηρούμε ότι οι πληροφορίες πρόθεσης συμβάλλουν σημαντικά στην πρόβλεψη της προσωπικότητας των συγγραφέων (σύνολο δεδομένων stream-of-consciousness essay), ενώ για το σύνολο δεδομένων του YouTube η συναισθηματική πηγή είναι η κυρίαρχη πηγή πληροφοριών.

Επιπλέον, παρατηρούμε ότι η ενσωμάτωση τόσο των συναισθημάτων όσο και των πληροφοριών πρόθεσης μέσω της προεκαίδευσης πολλαπλών εργασιών, επιτυγχάνει τα πιο συνεπή αποτελέσματα. Για το σύνολο δεδομένων stream-of-consciousness essay, οι δύο τύποι πληροφοριών φαίνονται να αλληλοσυμπληρώνονται μεταξύ τους. Αυτό γίνεται περισσότερο εμφανές για τη μέθοδο INIT-TL (Σχήμα 4.2a). Ωστόσο, για τη μέθοδο HTL τα κέρδη απόδοσης μειώνονται και σε ορισμένες περιπτώσεις η προσθήκη μεμονωμένων πηγών πληροφόρησης ξεπερνά τον συνδυασμό και των δύο. Αυτό μπορεί να αποδοθεί στο γεγονός ότι διατηρώντας τον κωδικοποιητή παγωμένο, ξεπερνάμε τις επιπτώσεις του προβλήματος του catastrophic forgetting, διαφυλάσσοντας τη γνώση από το source task. Αυτό το φαινόμενο είναι πιο εμφανές σε περιπτώσεις όπου υπάρχει ισχυρή συσχέτιση μεταξύ ενός χαρακτηριστικού προσωπικότητας και μιας συγκεκριμένης πηγής πληροφοριών. Ενώ οι σχετικές πληροφορίες αραιώνονται στη διαδικασία πολλαπλών εργασιών προπαρασκευής.

4.6 Συμπεράσματα

Σε αυτή την εργασία διερευνάμε την αναγνώριση προσωπικότητας σε επίπεδο εγγράφου στο πλαίσιο μεταφοράς γνώσης (Transfer Learning - TL). Προτείνουμε δύο μεθόδους TL για να χρησιμοποιήσουμε πληροφορίες σχετικές με τα συναισθήματα και τις προθέσεις του ατόμου σε επίπεδο προτάσεων, βελτιώνοντας την απόδοση του μοντέλου. Τα μοντέλα αξιολογούνται στα σύνολα δεδομένων του YouTube και stream-of-consciousness essay, αποδίδοντας state-of-the-art αποτελέσματα. Η ανάλυση των αποτελεσμάτων μας δείχνει ότι οι συναισθηματικές, προθετικές και ψυχολογικές πληροφορίες συμβάλλουν στην αναγνώριση της προσωπικότητας. Διαφορετικοί τύποι πληροφοριών κυριαρχούν περισσότερο για συγκεκριμένα χαρακτηριστικά και σύνολα δεδομένων. Ο συνδυασμός των πληροφοριών συναισθήματος και πρόθεσης στο πλαίσιο μάθησης πολλαπλών εργασιών προσφέρει περαιτέρω αύξηση της απόδοσης, ειδικά στην περίπτωση του συνόλου δεδομένων stream-of-consciousness essay.

Σε μελλοντικές εργασίες, σχεδιάζουμε να διερευνήσουμε πιο εξελιγμένες προσεγγίσεις TL για τη μεταφορά γνώσεων από προβλήματα επιπέδου πρότασης σε προβλήματα σε επίπεδο εγγράφου. Θα διερευνήσουμε περαιτέρω τον συνδυασμό των προτεινόμενων μεθόδων TL με χαρακτηριστικά που βασίζονται σε λεξικά για την αναγνώριση της προσωπικότητας και θα ενσωματώσουμε πρόσθετες πηγές πληροφοριών.

Κεφάλαιο 5

Συμπεράσματα

Σε αυτό το έργο, ερευνούμε μεθόδους βαθιάς μάθησης για την πρόβλεψη της αναγνώρισης πρόθεσης και προσωπικότητας των ομιλητών. Και τα δύο μοντέλα μπορούν να ενσωματωθούν σε ένα σύστημα διαλόγου (Dialogue System - DS) για τη βελτίωση της αλληλεπίδρασης ανθρώπου-μηχανής. Συγκεκριμένα, η αναγνώριση των διαλογικών πράξεων (Dialogue Acts - DAs) κάθε φράσης του ομιλητή συμβάλλει στη δημιουργία κατάλληλης απάντησης από το DS. Καταλαβαίνοντας τις προθέσεις του ομιλητή (π.χ. ενημέρωση, ερώτηση, ανατροφοδότηση, backchannel κλπ.) κάθε στιγμή, τα DSs μπορούν να περιορίσουν τον αριθμό των επιλογών που παράγονται ως πιθανές απαντήσεις. Ενώ η αναγνώριση προθέσεων μπορεί να βελτιώσει την αλληλεπίδραση ανθρώπου-μηχανής, η αναγνώριση της προσωπικότητας ενός ατόμου μπορεί να συμβάλει σε πιο εξατομικευμένα DSs. Αν ένα DS είναι σε θέση να αναγνωρίσει με επιτυχία τα χαρακτηριστικά της προσωπικότητας κάθε χρήστη, θα προσαρμόσει ανάλογα τη συμπεριφορά του. Επιπλέον, τα συγκεκριμένα χαρακτηριστικά προσωπικότητας μπορούν να υιοθετηθούν από το DS προκειμένου να αποδοθεί μια συγκεκριμένη «προσωπικότητα» στο σύστημα. Στο πλαίσιο αυτής της διατριβής, προτείνουμε δύο διαφορετικά μοντέλα για την πρόβλεψη των προθέσεων και των χαρακτηριστικών προσωπικότητας των ομιλητών, αντίστοιχα.

Πρώτον, εφαρμόζουμε ένα μοντέλο βαθιάς μάθησης για την ταξινόμηση των διαλογικών πράξεων για κάθε φράση στο πλαίσιο των διαλόγων. Το μοντέλο μας αποτελείται από ένα δίκτυο LSTM για την παράσταση της κάθε φράσης και, στη συνέχεια, ένα απλό Feed-forward Neural Network (FFNN) για την ταξινόμηση της τρέχουσας φράσης με βάση την αναπαράστασή της καθώς και εκείνες των δύο προηγούμενων φράσεων στα πλαίσια του διαλόγου. Το FFNN στοχεύει στην εκμετάλλευση της ιστορίας του διαλόγου προτού προβλέψει την DA της τρέχουσας φράσης. Η συμβολή μας σε αυτό το μοντέλο είναι η επέκταση των γενικών διανυσμάτων λέξεων (word embeddings) που τροφοδοτούνται στο δίκτυο. Συγκεκριμένα, ενσωματώνουμε ειδικές πληροφορίες DA στα γενικά word embeddings. Δοκιμάζουμε την προτεινόμενη μέθοδο χρησιμοποιώντας το corpus Switchboard-DAMSL, το οποίο επισημαίνεται με 42 ετικέτες DA και επιτυγχάνουμε αποτελέσματα συγκρίσιμα με τα αντίστοιχα state-of-the-art.

Δεύτερον, προτείνουμε δύο μεθόδους μεταφοράς μάθησης (Transfer Learning - TL) για την αναγνώριση προσωπικότητας σε επίπεδο εγγράφου. Το κίνητρό μας για τη χρήση του TL είναι ότι τα χαρακτηριστικά της προσωπικότητας συσχετίζονται με άλλες πληροφορίες, όπως το συναίσθημα, η πρόθεση και η ψυχολογολογία. Στην πραγματικότητα, εξ όσων γνωρίζουμε, εισάγουμε για πρώτη φορά την πρόθεση ως σχετική πηγή πληροφοριών στο πρόβλημα της αναγνώρισης της προσωπικότητας. Προηγούμενες προσεγγίσεις στην αναγνώριση της προσωπικότητας έχουν χρησιμοποιήσει αυτές τις σχετικές πληροφορίες για τη βελτίωση της απόδοσης των μοντέλων τους με τη μορφή χειροποίητων χαρακτηριστικών και λεξικών. Σε αυτήν την εργασία, προτείνουμε μια καινοφανή προσαρμογή δύο πολύ γνωστών μεθόδων TL για τη μεταφορά γνώσεων από εργασίες πηγής σε επίπεδο πρότασης στην εργασία στόχου σε επίπεδο εγγράφου. Το μοντέλο μας βασίζεται σε δίκτυα ιεραρχικής προσοχής. Εκπαιδευούμε ένα μοντέλο σε ένα πρόβλημα αναγνώρισης πρόθεσης, συναισθήματος ή και στα δύο με τη μάθηση πολλαπλών εργασιών. Στη συνέχεια, χρησιμοποιούμε τον κωδικοποιητή του προεκπαιδευμένου μοντέλου για την αρχικοποίηση του κωδικοποιητή προτάσεων του μοντέλου μας κα-

θώς και ως εξολκεία χαρακτηριστικών επιπέδου προτάσεων. Αξιολογούμε την προσέγγισή μας σε δύο σύνολα δεδομένων προσωπικότητας, συγκεκριμένα στα stream-of-consciousness essay και YouTube και επιτυγχάνουμε state-of-the-art αποτελέσματα. Επιπλέον, διεξάγουμε μια ανάλυση σχετικά με τη συμβολή των διαφόρων πηγών πληροφοριών (δηλαδή, συναίσθημα, πρόθεση, ψυχολογολογία) και αποδεικνύουμε ότι όλες οι πληροφορίες μπορούν να βελτιώσουν την πρόβλεψη της προσωπικότητας του ατόμου. Ένας διαφορετικός τύπος πληροφορίας είναι ο πιο κυρίαρχος για κάθε χαρακτηριστικό προσωπικότητας και σύνολο δεδομένων.

Στο μέλλον, μπορούμε να βελτιώσουμε το μοντέλο ταξινόμησης DA, αντικαθιστώντας το FFNN, το οποίο χρησιμοποιείται για την αξιοποίηση πληροφοριών σχετικών με το ιστορικό του διαλόγου, με ένα μοντέλο LSTM για τη διατήρηση μακράς βραχυπρόθεσμης ενημέρωσης των διαδοχικών πράξεων. Επίσης, θα αξιολογήσουμε την προτεινόμενη προσέγγιση σε άλλα σύνολα δεδομένων που χρησιμοποιούνται συνήθως για ταξινόμηση DA. Επιπλέον, για το πρόβλημα αναγνώρισης της προσωπικότητας, μπορούμε να πειραματιστούμε με περισσότερες πηγές πληροφοριών, εκτός από το συναίσθημα, την πρόθεση και την ψυχολογολογία, για την προεκπαίδευση του κωδικοποιητή προτάσεων του δικτύου μας. Επιπλέον, μπορούμε να δοκιμάσουμε την ενσωμάτωση των ψυχολογολογικών χαρακτηριστικών εξαγόμενων από λεξικά στην TL προσέγγισή μας.

Τέλος, τα μεμονωμένα προτεινόμενα μοντέλα μπορούν να συνδυαστούν στα πλαίσια ενός DS μαζί με άλλα πιθανά χαρακτηριστικά (π.χ., ενσυναίσθηση ή αναγνώριση δέσμευσης). Μπορούμε επίσης να υλοποιήσουμε ένα μοντέλο παραγωγής λόγου που να εξαρτάται από αυτά τα χαρακτηριστικά για την παραγωγή κατάλληλων απαντήσεων. Αυτό θα βοηθήσει στην εξατομίκευση και προσαρμογή των DSs σε μεμονωμένους χρήστες. Η συμβολή τέτοιων χαρακτηριστικών στην απόδοση ενός DS μπορεί να αξιολογηθεί με τη σύγκρισή του με το αντίστοιχο baseline σύστημα.

Βιβλιογραφία

- [1] J. M. Digman and J. Inouye, “Further specification of the five robust factors of personality.” *Journal of personality and social psychology*, vol. 50, no. 1, p. 116, 1986.
- [2] R. R. McCrae and P. T. Costa Jr, “A five-factor theory of personality,” *Handbook of personality: Theory and research*, vol. 2, pp. 139–153, 1999.
- [3] W. Mischel, “Toward a cognitive social learning reconceptualization of personality.” *Psychological review*, vol. 80, no. 4, p. 252, 1973.
- [4] C. Sarkar, S. Bhatia, A. Agarwal, and J. Li, “Feature analysis for computational personality recognition using youtube personality data set,” in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 11–14.
- [5] G. Farnadi, G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M.-F. Moens, and M. De Cock, “Computational personality recognition in social media,” *User modeling and user-adapted interaction*, vol. 26, no. 2-3, pp. 109–142, 2016.
- [6] N. Majumder, S. Poria, A. Gelbukh, and E. Cambria, “Deep learning-based document modeling for personality detection from text,” *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74–79, 2017.
- [7] J. R. Searle, *The construction of social reality*. Simon and Schuster, 1995.
- [8] M. Tomasello and M. Carpenter, “Shared intentionality,” *Developmental science*, vol. 10, no. 1, pp. 121–125, 2007.
- [9] D. Jacquette, “Brentano’s concept of intentionality,” *The Cambridge Companion to Brentano*, pp. 98–130, 2004.
- [10] F. Dretske, “The explanatory role of content,” 1988.
- [11] J. Conklin, “Dialog mapping: Reflections on an industrial strength case study,” in *Visualizing argumentation*. Springer, 2003, pp. 117–136.
- [12] V. E. Van Reijswoud, “The structure of business communication: theory, model and application,” Ph.D. dissertation, TU Delft, Delft University of Technology, 1996.
- [13] F. Dignum and H. Weigand, “Communication and deontic logic,” 1995.
- [14] H. H. Clark, *Using language*. Cambridge university press, 1996.
- [15] H. Weigand, F. Van Der Poll, A. De Moor *et al.*, “Coordination through communication,” in *Proc. of the 8th International Working Conference on the Language-Action Perspective on Communication Modelling (LAP 2003)*, 2003, pp. 1–2.
- [16] R. Stamper, K. Liu, M. Hafkamp, and Y. Ades, “Understanding the roles of signs and norms in organizations—a semiotic approach to information systems design,” *Behaviour & Information Technology*, vol. 19, no. 1, pp. 15–27, 2000.

- [17] J. Groenendijk and M. Stokhof, "Dynamic predicate logic," *Linguistics and philosophy*, vol. 14, no. 1, pp. 39–100, 1991.
- [18] J. L. Austin, *How to do things with words*. Oxford university press, 1975.
- [19] J. R. Searle, *Speech acts: An essay in the philosophy of language*. Cambridge university press, 1969, vol. 626.
- [20] D. Vanderveken, "Meaning and speech acts, vol. 2. formal semantics of success and satisfaction," 1991.
- [21] H. Bunt, "Interaction management functions and context representation requirements," in *Proceedings of the Twente Workshop on Language Technology: Dialogue Management in Natural Language Systems (TWLT 11)*, 1996, pp. 187–198.
- [22] M. Poesio and D. R. Traum, "Conversational actions and discourse situations," *Computational intelligence*, vol. 13, no. 3, pp. 309–347, 1997.
- [23] M. Poesio and D. Traum, "Towards an axiomatization of dialogue acts," in *Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues (13th Twente Workshop on Language Technology)*. Citeseer, 1998.
- [24] T. Winograd and F. Flores, *Understanding computers and cognition: A new foundation for design*. Intellect Books, 1986.
- [25] E. Bilange, "A task independent oral dialogue model," in *Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 1991, pp. 83–88.
- [26] J. Carletta, S. Isard, G. Doherty-Sneddon, A. Isard, J. C. Kowtko, and A. H. Anderson, "The reliability of a dialogue structure coding scheme," *Computational linguistics*, vol. 23, no. 1, pp. 13–31, 1997.
- [27] H. Alexandersson, H. Tuomenvirta, T. Schmith, and K. Iden, "Trends of storms in nw europe derived from an updated pressure data set," *Climate Research*, vol. 14, no. 1, pp. 71–73, 2000.
- [28] C. L. Sidner, "An artificial discourse language for collaborative negotiation," in *AAAI*, vol. 94, 1994, pp. 814–819.
- [29] F. Fipa, "specification part 2: Agent communication language," Technical report, FIPA-Foundation for Intelligent Physical Agents, Tech. Rep., 1997.
- [30] N. J. Allen and J. P. Meyer, "Affective, continuance, and normative commitment to the organization: An examination of construct validity," *Journal of vocational behavior*, vol. 49, no. 3, pp. 252–276, 1996.
- [31] P. Bretier and D. Sadek, "A rational agent as the kernel of a cooperative spoken dialogue system: Implementing a logical theory of interaction," in *International Workshop on Agent Theories, Architectures, and Languages*. Springer, 1996, pp. 189–203.
- [32] M. D. Sadek, "Dialogue acts are rational plans," in *The Structure of Multimodal Dialogue; Second VENACO Workshop*, 1991.

- [33] D. R. Initiative *et al.*, “Standards for dialogue coding in natural language processing,” Report, Tech. Rep., 1997.
- [34] J. Allwood, “Dialog as collective thinking,” in *New Directions in Cognitive Science. Publications of the Finnish Artificial Intelligence Society. International Conferences*, no. 2, 1997, pp. 222–226.
- [35] P. R. Cohen and H. J. Levesque, “Intention is choice with commitment,” *Artificial intelligence*, vol. 42, no. 2-3, pp. 213–261, 1990.
- [36] D. R. Traum and J. F. Allen, “A” speech acts” approach to grounding in conversation,” in *Second International Conference on Spoken Language Processing*, 1992.
- [37] D. Traum, J. Bos, R. Cooper, S. Larsson, I. Lewin, C. Matheson, and M. Poesio, “A model of dialogue moves and information state revision,” Tech. rept. Deliverable, Tech. Rep., 1999.
- [38] J. F. Allen and C. R. Perrault, “Analyzing intention in utterances,” *Artificial intelligence*, vol. 15, no. 3, pp. 143–178, 1980.
- [39] J. Allwood, “Obligations and options in dialogue,” *Think Quarterly*, vol. 3, pp. 9–18, 1994.
- [40] T. H. Bui, M. Rajman, and M. Melichar, “Rapid dialogue prototyping methodology,” in *International Conference on Text, Speech and Dialogue*. Springer, 2004, pp. 579–586.
- [41] J. Allen and M. Core, “Draft of DAMSL: Dialog act markup in several layers,” 1997.
- [42] D. Jurafsky, E. Shriberg, and D. Biasca, “Switchboard SWBD-DAMSL shallow-discourse-function annotation coders manual,” *Institute of Cognitive Science Technical Report*, pp. 97–102, 1997.
- [43] J. Ang, Y. Liu, and E. Shriberg, “Automatic dialog act segmentation and classification in multiparty meetings,” in *Proceedings of the ICASSP*, vol. 1, 2005, pp. I/1061–I/1064.
- [44] T. Bub and J. Schwinn, “Verbmobil: The evolution of a complex large speech-to-speech translation system,” in *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, vol. 4. IEEE, 1996, pp. 2371–2374.
- [45] A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller *et al.*, “The herc map task corpus,” *Language and speech*, vol. 34, no. 4, pp. 351–366, 1991.
- [46] J. Weizenbaum, “Eliza—a computer program for the study of natural language communication between man and machine,” *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.
- [47] G. W. Allport, “Pattern and growth in personality.” 1961.
- [48] R. Weinberg and D. Gould, “Foundations of exercise and sport psychology,” 1999.
- [49] H. J. Eysenck, “The science of personality: Nomothetic.” 1954.
- [50] A. Bandura and R. H. Walters, *Social learning theory*. Prentice-hall Englewood Cliffs, NJ, 1977, vol. 1.
- [51] S. Freud, *The ego and the id*. Courier Dover Publications, 2018.

- [52] H. J. Eysenck, "The scientific study of personality." 1952.
- [53] H. Eysenck, *The biological basis of personality*. Routledge, 2017.
- [54] H. J. Eysenck, "Personality, genetics, and behavior: Selected papers," 1982.
- [55] H. Eysenck, "Student selection by means of psychological tests—a critical survey," *British Journal of Educational Psychology*, vol. 17, no. 1, pp. 20–39, 1947.
- [56] ———, "Personality and experimental psychology." *Bulletin of the British Psychological Society*, 1966.
- [57] R. Cattell, "The scientific study of personality," *Harmondsworth: Penguin*, vol. 252, pp. 76–83, 1965.
- [58] G. W. Allport, "Personality: A psychological interpretation." 1937.
- [59] E. C. Tupes and R. E. Christal, "Recurrent personality factors based on trait ratings," *Journal of personality*, vol. 60, no. 2, pp. 225–251, 1992.
- [60] J. M. Digman, "Personality structure: Emergence of the five-factor model," *Annual review of psychology*, vol. 41, no. 1, pp. 417–440, 1990.
- [61] L. R. Goldberg, "The structure of phenotypic personality traits." *American psychologist*, vol. 48, no. 1, p. 26, 1993.
- [62] T. J. Trull and D. C. Geary, "Comparison of the big-five factor structure across samples of chinese and american adults," *Journal of Personality Assessment*, vol. 69, no. 2, pp. 324–341, 1997.
- [63] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. Watkins, "Text classification using string kernels," *Journal of Machine Learning Research*, vol. 2, no. Feb, pp. 419–444, 2002.
- [64] M. H. Bond, H. Nakazato, and D. Shiraishi, "Universality and distinctiveness in dimensions of japanese person perception," *Journal of Cross-Cultural Psychology*, vol. 6, no. 3, pp. 346–357, 1975.
- [65] F. M. Cheung, F. J. van de Vijver, and F. T. Leong, "Toward a new approach to the study of personality in culture." *American Psychologist*, vol. 66, no. 7, p. 593, 2011.
- [66] J. Oberlander and A. J. Gill, "Individual differences and implicit language: personality, parts-of-speech and pervasiveness," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 26, no. 26, 2004.
- [67] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *Journal of artificial intelligence research*, vol. 30, pp. 457–500, 2007.
- [68] C. G. DeYoung, "Personality neuroscience and the biology of traits," *Social and Personality Psychology Compass*, vol. 4, no. 12, pp. 1165–1180, 2010.
- [69] J. S. Adelstein, Z. Shehzad, M. Mennes, C. G. DeYoung, X.-N. Zuo, C. Kelly, D. S. Margulies, A. Bloomfield, J. R. Gray, F. X. Castellanos *et al.*, "Personality is reflected in the brain's intrinsic functional architecture," *PloS one*, vol. 6, no. 11, p. e27633, 2011.

- [70] S. Dhuria, “Natural language processing: An approach to parsing and semantic analysis.”
- [71] C. D. Manning and H. Schütze, *Foundations of statistical natural language processing*. MIT press, 1999.
- [72] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,, 2016.
- [73] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, “Playing atari with deep reinforcement learning,” *arXiv preprint arXiv:1312.5602*, 2013.
- [74] M. Hu and B. Liu, “Mining opinion features in customer reviews,” in *AAAI*, vol. 4, no. 4, 2004, pp. 755–760.
- [75] G. Somprasertsri and P. Lalitrojwong, “Automatic product feature extraction from online product reviews using maximum entropy with lexical and syntactic features,” in *Information Reuse and Integration, 2008. IRI 2008. IEEE International Conference on*. IEEE, 2008, pp. 250–255.
- [76] N. Kobayashi, K. Inui, and Y. Matsumoto, “Extracting aspect-evaluation and aspect-of relations in opinion mining,” in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, 2007.
- [77] A. Swami, A. Mete, S. Bhosle, N. Nimbalkar, and S. Kale, “Feature extraction and refinement for opinion mining,” 2017.
- [78] G. Mishne *et al.*, “Experiments with mood classification in blog posts,” in *Proceedings of ACM SIGIR 2005 workshop on stylistic analysis of text for information access*, vol. 19, 2005, pp. 321–327.
- [79] W. Zhang, H. Xu, and W. Wan, “Weakness finder: Find product weakness from chinese reviews by using aspects based sentiment analysis,” *Expert Systems with Applications*, vol. 39, no. 11, pp. 10 283–10 291, 2012.
- [80] S. Stymne, “Pre-and postprocessing for statistical machine translation into germanic languages,” in *Proceedings of the ACL 2011 Student Session*. Association for Computational Linguistics, 2011, pp. 12–17.
- [81] E. Lloret, H. Saggion, and M. Palomar, “Experiments on summary-based opinion classification,” in *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*. Association for Computational Linguistics, 2010, pp. 107–115.
- [82] C. C. Aggarwal and C. Zhai, *Mining text data*. Springer Science & Business Media, 2012.
- [83] W. Wei, H. Liu, J. He, H. Yang, and X. Du, “Extracting feature and opinion words effectively from chinese product reviews,” in *Fuzzy Systems and Knowledge Discovery, 2008. FSKD’08. Fifth International Conference on*, vol. 4. IEEE, 2008, pp. 170–174.
- [84] R. Dale, H. Moisl, and H. Somers, *Handbook of natural language processing*. CRC Press, 2000.

- [85] S. Bird and E. Loper, “Nltk: the natural language toolkit,” in *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions*. Association for Computational Linguistics, 2004, p. 31.
- [86] D. Ploch, “Exploring entity relations for named entity disambiguation,” in *Proceedings of the ACL 2011 Student Session*. Association for Computational Linguistics, 2011, pp. 18–23.
- [87] R. R. Larson, “Introduction to information retrieval,” *Journal of the American Society for Information Science and Technology*, vol. 61, no. 4, pp. 852–853, 2010.
- [88] H. P. Luhn, “The automatic creation of literature abstracts,” *IBM Journal of research and development*, vol. 2, no. 2, pp. 159–165, 1958.
- [89] J. Savoy, “Ir multilingual resources at unine,” URL: <http://members.unine.ch/jacques.savoy/clef/index.html> [Stand: 10.04. 2014], 2011.
- [90] H. Liu and H. Motoda, *Computational methods of feature selection*. CRC Press, 2007.
- [91] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics New York, 2001, vol. 1.
- [92] L. Song, A. Smola, A. Gretton, K. M. Borgwardt, and J. Bedo, “Supervised feature selection via dependence estimation,” in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, pp. 823–830.
- [93] J. Weston, A. Elisseeff, B. Schölkopf, and M. Tipping, “Use of the zero-norm with linear models and kernel methods,” *Journal of machine learning research*, vol. 3, no. Mar, pp. 1439–1461, 2003.
- [94] P. Mitra, C. Murthy, and S. K. Pal, “Unsupervised feature selection using feature similarity,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 3, pp. 301–312, 2002.
- [95] J. G. Dy and C. E. Brodley, “Feature selection for unsupervised learning,” *Journal of machine learning research*, vol. 5, no. Aug, pp. 845–889, 2004.
- [96] Z. Xu, I. King, M. R.-T. Lyu, and R. Jin, “Discriminative semi-supervised feature selection via manifold regularization,” *IEEE Transactions on Neural networks*, vol. 21, no. 7, pp. 1033–1047, 2010.
- [97] Z. Zhao and H. Liu, “Semi-supervised feature selection via spectral analysis,” in *Proceedings of the 2007 SIAM international conference on data mining*. SIAM, 2007, pp. 641–646.
- [98] M. Robnik-Šikonja and I. Kononenko, “Theoretical and empirical analysis of relieff and rrelieff,” *Machine learning*, vol. 53, no. 1-2, pp. 23–69, 2003.
- [99] R. O. Duda, P. E. Hart, D. G. Stork *et al.*, *Pattern classification*. Wiley New York, 1973, vol. 2.
- [100] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.
- [101] H. Liu and L. Yu, “Toward integrating feature selection algorithms for classification and clustering,” *IEEE Transactions on knowledge and data engineering*, vol. 17, no. 4, pp. 491–502, 2005.

- [102] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.
- [103] G. C. Cawley, N. L. Talbot, and M. Girolami, “Sparse multinomial logistic regression via bayesian l1 regularisation,” in *Advances in neural information processing systems*, 2007, pp. 209–216.
- [104] J. G. Dy and C. E. Brodley, “Feature subset selection and order identification for unsupervised learning,” in *ICML*. Citeseer, 2000, pp. 247–254.
- [105] S. Alelyani, J. Tang, and H. Liu, “Feature selection for clustering: A review.” *Data Clustering: Algorithms and Applications*, vol. 29, pp. 110–121, 2013.
- [106] T. Hastie and R. Tibshirani, “Discriminant adaptive nearest neighbor classification and regression,” in *Advances in Neural Information Processing Systems*, 1996, pp. 409–415.
- [107] C. Domeniconi and D. Gunopulos, “Local feature selection for classification,” *Computational Methods of Feature Selection*, p. 211, 2007.
- [108] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz, “A bayesian approach to filtering junk e-mail,” in *Learning for Text Categorization: Papers from the 1998 workshop*, vol. 62, 1998, pp. 98–105.
- [109] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in *Advances in neural information processing systems*, 2013, pp. 3111–3119.
- [110] J. Pennington, R. Socher, and C. Manning, “Glove: Global vectors for word representation,” in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.
- [111] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [112] Y. Nesterov, “A method of solving a convex programming problem with convergence rate $o(1/k^2)$,” in *Soviet Mathematics Doklady*, vol. 27, no. 2, 1983, pp. 372–376.
- [113] J. Duchi, E. Hazan, and Y. Singer, “Adaptive subgradient methods for online learning and stochastic optimization,” *Journal of Machine Learning Research*, vol. 12, no. Jul, pp. 2121–2159, 2011.
- [114] M. D. Zeiler, “Adadelta: an adaptive learning rate method,” *arXiv preprint arXiv:1212.5701*, 2012.
- [115] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [116] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *nature*, vol. 323, no. 6088, p. 533, 1986.
- [117] A. Graves, “Generating sequences with recurrent neural networks,” *arXiv preprint arXiv:1308.0850*, 2013.

- [118] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” *arXiv preprint arXiv:1409.0473*, 2014.
- [119] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [120] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *arXiv preprint arXiv:1301.3781*, 2013.
- [121] H. Bunt, J. Alexandersson, J.-W. Choe, A. C. Fang, K. Hasida, V. Petukhova, A. Popescu-Belis, and D. R. Traum, “Iso 24617-2: A semantically-based standard for dialogue annotation.” in *LREC*. Citeseer, 2012, pp. 430–437.
- [122] H. Bunt, “The dit++ taxonomy for functional dialogue markup,” in *AAMAS 2009 Workshop, Towards a Standard Markup Language for Embodied Dialogue Acts*, 2009, pp. 13–24.
- [123] J. R. Searle, *Speech acts: An essay in the philosophy of language*, 1969, vol. 626.
- [124] A. Stolcke, K. Ries, N. Coccaro, E. Shriberg, R. Bates, D. Jurafsky, P. Taylor, R. Martin, C. Van Ess-Dykema, and M. Meteer, “Dialogue act modeling for automatic tagging and recognition of conversational speech,” *Computational linguistics*, vol. 26, no. 3, pp. 339–373, 2000.
- [125] Y. Ji, G. Haffari, and J. Eisenstein, “A latent variable recurrent neural network for discourse relation language models,” *arXiv preprint arXiv:1603.01913*, 2016.
- [126] R. Fernandez and R. W. Picard, “Dialog act classification from prosodic features using support vector machines,” in *Speech Prosody 2002, International Conference*, 2002.
- [127] N. Kalchbrenner and P. Blunsom, “Recurrent convolutional neural networks for discourse compositionality,” *arXiv preprint arXiv:1306.3584*, 2013.
- [128] J. Y. Lee and F. Deroncourt, “Sequential short-text classification with recurrent and convolutional neural networks,” *arXiv preprint arXiv:1603.03827*, 2016.
- [129] H. Khanpour, N. Guntakandla, and R. Nielsen, “Dialogue Act Classification in Domain-Independent Conversations Using a Deep Recurrent Neural Network,” *Proceedings of COLING*, pp. 2012–2021, 2016.
- [130] A. Qadir and E. Riloff, “Classifying sentences as speech acts in message board posts,” in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2011, pp. 748–758.
- [131] D. Jurafsky, R. Bates, N. Coccaro, R. Martin, M. Meteer, K. Ries, E. Shriberg, A. Stolcke, P. Taylor, and C. V. Ess-Dykema, “Automatic detection of discourse structure for speech recognition and understanding,” *1997 IEEE Workshop on Speech Recognition and Understanding*, pp. 88–95, 1997.
- [132] R. Klaus, C. Noah, S. Elizabeth, B. Rebecca, J. Daniel, T. Paul, M. Rachel, V. E.-D. Carol, V. E.-D. Carol, and M. Marie, “Automatic detection of discourse structure for speech recognition and understanding,” *1997 IEEE Workshop on Speech Recognition and Understanding*, pp. 88–95, 1997.

- [133] S. N. Kim, L. Cavedon, and T. Baldwin, “Classifying dialogue acts in one-on-one live chats,” in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, 2010, pp. 862–871.
- [134] N. Novielli and C. Strapparava, “The role of affect analysis in dialogue act identification,” *IEEE Transactions on Affective Computing*, vol. 4, no. 4, pp. 439–451, 2013.
- [135] N. Malandrakis, A. Potamianos, E. Iosif, and S. Narayanan, “Distributional semantic models for affective text analysis,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 11, pp. 2379–2392, 2013.
- [136] S. Kim, D. Luis Fernando, R. E. Banchs, J. Williams, M. Henderson, and K. Yoshino, “Dialog State Tracking Challenge 4: Handbook,” 2015.
- [137] M. Henderson, B. Thomson, and J. Williams, “The second dialog state tracking challenge,” in *15th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, vol. 263, 2014.
- [138] H. Bunt, J. Alexandersson, J.-W. Choe, A. C. Fang, K. Hasida, V. Petukhova, A. Popescu-Belis, and D. R. Traum, “ISO 24617-2: A semantically-based standard for dialogue annotation.” in *Proceedings of LREC*, 2012, pp. 430–437.
- [139] H. Bunt, J. Alexandersson, J. Carletta, J.-W. Choe, A. C. Fang, K. Hasida, K. Lee, V. Petukhova, A. Popescu-Belis, L. Romary *et al.*, “Towards an ISO standard for dialogue act annotation,” in *Proceedings of LREC*, 2010.
- [140] H. Bunt, V. Petukhova, D. Traum, and J. Alexandersson, “Dialogue Act Annotation with the ISO 24617-2 Standard,” in *Multimodal Interaction with W3C Standards*, 2017, pp. 109–135.
- [141] T. Lei, R. Barzilay, and T. Jaakkola, “Molding cnns for text: non-linear, non-consecutive convolutions,” *arXiv preprint arXiv:1508.04112*, 2015.
- [142] A. Graves, N. Jaitly, and A.-r. Mohamed, “Hybrid speech recognition with deep bidirectional LSTM,” in *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on*, 2013, pp. 273–278.
- [143] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” in *Advances in neural information processing systems*, 2014, pp. 3104–3112.
- [144] A. L. Gorin, “Processing of semantic information in fluently spoken language,” in *Proceedings of ICSLP*, vol. 2, 1996, pp. 1001–1004.
- [145] R. Iyer, M. Ostendorf, and H. Gish, “Using out-of-domain data to improve in-domain language models,” *IEEE Signal processing letters*, vol. 4, no. 8, pp. 221–223, 1997.
- [146] D. Povey, P. C. Woodland, and M. J. Gales, “Discriminative MAP for acoustic model adaptation,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. (ICASSP’03).*, vol. 1, 2003, pp. I–I.
- [147] T. Lei, H. Joshi, R. Barzilay, T. Jaakkola, K. Tymoshenko, A. Moschitti, and L. Marquez, “Semi-supervised question retrieval with gated convolutions,” *arXiv preprint arXiv:1512.05726*, 2015.
- [148] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, “Improving neural networks by preventing co-adaptation of feature detectors,” *arXiv preprint arXiv:1207.0580*, 2012.

- [149] M. D. Ekstrand, J. T. Riedl, J. A. Konstan *et al.*, “Collaborative filtering recommender systems,” *Foundations and Trends® in Human–Computer Interaction*, vol. 4, no. 2, pp. 81–173, 2011.
- [150] R. Gao, B. Hao, S. Bai, L. Li, A. Li, and T. Zhu, “Improving user profile with personality traits predicted from social media content,” in *Proceedings of the 7th ACM conference on Recommender systems*. ACM, 2013, pp. 355–358.
- [151] A. Aly and A. Tapus, “Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction,” *Autonomous Robots*, vol. 40, no. 2, pp. 193–209, 2016.
- [152] P. Fung, A. Dey, F. B. Siddique, R. Lin, Y. Yang, Y. Wan, and H. Y. R. Chan, “Zara the supergirl: An empathetic personality recognition system,” in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, 2016, pp. 87–91.
- [153] C. S. Carver, S. K. Sutton, and M. F. Scheier, “Action, emotion, and personality: Emerging conceptual integration,” *Personality and social psychology bulletin*, vol. 26, no. 6, pp. 741–751, 2000.
- [154] R. M. Ryan, “A motivational approach to self: Integration in personality edward l., deci and,” *Perspectives on motivation*, vol. 38, no. 237, pp. 237–288, 1991.
- [155] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “CNN features off-the-shelf: an astounding baseline for recognition,” *CoRR*, vol. abs/1403.6382, 2014.
- [156] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” *CoRR*, vol. abs/1411.4038, 2014.
- [157] B. Felbo, A. Misllove, A. Søgaard, I. Rahwan, and S. Lehmann, “Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm,” in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics, September 2017, pp. 1615–1625. [Online]. Available: <https://www.aclweb.org/anthology/D17-1169>
- [158] J. Howard and S. Ruder, “Universal Language Model Fine-tuning for Text Classification,” *ArXiv e-prints*, Jan. 2018.
- [159] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, “Hierarchical attention networks for document classification,” in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 1480–1489.
- [160] D. Xue, Z. Hong, S. Guo, L. Gao, L. Wu, J. Zheng, and N. Zhao, “Personality recognition on social media with label distribution learning,” *IEEE Access*, vol. 5, pp. 13 478–13 488, 2017.
- [161] J.-I. Biel and D. Gatica-Perez, “The youtube lens: Crowdsourced personality impressions and audiovisual analysis of vlogs,” *IEEE Transactions on Multimedia*, vol. 15, no. 1, pp. 41–55, 2013.
- [162] J. W. Pennebaker and L. A. King, “Linguistic styles: Language use as an individual difference,” *Journal of personality and social psychology*, vol. 77, no. 6, p. 1296, 1999.

- [163] S. Poria, A. Gelbukh, B. Agarwal, E. Cambria, and N. Howard, “Common sense knowledge based personality recognition from text,” in *Mexican International Conference on Artificial Intelligence*. Springer, 2013, pp. 484–496.
- [164] M. P. Kalghatgi, M. Ramannavar, and D. N. S. Sidnal, “Social-network-sourced big data analytics for personality prediction: A review,” 2015.
- [165] J. W. Pennebaker, M. E. Francis, and R. J. Booth, “Linguistic inquiry and word count: Liwc 2001,” *Mahway: Lawrence Erlbaum Associates*, vol. 71, no. 2001, p. 2001, 2001.
- [166] S. M. Mohammad and S. Kiritchenko, “Using hashtags to capture fine emotion categories from tweets,” *Computational Intelligence*, vol. 31, no. 2, pp. 301–326, 2015.
- [167] F. Liu, J. Perez, and S. Nowson, “A language-independent and compositional model for personality trait recognition from short texts,” *arXiv preprint arXiv:1610.04345*, 2016.
- [168] S. Gievska and K. Koroveshevski, “The impact of affective verbal content on predicting personality impressions in youtube videos,” in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 19–22.
- [169] M. Coltheart, “The mrc psycholinguistic database,” *The Quarterly Journal of Experimental Psychology Section A*, vol. 33, no. 4, pp. 497–505, 1981.
- [170] M. Wilson, “Mrc psycholinguistic database: Machine-usable dictionary, version 2.00,” *Behavior research methods, instruments, & computers*, vol. 20, no. 1, pp. 6–10, 1988.
- [171] J. Yu and K. Markov, “Deep learning based personality recognition from facebook status updates,” in *Awareness Science and Technology (iCAST), 2017 IEEE 8th International Conference on*. IEEE, 2017, pp. 383–387.
- [172] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “Deepface: Closing the gap to human-level performance in face verification,” 09 2014.
- [173] A. Agrawal, J. Lu, S. Antol, M. Mitchell, C. L. Zitnick, D. Parikh, and D. Batra, “Vqa: Visual question answering,” *Int. J. Comput. Vision*, vol. 123, no. 1, pp. 4–31, May 2017.
- [174] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and F.-F. Li, “Imagenet: a large-scale hierarchical image database,” pp. 248–255, 06 2009.
- [175] C. Baziotis, N. Athanasiou, A. Chronopoulou, A. Kolovou, G. Paraskevopoulos, N. Ellinas, S. Narayanan, and A. Potamianos, “Ntua-slp at semeval-2018 task 1: Predicting affective content in tweets with deep attentive rnns and transfer learning,” *arXiv preprint arXiv:1804.06658*, 2018.
- [176] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [177] B. Hariharan, P. Arbeláez, R. Girshick, and J. Malik, “Hypercolumns for object segmentation and fine-grained localization,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 447–456.
- [178] Y. Li, H. Su, X. Shen, W. Li, Z. Cao, and S. Niu, “Dailydialog: A manually labelled multi-turn dialogue dataset,” *arXiv preprint arXiv:1710.03957*, 2017.

- [179] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting.” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [180] Y. R. Tausczik and J. W. Pennebaker, “The psychological meaning of words: Liwc and computerized text analysis methods,” *Journal of language and social psychology*, vol. 29, no. 1, pp. 24–54, 2010.

Παράρτημα Α

Συντομογραφίες

(DA): Dialogue Act
(SDS): Spoken Dialogue System
(DNN): Deep Neural Network
(LSTM): Long Short-Term Memory
(CNN): Convolutional Neural Network
(AI): Artificial Intelligence
(DS): Dialogue System
(NLP): Natural Language Processing
(ML): Machine Learning
(DL): Deep Learning
(BiLSTM): Bidirectional LSTM
(TL): Transfer Learning
(LAP): Language/Action Perspective
(TPM): Transaction Process Model
(CA): Communicative Act
(RPDM): Rapid Dialogue Prototype Methodology
(ASR): Automatic Speech Recognition
(DAMSL): Dialogue Act Markup in Several Layers
(MRDA): Meeting Recorder DA
(ANS): Automatic Nervous System
(IR): Information Retrieval
(MT): Machine Translation
(IE): Information Extraction
(QA): Question Answering
(POS): Part-Of-Speech
(kNN): k-Nearest Neighbor
(MAP): Maximum A Posteriori
(SVM): Support Vector Machine
(CBOW): Continuous Bag-Of-Words
(ANN): Artificial Neural Network
(MLP): Multi-Layer Perceptron
(FFNN): Feed-Forward Neural Network
(ReLU): Rectified Linear Unit
(GD): Gradient Descent
(SGD): Stochastic Gradient Descent
(RNN): Recurrent Neural Network
(BiRNN): Bidirectional Recurrent Neural Network
(CV): Computer Vision

(INIT): Parameter Initialization
(MULT): MUlti-Task learning
(HMM): Hidden Markov Models
(HAN): Hierarchical Attention Network
(FFM): Five Factor Model
(LIWC): Linguistic Inquiry Word Count
(LM): Language Model
(SE): Sentence Encoder
(INIT-TL): Initialization of sentence encoder
(HTL): Hypercolumns transfer learning of sentence encoder
(MSE): Mean Square Error
(RMSE): Root Mean Square Error