



ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

# Αξιολόγηση και ανίχνευση ψευδών ειδήσεων

Όνοματεπώνυμο: ΘΩΜΟΓΙΑΝΝΟΠΟΥΛΟΥ ΣΠΥΡΙΔΟΥΛΑ

ΑΜ:1693

Όνοματεπώνυμο: ΔΡΙΤΣΑ ΧΡΥΣΟΥΛΑ

ΑΜ:1377

Επιβλέπων Καθηγητής: ΤΡΙΑΝΤΑΦΥΛΛΟΥ ΒΑΣΙΛΕΙΟΣ

Πάτρα, 2020



## ΠΕΡΙΕΧΟΜΕΝΑ

<b>ΠΕΡΙΕΧΟΜΕΝΑ.....</b>	<b>3</b>
<b>ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ.....</b>	<b>5</b>
<b>ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΩΝ.....</b>	<b>6</b>
<b>ΠΕΡΙΛΗΨΗ.....</b>	<b>8</b>
<b>1ΚΕΦΑΛΑΙΟ - ΕΙΣΑΓΩΓΗ.....</b>	<b>9</b>
ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ.....	9
ΣΗΜΑΣΙΑ ΤΩΝ ΕΙΔΗΣΕΩΝ ΣΤΗΝ ΑΝΘΡΩΠΟΤΗΤΑ.....	9
ΨΕΥΔΕΙΣ ΕΙΔΗΣΕΙΣ (FAKE NEWS).....	11
<i>Ορισμός</i> .....	11
<i>Είδη ψευδών ειδήσεων</i> .....	11
<i>Χαρακτηριστικά γνωρίσματα</i> .....	11
ΕΝΤΟΠΙΣΜΟΣ ΤΑΧΕΩΝ ΔΙΑΔΙΔΟΜΕΝΩΝ ΘΕΜΑΤΩΝ ΚΑΙ ΕΝΤΟΠΙΣΜΟΣ ΨΕΥΔΩΝ ΕΙΔΗΣΕΩΝ.....	12
ΦΗΜΕΣ.....	13
<i>Χαρακτηρισμός φημών (rumor characterization)</i> .....	13
<i>Ανίχνευση φημών (rumor detection)</i> .....	14
<i>Περιορισμός φημών (rumor containment)</i> .....	15
<i>Σάτιρα</i> .....	16
ΠΡΟΠΑΓΑΝΔΑ.....	16
<i>Ορισμός</i> .....	16
<i>Χαρακτηριστικά</i> .....	17
<b>2ΚΕΦΑΛΑΙΟ - ΕΡΕΥΝΗΤΙΚΟ ΥΠΟΒΑΘΡΟ STATE OF THE ART.....</b>	<b>18</b>
ΜΕΘΟΔΟΙ ΑΝΙΧΝΕΥΣΗΣ ( DETECTING RUMORS IN SOCIAL MEDIA: A SURVEY).....	18
ΤΕΧΝΙΚΕΣ.....	18
<i>Γλωσσική Ανάλυση (Profile Compatibility)</i> .....	18
Σύνολο των δεδομένων.....	19
Επιλογή κατάλληλων γνωρισμάτων.....	19
Μορφολογική ανάλυση.....	21
Συντακτική ανάλυση.....	21
<i>Διαδικτυακή Διερεύνηση</i> .....	21
<i>Μηχανική Μάθηση (Machine Learning)</i> .....	22
Ιστορική αναδρομή.....	22
Εισαγωγή στη Μηχανική Μάθηση.....	23
Μέθοδοι μηχανικής μάθησης.....	26
Εξόρυξη Δεδομένων & Μηχανική Μάθηση.....	34
<b>3ΚΕΦΑΛΑΙΟ - ΤΕΧΝΟΛΟΓΙΕΣ ΥΛΟΠΟΙΗΣΗΣ.....</b>	<b>35</b>

ΥΠΟΛΟΓΙΣΤΙΚΑ ΕΡΓΑΛΕΙΑ.....	35
ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ (DATA MINING).....	35
ΜΕΤΡΙΚΕΣ ΑΠΟΔΟΣΗΣ.....	36
Πίνακας σύγχυσης ( <i>confusion matrix</i> ).....	37
Ακρίβεια ( <i>Accuracy</i> ).....	38
Ανάκληση ( <i>Recall</i> ).....	39
Ακρίβεια προσέγγισης ( <i>Precision</i> ).....	39
<i>F1- Score</i> .....	40
<b>4ΚΕΦΑΛΑΙΟ – ΥΛΟΠΟΙΗΣΗ.....</b>	<b>42</b>
ΠΡΟΒΛΗΜΑ.....	42
<i>Δεδομένα</i> .....	42
<i>Κατηγοριοποίηση του προβλήματος</i> .....	42
NATURAL LANGUAGE PROCESSING.....	43
<i>Θεωρία Νευρωνικών δικτύων</i> .....	43
<i>Πιθανοτικά Μοντέλα</i> .....	70
<i>Μοντέλα βασισμένα σε Νευρωνικά Δίκτυα</i> .....	71
<i>One Hot Κωδικοποίηση – Count Vectorization</i> .....	71
<i>Νεύρο – Πιθανοτικό Μοντέλο (Neuro Probabilistic)</i> .....	73
<i>TF – IDF Μετασχηματισμός</i> .....	74
<i>Word2Vec</i> .....	75
One – Word Context.....	75
Multi – Word context.....	76
Skip – Gram Μοντέλο.....	77
<i>GloVe (Global Vector for Word Representation)</i> .....	79
<i>FastText</i> .....	80
ΑΝΑΛΥΣΗ ΤΗΣ ΠΡΟΣΕΓΓΙΣΗΣ.....	82
<i>Προ επεξεργασία Κειμένου</i> .....	82
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ.....</b>	<b>91</b>

## ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

<a href="#">Εικόνα 1 - Περιγραφή συστήματος αντίγνωσης φημών</a>	15
<a href="#">Εικόνα 2 - Αποτελέσματα έρευνας για την αναγνώριση των ψευδών ειδήσεων</a>	22
<a href="#">Εικόνα 3 - Διεργασίες που εφαρμόζονται στη Μηχανική Μάθηση</a>	25
<a href="#">Εικόνα 4 - Παράδειγμα υπερεπιπέδου διαχωρισμού</a>	31
<a href="#">Εικόνα 5 - Τεχνητός Νευρώνας</a>	33
<a href="#">Εικόνα 6 - Τεχνητό Νευρωνικό Δίκτυο</a>	34
<a href="#">Εικόνα 7 - Παράδειγμα πίνακα σύγκρισης</a>	39
<a href="#">Εικόνα 8 - Επεξήγηση των Accuracy και Precision</a>	42
<a href="#">Εικόνα 9: Underfit/Overfit</a>	48
<a href="#">Εικόνα 10: Απλή γραμμική παλινδρόμηση</a>	49
<a href="#">Εικόνα 11: Επιρροή των outliers</a>	50
<a href="#">Εικόνα 12: Σιγμοειδής Συνάρτηση</a>	51
<a href="#">Εικόνα 13: kNN Αλγόριθμος</a>	52
<a href="#">Εικόνα 14: Reversed Tree</a>	54
<a href="#">Εικόνα 15: Decision Tree</a>	55
<a href="#">Εικόνα 16: Information gain</a>	56
<a href="#">Εικόνα 17: SVM</a>	57
<a href="#">Εικόνα 18: Neural Network</a>	58
<a href="#">Εικόνα 19: Neuron</a>	59
<a href="#">Εικόνα 20: Σιγμοειδή Συνάρτηση</a>	60
<a href="#">Εικόνα 21: Gradient Descent</a>	62
<a href="#">Εικόνα 22: ReLU</a>	63
<a href="#">Εικόνα 23: Deep Neural Network</a>	65
<a href="#">Εικόνα 24: Συνέλιξη</a>	66
<a href="#">Εικόνα 25: Αναπαράσταση RNN</a>	68

<a href="#">Εικόνα 26: RNN</a>	69
<a href="#">Εικόνα 27: Διαφορετικοί τύποι των RNNs</a>	70
<a href="#">Εικόνα 28: Τα βασικά στάδια μίας NLP διαδικασίας</a>	71
<a href="#">Εικόνα 29: Βασικό Pipeline Επεξεργασίας</a>	72
<a href="#">Εικόνα 30: One Hot Encoding</a>	75
<a href="#">Εικόνα 31: Co-Occurrence Πίνακας</a>	76
<a href="#">Εικόνα 32: Παράδειγμα Νεύρο-Πιθανοτικού Μοντέλου</a>	77
<a href="#">Εικόνα 33: High Level απεικόνιση του CBOW</a>	79
<a href="#">Εικόνα 34: Multi Word</a>	79
<a href="#">Εικόνα 35: Skip Gram μοντέλο</a>	80
<a href="#">Εικόνα 36: BackPropagation</a>	82
<a href="#">Εικόνα 37: Συνάρτηση κόστους</a>	83
<a href="#">Εικόνα 38: FastText</a>	84
<a href="#">Εικόνα 39: Poicare μοντέλο</a>	85
<a href="#">Εικόνα 40: Σειρά δεδομένων</a>	86
<a href="#">Εικόνα 41: Φόρτωμα διαφορετικών tokenizers</a>	87
<a href="#">Εικόνα 42: Φόρτωμα Transformers στον κώδικα</a>	87
<a href="#">Εικόνα 43: Encoders – Decoders</a>	88
<a href="#">Εικόνα 44: Δομή του Encoder</a>	89
<a href="#">Εικόνα 45: Softmax συνάρτηση</a>	90
<a href="#">Εικόνα 46: Δημιουργία μασκών</a>	91
<a href="#">Εικόνα 47: Next Sentence Prediction</a>	92
<a href="#">Εικόνα 48: Λειτουργία του EMD</a>	93
<a href="#">Εικόνα 49: Συγκρίσεις μεταξύ των διαφορετικών μοντέλων</a>	93
<a href="#">Εικόνα 50: Παραγωγή προβλέψεων</a>	94



## ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΩΝ

Data mining – Εξόρυξη Δεδομένων

Machines Learning - Μηχανική Μάθηση

Fake news – Ψευδής ειδήσεις

Rumors – Φήμες

Supervised based approaches – Εποπτευόμενη Μάθηση

Unsupervised based approaches - Μη εποπτευόμενη Μέθοδος Μάθησης

Reinforcement Learning - Ενισχυτική Μάθηση

Planning – Σχεδιασμός

Profile Compatibility - Γλωσσική Ανάλυση

Rumor containment - Περιορισμός φημών

Kaggle – Γνωστό σύνολο δεδομένων

LIWC (Linguistic Enquiry and Word Count) - πρόγραμμα ανάλυσης κειμένου διαθέσιμο προς αγορά

Parse trees - Δενδροειδής δομή

Support Vector Machines - Μηχανές Διανυσμάτων Υποστήριξης

Regression - Παλινδρόμηση

Neural Networks - Νευρωνικά Δίκτυα

Recall -Ανάκληση, Μετρική απόδοσης

Precision - Ακρίβεια , Μετρική απόδοσης

F1-Score – Μετρική απόδοσης

FN (False Negative)

TP (True Positives)

TN (True Negatives)

FP (False Positives)

Decision Trees - Δέντρα Αποφάσεων

Document classification - κατηγοριοποίηση των εγγράφων

Deep Learning - Βαθεία Μηχανική Μάθηση

Multi – variate Bernouli – Μοντέλο Κατηγοριοποίησης

Multinomial - Μοντέλο Κατηγοριοποίησης

NETtalk – Πρόγραμμα μετατροπής κειμένου σε φωνή

Hidden layers - κρυφά επίπεδα ενός νευρωνικού δικτύου



Database – Βάση Δεδομένων

InVID – πρόγραμμα για την εξάλειψη της παραπληροφόρησης και των ψευδών ειδήσεων

Confusion matrix - Πίνακας σύγχυσης

Accuracy - Ακρίβεια

## ΠΕΡΙΛΗΨΗ

Το αντικείμενο απασχόλησης και μελέτης της συγκεκριμένης πτυχιακής εργασίας είναι το φαινόμενο της παραπληροφόρησης και των ψευδών ειδήσεων στο διαδίκτυο και συγκεκριμένα στα μέσα κοινωνικής δικτύωσης. Μελετάτε σε βάθος το φαινόμενο των ψευδών ειδήσεων, το οποίο έχει όλο ένα και περισσότερες εμφανίσεις τα τελευταία χρόνια.

Στη σύγχρονη κοινωνία στην οποία ζούμε, τα μέσα κοινωνικής δικτύωσης έχουν καθημερινή επίδραση στους ανθρώπους. Εκτός από τα πολλά θετικά στοιχεία που παρουσιάζονται με τη χρήση του διαδικτύου, υπάρχουν και αρκετά μειονεκτήματα από την αλόγιστη χρήση του. Ένα από αυτά είναι και η παραπληροφόρηση. Το φαινόμενο των ψευδών ειδήσεων αποτελεί πλέον μια μάστιγα, η οποία βρίσκεται σε όλο το εύρος του διαδικτύου και η μετάδοσή τους πραγματοποιείται με ταχύτερους ρυθμούς.

Στόχος της παρούσας πτυχιακής εργασίας είναι η μελέτη του φαινομένου των ψευδών ειδήσεων. Συγκεκριμένα, αναπτύσσονται οι μέθοδοι που βοηθούν και συμβάλουν στην διαδικασία ανίχνευσης αυτών. Για το λόγο αυτό θεωρείται σχεδόν απαραίτητο να αναφερθούμε στον κλάδο της Μηχανικής Μάθησης – Machines Learning και στις αντίστοιχες μεθόδους της.

## 1 ΚΕΦΑΛΑΙΟ - ΕΙΣΑΓΩΓΗ

Το αντικείμενο με το οποίο ασχολείται το πρώτο κεφάλαιο της συγκεκριμένης πτυχιακής εργασίας είναι η αναφορά και η ανάλυση των βασικών εννοιών όσον αφορά τις ψευδείς ειδήσεις. Συγκεκριμένα, θα πραγματοποιηθεί η περιγραφή της έννοιας των ειδήσεων για την σύγχρονη κοινωνία. Εν συνεχεία, θα αναφερθούμε στις επιπτώσεις που αυτές έχουν για τα μέσα κοινωνικής δικτύωσης καθώς και στον άμεσο εντοπισμό τους. Θα πραγματοποιηθεί η αναφορά στην εξόρυξη δεδομένων – data mining. Εν συνεχεία, θα αναφερθούμε στα διάφορα είδη φημών και στα χαρακτηριστικά τους.

### Ιστορική Αναδρομή

Όλη η ανθρωπότητα από τις μεγαλουπόλεις έως τα πιο απομακρυσμένα μέρη, έχουν την ανάγκη να ενημερώνονται για τις ειδήσεις και την επικαιρότητα. Η ανάγκη αυτή ξεκίνησε εδώ και πάρα πολλά χρόνια. Αρχικά, τα νέα διαδίδονταν από στόμα σε στόμα. Οι άνθρωποι δεν είχαν τη δυνατότητα να ενημερώνονται για παγκόσμιες ειδήσεις, παρά μόνο για τα νέα που υπήρχαν στις τοπικές κοινωνίες.

Περνώντας τα χρόνια, ήρθε στο προσκήνιο η εφημερίδα. Ήταν ένα πολύ σημαντικό μέσο με το οποίο μπορούσαν πλέον οι κοινωνίες να ενημερώνονται για ένα μεγαλύτερο φάσμα ειδήσεων. Πολλοί είναι αυτοί, που επιλέγουν τις εφημερίδες και σήμερα για την καθημερινή τους ενημέρωση. Εξίσου σημαντική για την ανθρωπότητα ήταν η εφεύρεση της τηλεόρασης. Πλέον, οι πολίτες είχαν τη δυνατότητα να παρακολουθούν σε καθημερινή βάση τα νέα που υπάρχουν σε παγκόσμιο επίπεδο.

Στις μέρες μας, η ενημέρωση των ειδήσεων μπορεί να πραγματοποιηθεί από διάφορα μέσα. Για πολλούς τα κυριότερα μέσα μετάδοσης των ειδήσεων παραμένουν τόσο οι εφημερίδες, όσο και η τηλεόραση. Σήμερα, ένα μεγάλο ποσοστό θεωρεί σημαντικό και πολύ χρησιμοποιούμενο μέσο μετάδοσης των ειδήσεων τα κοινωνικά δίκτυα.

### Σημασία των ειδήσεων στην ανθρωπότητα

Στην σύγχρονη κοινωνία στην οποία ζούμε, όλοι οι άνθρωποι κατακλύζονται διαρκώς από διάφορου τύπου πληροφορίες και ειδήσεις. Με τους όρους πληροφορίες και ειδήσεις αναφερόμαστε σε κάθε ενημέρωση προς τους πολίτες είτε αυτή είναι δημοσιογραφική είτε όχι.

Οι ειδήσεις για την ανθρωπότητα παίζουν πολύ σημαντικό ρόλο. Ο κυριότερος λόγος είναι η σωστή ενημέρωση. Οι πολίτες έχουν την δυνατότητα να ενημερώνονται διαρκώς για θέματα που τους ενδιαφέρουν και τους απασχολούν.

Όσο η τεχνολογία συνεχώς εξελίσσεται τόσο ο κλάδος των επικοινωνιών καθώς και του διαδικτύου έχουν παρουσιάσει τεράστια ανάπτυξη. Η ανάπτυξη αυτή έχει επηρεάσει και έχει αλλάξει σε μεγάλο βαθμό τη ζωή μας και την καθημερινότητα μας. Οι συγκεκριμένες αλλαγές έχουν συμβάλει τόσο θετικά όσο και αρνητικά όσον αφορά το συμφέρον της κοινής γνώμης. Πιο συγκεκριμένα, με την καθημερινή πληροφόρηση από διάφορες πηγές ελλοχεύει ο κίνδυνος της χειραγώγησης των πολιτών. (1)

Υπάρχουν πολλά πλεονεκτήματα όσον αφορά την επιρροή των ειδήσεων στην ανθρωπότητα. Αρχικά, οι πολίτες μέσω των ειδήσεων από τα διάφορα μέσα ενημέρωσης έχουν τη δυνατότητα να ενημερώνονται για γεγονότα και πράξεις που συμβαίνουν σε όλο τον κόσμο. Οποιαδήποτε στιγμή της ημέρας μπορούν να ενημερώνονται εύκολα και γρήγορα για ότι τους ενδιαφέρει. Η ενημέρωση πραγματοποιείται εύκολα διότι η πρόσβαση στις ειδήσεις μπορεί να γίνει από οποιοδήποτε σημείο και από όλες τις ηλεκτρονικές συσκευές που υπάρχουν στην αγορά, όπως λόγου χάριν τα κινητά τηλέφωνα και τα tablets. (2)

Από την άλλη μεριά, οι ειδήσεις μπορούν να συνεισφέρουν αρνητικά στην ανθρωπότητα και πιο συγκεκριμένα στην ψυχολογία τόσο των ενηλίκων, όσο και των ανήλικων ανθρώπων που τυχαίνει να παρακολουθήσουν διαφόρου τύπου ειδήσεις. Στις μέρες μας, οι ειδήσεις ενδέχεται να περιέχουν σκηνές βίας, μεγάλες καταστροφές καθώς και εγκληματικότητα. Τέτοιες ακατάλληλες σκηνές που συμβαίνουν συνεχώς γύρω μας, προβάλλονται καθημερινά σε όλα τα μέσα μαζικής ενημέρωσης. Οι συνέπειες από τη συγκεκριμένη προβολή μπορεί να παίζουν καθοριστικό ρόλο τόσο στην ψυχολογία όσο και στην συμπεριφορά των ανθρώπων.

Από την έναρξη λειτουργίας των σύγχρονων κοινωνικών μέσων μαζικής ενημέρωσης έχουν περάσει περίπου δύο δεκαετίες. Τα συγκεκριμένα μέσα ενημέρωσης τα χρόνια αυτά έχουν προσελκύσει μεγάλο αριθμό χρηστών. Μία από τις πιο διαδεδομένες μορφές κοινωνικών μέσων είναι τα κοινωνικά δίκτυα. Πιο συγκεκριμένα, τα κοινωνικά δίκτυα είναι ιστοσελίδες που επιτρέπουν στους χρήστες να δημοσιεύουν και να αλληλεπιδρούν με άλλους διαδικτυακούς χρήστες και να μοιράζονται διάφορα θέματα. (3)

Συγκεκριμένα, το 2003 πρωτοεμφανίστηκε το LinkedIn και ακολούθησαν το 2004 τα Myspace και Facebook. Ένα χρόνο αργότερα, στα μέσα του 2005 εμφανίστηκε το YouTube .

Εν συνεχεία εμφανίστηκαν το Twitter, το Blogging, το Google+, το WhatsApp καθώς και το Instagram. (4)

Τα συγκεκριμένα κοινωνικά δίκτυα συνέβαλαν χρόνο με το χρόνο στην εξάπλωση των ειδήσεων και των πληροφοριών. Λόγω της ανάπτυξης της τεχνολογίας η πρόσβαση στα κοινωνικά δίκτυα ήταν διαθέσιμα από τους χρήστες όλη την ώρα καθ' όλη τη διάρκεια της ημέρας. Γεγονός που βοήθησε ακόμα περισσότερο στην εξάπλωση και τη μετάδοση των πληροφοριών και των ειδήσεων. (2) Μάλιστα η διαδικασία της μετάδοσης των ειδήσεων αποτελούσε πλέον μια αρκετά γρήγορα διαδικασία.

## Ψευδείς ειδήσεις (fake news)

### Ορισμός

Ένας ορισμός που μπορεί να δοθεί για το όρο των ψευδών ειδήσεων είναι ο ακόλουθος:

Ορίζουμε ως Ψευδής ειδήσεις ή Fake news τις παραπλανητικές πληροφορίες, οι οποίες διαδίδονται εσκεμμένα στοχεύοντας στην επιρροή της κοινής γνώμης. (5)

### Είδη ψευδών ειδήσεων

Οι ψευδής ειδήσεις ανάλογα με το περιεχόμενό τους μπορούν να κατηγοριοποιηθούν στις εξής κατηγορίες:

- Δημοσιογραφικό Σφάλμα
- Παρωδία
- Προπαγάνδα
- Ψευδοεπιστήμη
- Ιδεολογία ατόμων με συγκριμένα «πιστεύω»

### Χαρακτηριστικά γνωρίσματα

Στις μέρες μας τα μέσα κοινωνικής δικτύωσης έχουν παρουσιάσει μεγάλη δημοτικότητα σε όλο το φάσμα των ηλικιών. Το διαδίκτυο και οι υπηρεσίες που αυτό προσφέρει στους δισεκατομμύρια χρήστες του έχει βοηθήσει την διαδικασία μετάδοσης των ειδήσεων.

Τα μέσα κοινωνικής δικτύωσης έχουν δημιουργήσει πολλά οφέλη και πλεονεκτήματα στο τομέα της μετάδοσης των ειδήσεων. Όμως η αλόγιστη χρήση του τόσο από τους δημοσιογράφους όσο και από τους υπόλοιπους χρήστες του διαδικτύου, έχει οδηγήσει στην

αύξηση του φαινομένου των ψευδών ειδήσεων – fake news. Το φαινόμενο αυτό εμφανίζεται πλέον σε καθημερινή βάση τόσο σε πολιτικές και κοινωνικές όσο και σε οικονομικές εξελίξεις. (2)

Τα πλεονεκτήματα που παρέχουν οι υπηρεσίες του διαδικτύου στον κλάδο της δημοσιογραφίας είναι πάρα πολλά. Ενδεικτικά μπορούμε να αναφέρουμε μερικά από αυτά τα πλεονεκτήματα. Αρχικά θα πρέπει να αναφερθούμε στην αρκετά εύκολη διαδικασία όσον αφορά την ανεύρεση των πληροφοριών καθώς και την ταχύτατη δημοσίευση των ειδήσεων και των πληροφοριών αυτών. Επίσης, ένα πολύ σημαντικό πλεονέκτημα που παρατηρείται από τη χρήση των μέσων κοινωνικής δικτύωσης είναι η συνεχής ανεύρεση καινούργιων τρόπων διαδραστικής επικοινωνίας και επαφής με το αναγνωστικό κοινό τους.

Τα τελευταία χρόνια κατά τα οποία η διαδικτυακή δημοσιογραφία έχει παρουσιάσει ραγδαία ανάπτυξη, παρουσιάζετε ολοένα και περισσότερο το πρόβλημα των ψευδών ειδήσεων, γνωστά και ως «fake news». Με τον όρο ψευδής ειδήσεις αναφερόμαστε στις πληροφορίες εκείνες οι οποίες χαρακτηρίζονται ως αναξιόπιστες και προέρχεται από μη έγκυρες πηγές. (3)

Το κοινό έχει δείξει ιδιαίτερο ενδιαφέρον όσον αφορά τις ειδήσεις στα μέσα κοινωνικής δικτύωσης. Πιο συγκεκριμένα, έρευνες έχουν δείξει ότι το μεγαλύτερο ποσοστό των ανθρώπων στις μέρες μας προτιμούν να ενημερώνονται για τις ειδήσεις που τους ενδιαφέρουν από τα μέσα κοινωνικής δικτύωσης. Υπάρχουν πολλοί λόγοι για τους οποίους συμβαίνει αυτό. Αρχικά το χαμηλό κόστος χρήσης του είναι ο κυριότερος λόγος που προτιμάται από το κοινό έναντι των παραδοσιακών μέσων ενημέρωσης. Επίσης, τόσο η εύκολη πρόσβαση όσο και η ταχύτητα διάδοσης των ειδήσεων είναι δύο λόγοι που συμβάλλουν στην ολοένα και αυξανόμενη χρήση των μέσων κοινωνικής δικτύωσης. (1)

## **Εντοπισμός ταχέων διαδιδόμενων θεμάτων και εντοπισμός ψευδών ειδήσεων**

Με την συνεχή ανάπτυξη της τεχνολογίας και ειδικότερα η ραγδαία ανάπτυξη των ηλεκτρονικών συσκευών, έχουν επιφέρει πολλά οφέλη στους χρήστες του διαδικτύου. Το σημαντικότερο γεγονός από όλα είναι η εύκολη πρόσβαση στα κοινωνικά μέσα ενημέρωσης όλο το 24ωρο.

Οι πληροφορίες διαδίδονται με ραγδαίου ρυθμούς συγκριτικά με τα προηγούμενα χρόνια. Το ζήτημα που προκύπτει είναι κατά πόσο οι πληροφορίες αυτές είναι αξιόπιστες. Η αυτόματη ανίχνευση και εξακρίβωση των φημών στα μέσα κοινωνικής δικτύωσης είναι δύο εξαιρετικά δύσκολες διεργασίες. (1) Η αυτόματη ανίχνευση των φημών στα μέσα κοινωνικής δικτύωσης,

παραδείγματος χάριν σε ιστοσελίδες ενημέρωσης έχει αποτελέσει σημαντικό θέμα έρευνας και προσπάθειας επίλυσης από τους ειδικούς.

## Φήμες

Πριν από οποιαδήποτε αναφορά για τις φήμες και τους τρόπους με τους οποίους μπορεί κανείς να ανιχνεύσει και κατ' επέκταση να προβεί στην εξάλειψή τους, πρέπει να δοθεί ένας σαφής ορισμός τους.

Η φήμη ορίζεται ως ένας ισχυρισμός, ο οποίος θεωρείται μη επαληθευμένος. Ο ισχυρισμός αυτός μπορεί να ξεκινήσει από μια ή και περισσότερες πηγές. Με την πάροδο του χρόνου η συγκεκριμένη πηγή εξαπλώνεται από διάφορες πηγές στο διαδίκτυο.

Παραδείγματος χάριν στο Twitter, μια φήμη είναι μια συλλογή από tweets. Αυτά τα tweets αναφέρονται στο ίδιο γεγονός, το οποίο όμως δεν έχει επαληθευτεί με κάποιον τρόπο. Οι φήμες στο Twitter διαδίδονται με πολύ μεγάλη ευκολία και σε πολύ μικρό χρονικό διάστημα. Όπως συμβαίνει στο Twitter, αντίστοιχα διαδίδονται και αντιμετωπίζονται οι φήμες και από τα υπόλοιπα μέσα κοινωνικής δικτύωσης, όπως είναι και το Facebook. (3)

Οι Allport και Postman όρισαν τον όρο φήμη ως ορισμένες προτάσεις που αφορούν συγκεκριμένα ή επίκαιρα θέματα, τα οποία περνούν από άτομο σε άτομο, συγκεκριμένα από στόμα σε στόμα, χωρίς καμία απόδειξη ότι οι προτάσεις αυτές είναι αληθής. (4)

Ένας ακόμα ορισμός για τις φήμες έρχεται από τους DiFonzo και Bordia. Συγκεκριμένα, ορίζουν ως φήμη μια μη επαληθευμένη και συναφής πληροφορία στα πλαίσια κινδύνου ή απειλής που μεταδίδεται και στοχεύει στο να ενθαρρύνει τους χρήστες ούτως ώστε να διαχειρίζονται όσο το δυνατόν καλύτερα τον κίνδυνο. (4)

Ο ορισμός που παρουσιάζουν οι Dunn και Allen θέλουν τη φήμη να είναι μια υπόθεση που μεταδίδεται απουσία επαληθεύσιμων πληροφοριών που σχετίζονται με αβέβαιες καταστάσεις και γεγονότα. Οι καταστάσεις αυτές είναι ιδιαίτερα σημαντικές για τους ανθρώπους που τις μεταδίδουν ενώ παράλληλα ανησυχούν για την εγκυρότητα τους. (4)

## Χαρακτηρισμός φημών (rumor characterization)

Στην υποενότητα αυτή αναφέρονται και αναλύονται τα χαρακτηριστικά των φημών που υπάρχουν σε πολύ μεγάλο βαθμό στις μέρες μας. Συγκεκριμένα θα γίνει αναφορά στη γλωσσολογία, τους χρήστες και το χρονικό μετάδοσης των φημών στους διαδικτυακούς τόπους.

Όπως έχει ήδη αναφερθεί μια φήμη μπορεί πολύ εύκολα να μεταδοθεί από άτομο σε άτομο σε πολύ μικρό χρονικό διάστημα με τη πολύτιμη συμβολή των σύγχρονων μέσων κοινωνικής δικτύωσης. Έτσι ένα μη επιβεβαιωμένο γεγονός μπορεί σε ελάχιστα λεπτά να μεταδοθεί σε μια πολύ μεγάλη μερίδα ανθρώπων που χρησιμοποιούν το διαδίκτυο και ενημερώνονται μέσω αυτού για τις ειδήσεις της επικαιρότητας. (7)

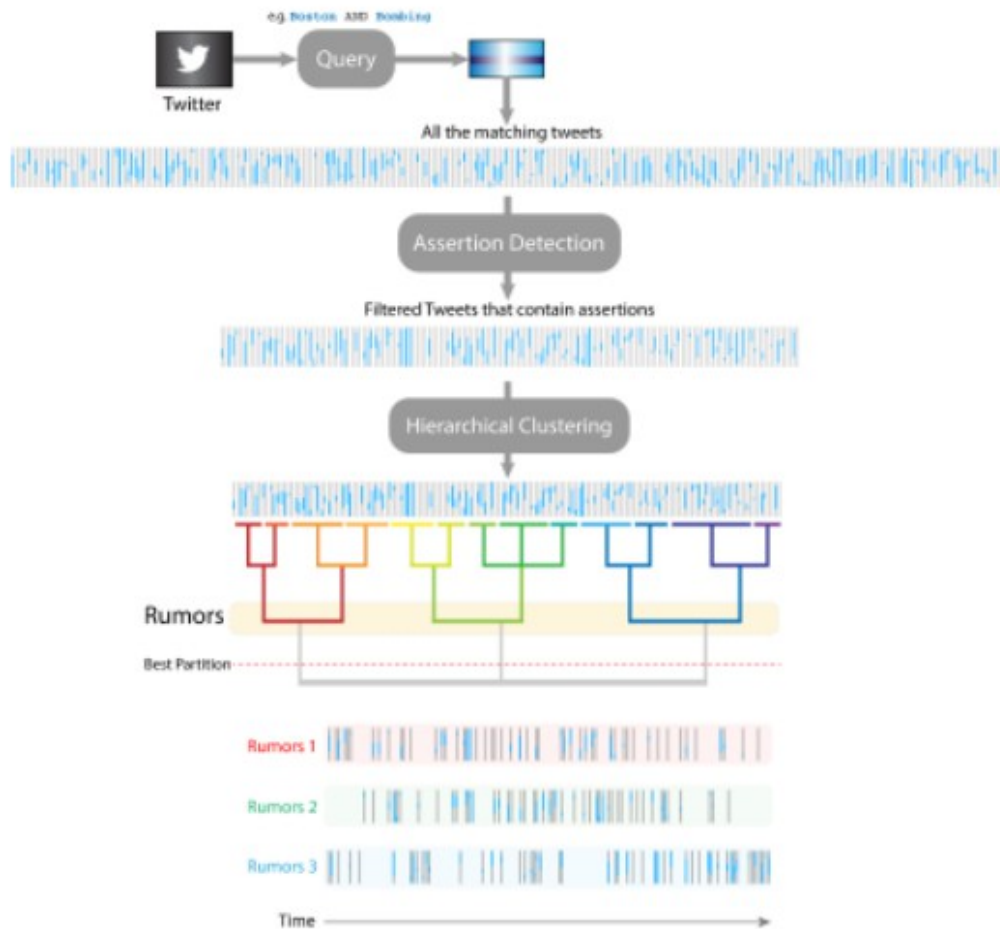
Συνήθως οι δημοσιογράφοι καθώς και οι διαχειριστές διαφόρων blog του διαδικτύου αρχικά δημοσιεύουν μια μη επιβεβαιωμένη είδηση. Εν συνεχεία, οι εκατοντάδες χρήστες μέσω του διαδικτύου μπορούν εύκολα να αναμεταδώσουν τις ειδήσεις αυτές από άτομο σε άτομο.

### **Ανίχνευση φημών (rumor detection)**

Στην ενότητα αυτή θα αναφερθούμε στην ανίχνευση και επαλήθευση των φημών. Συγκεκριμένα, θα αναφερθεί και θα περιγράψει το σύστημα ανίχνευσης των φημών.

Το σύστημα ανίχνευσης και επαλήθευσης φημών γίνεται περισσότερο κατανοητό με τη βοήθεια της εικόνας 1 που βρίσκεται στη συνέχεια. Όπως μπορεί κανείς να παρατηρήσει από την παρακάτω εικόνα αρχικά πραγματοποιείται η ανίχνευση των ισχυρισμών μέσω του κοινωνικού δικτύου Tweeter. Εν συνεχεία, υλοποιείται ένα φιλτράρισμα στα περιεχόμενα των συγκεκριμένων ισχυρισμών. Μέσω ενός ταξινομητή πραγματοποιείται η ομαδοποίηση τους. Οι φήμες που είχαν αρχικά ανιχνευθεί από το Tweeter έχουν πλέον ταξινομηθεί.





Εικόνα 1 - Περιγραφή συστήματος ανίχνευσης φημών

Συγκεκριμένα, η διαδικασία της ανίχνευσης των φημών σχετικά με ένα γεγονός πραγματοποιείται μέσω της μεθόδου της ταξινόμησης καθώς και της συγκέντρωσης όλων των σχετικών ισχυρισμών για το γεγονός αυτό. Για το σκοπό αυτό, χρησιμοποιείται ένας ταξινομητής, ο οποίος χρησιμοποιώντας τόσο συντακτικά όσο και σημασιολογικά χαρακτηριστικά είναι ικανός να αναγνωρίζει ισχυρισμούς σε ποσοστό που φτάνει το 91%. (8)

### Περιορισμός φημών (rumor containment)

Με τον όρο φήμη αναφερόμαστε σε μια ιστορία η οποία μπορεί να είναι είτε αληθινή είτε κατασκευασμένη. Η ιστορία αυτή θεωρείται μια ανεπίσημη ιστορία, η οποία μεταδίδεται πολύ γρήγορα από άτομο σε άτομο.

Υπάρχουν τρεις τρόποι με τους οποίους μπορεί να τερματίσει η διάδοση μιας φήμης. Αρχικά, μπορεί να προκύψει η επίλυση μιας φήμης είτε ως αληθινή ότι δηλαδή είναι ένα πραγματικό γεγονός, είτε ως ψευδής, ότι είναι ένα μη πραγματικό γεγονός, είτε τέλος ως ένα άλλο

πρόβλημα. Είναι δυνατόν να υπάρξουν πολλές φήμες που να αφορούν το ίδιο θέμα. Πολλές από αυτές μπορεί να είναι ψευδής, ενώ άλλες μπορεί να είναι αληθής. (8)

Στην περίπτωση όπου μια ή περισσότερες φήμες φτάσουν στο σημείο της επίλυσης, αυτό σημαίνει ότι οι υπόλοιπες φήμες που υπάρχουν για το συγκεκριμένο θέμα επιλύονται αυτόματα.

## Σάτιρα

Η Σάτιρα είναι το κυριότερο είδος φημών που υπάρχει. Πρόκειται είτε για ένα κείμενο είτε για ένα σκίτσο, κοινού σχολιασμού. Αναφερόμαστε σε σάτιρα όταν το κείμενο ή το σκίτσο αυτό προκαλεί τη γελοιοποίηση τόσο πραγματικών γεγονότων όσο και πραγματικών προσώπων.

### ΟΡΙΣΜΟΣ:

Η Σάτιρα (Satire) είναι η τέχνη του να υποβιβάζει και να υποβαθμίζει ένα θέμα. Αυτό επιτυγχάνεται μέσω της γελοιοποίησης του συγκεκριμένου θέματος. Κάτι τέτοιο έχει ως αποτέλεσμα το ευρύ κοινό είτε να διασκεδάσει είτε να θυμώσει και κατ' επέκταση να το περιφρονήσει. (9)

Υπάρχει ουσιώδη διαφορά ανάμεσα στη σάτιρα και το κωμικό. Πιο συγκεκριμένα, ένα κωμικό σχόλιο είναι ικανό να προκαλέσει γέλιο στον αναγνώστη, το οποίο αποτελεί και τον κυριότερο στόχο. Από την άλλη μεριά η σάτιρα διαχειρίζεται το γέλιο ως κύριο όπλο της. Επιπρόσθετα, η σάτιρα καυτηριάζει, στοχεύοντας σε έναν συγκεκριμένο στόχο. Ο στόχος αυτός μπορεί να είναι ένα άτομο, ένας θεσμός, μια συγκεκριμένη κοινωνική τάξη κλπ. (9)

Ο δημοσιογράφος ή και οποιοδήποτε άλλο πρόσωπο διαχειρίζεται τη σάτιρα για να πετύχει το στόχο του χρησιμοποιεί συγκεκριμένες μεθόδους όπως λόγου χάριν την παρωδία, την σύγκριση, την ειρωνεία καθώς και σε πολλές περιπτώσεις την υπερβολή.

## Προπαγάνδα

### Ορισμός

Προπαγάνδα είναι η συνεχή και διαδοχική μετάδοση συγκεκριμένων ιδεών και αντιλήψεων. Οι ιδέες και οι αντιλήψεις αυτές μπορεί να αφορούν είτε θρησκευτικό, ιδεολογικό και πολιτικό χαρακτήρα είτε οποιοδήποτε άλλο τομέα. (10)

## Χαρακτηριστικά

Η προπαγάνδα έχει ως στόχο της τον επηρεασμό της κοινής γνώμης. Στοχεύει στη διαμόρφωση και στρέβλωση της κοινής γνώμης μέσω της ελλιπής μετάδοσης ειδήσεων καθώς και της παραπληροφόρησης.

Επίσης, μπορούμε να αναφερθούμε στην προπαγάνδα ως η διαδικασία στρέβλωσης της αλήθειας. Η διαδικασία αυτή έχει ως σκοπό είτε να βοηθήσει ένα άτομο, μια επιχείρηση και ένα έθνος είτε να επιφέρει αρνητικά αποτελέσματα και κατ'επέκταση να το βλάψει.

Η λέξη προπαγάνδα εδώ και πολλά χρόνια χρησιμοποιείται κατά κύριο λόγο από το πολιτικό πλαίσιο. Πιο συγκεκριμένα, με την προπαγάνδα στην πολιτική αναφερόμαστε σε ορισμένες κινήσεις, τις οποίες προωθούν οι πολιτικές ομάδες με σκοπό να επηρεάσουν την κοινή γνώμη. (10)

## 2 ΚΕΦΑΛΑΙΟ - ΕΡΕΥΝΗΤΙΚΟ ΥΠΟΒΑΘΡΟ STATE OF THE ART

Το αντικείμενο με το οποίο ασχολείται το δεύτερο κεφάλαιο είναι η αναφορά και η ανάλυση των θεμελιωδών εννοιών και τεχνικών που χρησιμοποιούνται στην έρευνα για τον εντοπισμό ψευδών ειδήσεων. Συγκεκριμένα, θα πραγματοποιηθεί η περιγραφή των μεθόδων που χρησιμοποιούνται για την ανίχνευση των ψευδών ειδήσεων στο διαδίκτυο. Εν συνεχεία, θα πραγματοποιηθεί η γλωσσική ανάλυση και η διαδικτυακή διερεύνηση του συγκεκριμένου κλάδου. Τέλος, θα αναφερθούμε στον ορισμό καθώς και την περιγραφή του επιστημονικού κλάδου της μηχανικής μάθησης – Machine Learning και κατά πόσο αυτή συνέβαλε στην διαδικασία ανίχνευσης ψευδών ειδήσεων στο διαδίκτυο μέχρι και σήμερα.

### Μέθοδοι Ανίχνευσης ( Detecting rumors in social media: A survey)

Στο κεφάλαιο αυτό πραγματοποιείτε μια ανάλυση των μεθόδων ανίχνευσης φημών. Πιο συγκεκριμένα, αναφερόμαστε στις μεθόδους ανίχνευσης των φημών που υπάρχουν στα πολλά και διαφορετικά μέσα κοινωνικής δικτύωσης, όπως είναι λόγω χάριν το Facebook.

Υπάρχουν τρεις διαφορετικές προσεγγίσεις όσον αφορά την ανίχνευση των φημών. Ύστερα από έρευνες που έχουν πραγματοποιηθεί γύρω από το πρόβλημα της ανίχνευσης ψευδών ειδήσεων στα μέσα κοινωνικής δικτύωσης, κατέληξαν στις ακόλουθες τρεις προσεγγίσεις οι οποίες βοηθούν στο να γίνει η διαδικασία εντοπισμού τους ευκολότερη. (4)

- **Εποπτευόμενες - supervised based approaches**
- **Μη εποπτευόμενες - Unsupervised based approaches**
- **Ενισχυτική Μάθηση - Reinforcement Learning**

### Τεχνικές

#### Γλωσσική Ανάλυση (Profile Compatibility)

Στην υποενότητα αυτή θα πραγματοποιηθεί η γλωσσική ανάλυση των κειμένων. Πιο συγκεκριμένα, θα αναλυθούν τόσο τα συντακτικά και γραμματικά όσο και τα μορφολογικά χαρακτηριστικά των κειμένων που αποτελούν τις ειδήσεις που καλούνται να ανιχνευτούν και να χαρακτηριστούν ως ψευδής ή αληθής.

Μια από τις πιο σημαντικές διεργασίες που πρέπει να πραγματοποιηθεί για τη διαδικασία ανίχνευσης ψευδών ειδήσεων μέσω της Μηχανικής Μάθησης είναι η κατηγοριοποίηση των

κειμένων καθώς και ο εντοπισμός των χαρακτηριστικών γνωρισμάτων τους για την περαιτέρω εξέτασή τους.

### *Σύνολο των δεδομένων*

Το σύνολο των δεδομένων απαρτίζεται από κείμενα διαφόρων ειδήσεων, τα οποία χαρακτηρίζονται είτε ως ψευδής, είτε ως αληθής. Η διαδικασία της κατηγοριοποίησης των δεδομένων αυτών είναι μια διαδικασία, η οποία απαιτεί πολύ χρόνο. Για το λόγο αυτό, είναι ιδιαίτερα δύσκολο να βρεθεί ένα σύνολο, το οποίο να είναι αξιόλογο.

Η διαδικασία της εύρεσης των ψευδών ειδήσεων στα μέσα κοινωνικής δικτύωσης έχει κεντρίσει ιδιαίτερα τον τελευταίο καιρό το ενδιαφέρον του επιστημονικού κλάδου. Ως συνέπεια της συνεχούς έρευνας γύρω από τον συγκεκριμένο κλάδο έχουν διατεθεί στο διαδίκτυο αρκετά δείγματα. Ανάμεσα σε αυτά τα δείγματα που κυκλοφορούν παρατηρήθηκε ένα κοινό τους γνώρισμα. Πιο συγκεκριμένα, τα δείγματα στηρίζονται στο σύνολο δεδομένων Kaggle. (16)

Ένα από τα πιο σημαντικά πλεονεκτήματα του συγκεκριμένου συνόλου είναι ότι αποτελεί ένα αρκετά μεγάλο σύνολο. Παρ' όλα αυτά παρατηρείται και ένα εξίσου σημαντικό μειονέκτημα του. Συγκεκριμένα, με σκοπό να κατασκευαστεί το σύνολο δεδομένων Kaggle χρειάστηκε να βασιστεί σε πηγές, οι οποίες όμως δεν ξεχωρίζει με οποιονδήποτε τρόπο τις ιστοσελίδες και τις πηγές του διαδικτύου που παλαιότερα έχουν αναρτήσει άρθρα και ειδήσεις που έχουν χαρακτηριστεί ως αληθής ειδήσεις.

### *Επιλογή κατάλληλων γνωρισμάτων*

Ένας αλγόριθμος ταξινόμησης δεν είναι δυνατόν να δεχθεί ως εισόδους τα στοιχεία των ειδήσεων όπως αυτά εμφανίζονται στα διάφορα ειδησεογραφικά sites καθώς και στις αναρτήσεις των χρηστών στα διάφορα μέσα κοινωνικής δικτύωσης. Επομένως, θα πρέπει τα συγκεκριμένα δεδομένων να υποστούν μια τροποποίηση με σκοπό να φτάσουν στην ιδανική μορφή. Ο πιο συνηθισμένος και πολύ χρησιμοποιούμενος τρόπος αναπαράστασης των κειμένων αυτών είναι τα διανύσματα βαρών. Τα βάρη ανάλογα με την περίπτωση που καλείτε να υλοποιηθεί μπορούν να αντιστοιχούν σε διάφορα χαρακτηριστικά γνωρίσματα. Λόγου χάριν, σε ορισμένες περιπτώσεις θα μπορούσαν τα βάρη να παίρνουν τις τιμές 0 και 1. Οι τιμές αυτές θα σήμαιναν την παρουσία ή την απουσία μιας συγκεκριμένης λέξης μέσα στο σύνολο του κειμένου. Τέτοιου είδους λύσεις, πολλές φορές δεν είναι αποδοτικές εάν τα χαρακτηριστικά του κειμένου δεν είναι τόσο απλοϊκά, όπως για παράδειγμα τόσο η

συντακτική όσο και η λεξικολογική πληροφορία που υπάρχει μέσα στο σύνολο των ειδήσεων προς διερεύνηση.

Οι ερευνητές που έχουν μελετήσει σε βάθος το φαινόμενο των ψευδών ειδήσεων στο διαδίκτυο και συγκεκριμένα στα μέσα κοινωνικής δικτύωσης έχουν καταλήξει στο συμπέρασμα ότι είναι εφικτό να κατηγοριοποιηθούν με επιτυχία μη έγκυρες ειδήσεις εάν αναλυθούν και αξιολογηθούν τόσο τα μορφολογικά όσο και τα συντακτικά χαρακτηριστικά των κειμένων αυτών. (17)

Για να γίνει πλήρως κατανοητό, είναι σκόπιμο να αναφέρουμε ένα χαρακτηριστικό παράδειγμα. Αναλύοντας τις ψευδής ειδήσεις που κατά καιρούς έχουν εμφανιστεί στο διαδίκτυο, παρατηρήθηκε ότι οι τίτλοι αυτών των παραπλανητικών ειδήσεων περιέχουν λιγότερα ουσιαστικά συγκριτικά με τις έγκυρες ειδήσεις. (3) Επίσης, παρατηρήθηκε ότι οι τίτλοι ψευδών ειδήσεων είναι πιο μεγάλες όσον αφορά το συνολικό τους μέγεθος συγκρινόμενοι με τις αληθής ειδήσεις.

Υπάρχουν τέσσερις μεγάλες κατηγορίες στις οποίες μπορεί κανείς να κατατάξει τα χαρακτηριστικά γνωρίσματα.

➤ Υφολογική ανάλυση

Στην κατηγορία Υφολογικής ανάλυσης περιέχονται όλα τα χαρακτηριστικά γνωρίσματα που αφορούν το συντακτικό, τα γραμματικά δεδομένα καθώς και το ύφος του κειμένου των ειδήσεων. Λόγου χάριν, στην συγκεκριμένη κατηγορία εντάσσετε τόσο η καταμέτρηση των λέξεων, ουσιαστικών, σημείων στίξης όσο και των λέξεων που περιέχουν κεφαλαία γράμματα.

➤ Συναισθηματική ανάλυση

Η συγκεκριμένη κατηγορία αναφέρεται σε γνωρίσματα τα οποία αφορούν τόσο συναισθηματικούς όσο και ψυχολογικούς χαρακτηρισμούς. Στην περίπτωση αυτή, χρησιμοποιείτε ένας αναλυτής κειμένου και ένα σύνολο λεξικών.

➤ Πολυπλοκότητα κειμένου

Τα χαρακτηριστικά γνωρίσματα αυτής της κατηγορίας αναφέρονται στην πολυπλοκότητα του κειμένου. Αξίζει να σημειωθεί ότι η πολυπλοκότητα μετριέται σε επίπεδο των λέξεων καθώς και σε επίπεδο προτάσεων.

➤ Σχετικότητα

Στην τέταρτη και τελευταία κατηγορία εμφανίζονται τα χαρακτηριστικά τα οποία εκφράζουν σχετικότητα σύμφωνα με το λεξικό LIWC. (18) Καθώς επίσης και όσα γνωρίσματα δεν κατατάσσονται σε κάποια άλλη κατηγορία.

### *Μορφολογική ανάλυση*

Ο κύριος στόχος της μορφολογικής ανάλυσης είναι να προσδιορίζει για κάθε λέξη μέσα σε οποιαδήποτε πρόταση το μέρος του λόγου. Προκειμένου να πραγματοποιηθεί η συγκεκριμένη διεργασία, αρχικά θα πρέπει ο αναλυτής να χωρίσει την κάθε πρόταση. Εν συνεχεία, πραγματοποιείται η ανάθεση ετικετών σε όλα τα επιμέρους τμήματα που προέκυψαν μετά τον διαχωρισμό. (17) Οι ετικέτες αυτές προκύπτουν από ένα συγκεκριμένο σύνολο ετικετών. Πιο συγκεκριμένα, κάθε ετικέτα αντιστοιχεί σε ένα μέρος του λόγου παραδείγματος χάριν το ουσιαστικό, το ρήμα, το επίρρημα κλπ.

Η διαδικασία που αναφέρθηκε προηγουμένως καλείται ετικετοποίηση. Θεωρείται μια πολύ σημαντική εργασία διότι παρέχει χρήσιμες πληροφορίες για τις λέξεις.

### *Συντακτική ανάλυση*

Με τον όρο συντακτική ανάλυση αναφερόμαστε στη διαδικασία κατά την οποία πραγματοποιείται αναγνώριση της πρότασης καθώς και σημασίας της. Κάτι τέτοιο επιτυγχάνεται με τη σωστή κατανόηση της συντακτικής της δομής. Για τη σωστή υλοποίηση της διαδικασίας χρησιμοποιείται η ιεραρχική δενδροειδή δομή – parse trees. (3)

### **Διαδικτυακή Διερεύνηση**

Η δυσκολία στην αναγνώριση των ψευδών ειδήσεων – fake news γίνεται απόλυτα κατανοητή αν παρατηρήσει κανείς τα αποτελέσματα της έρευνας που πραγματοποίησαν οι “YouGov” και “The Economist”. Όπως φαίνεται και στην εικόνα που βρίσκεται στη συνέχεια, οι δύο οργανισμοί έφτασαν στο συμπέρασμα ότι ένα πολύ μικρό ποσοστό των χρηστών του διαδικτύου είναι σε θέση να αναγνωρίσει με σιγουριά μια αληθής είδηση από μια ψευδή. Συγκεκριμένα, σε ποσοστό που φτάνει μόλις το 22% από ένα σύνολο 1.485 χρηστών που συμμετείχαν στην συγκεκριμένη έρευνα δήλωσαν σίγουροι ότι μπορούν να διακρίνουν μια ψευδή είδηση στο διαδίκτυο. (19)

	Share of respondents
Very confident	22%
Somewhat confident	44%
Not very confident	16%
Not at all confident	5%
Not sure	13%

Showing entries 1 to 5 (5 entries in total)

© Statista 2017

**Additional Information:** United States; YouGov; The Economist; October 7 to 10, 2017; 1,485 respondents; 18 years and older

**Sources:** YouGov; The Economist

### Εικόνα 2 - Αποτελέσματα έρευνας για την αναγνώριση των ψευδών ειδήσεων

Το πρόβλημα των ψευδών ειδήσεων μπορεί να διορθωθεί αντιμετωπίζοντας δύο διαφορετικά μέρη του προβλήματος. Αρχικά, το πρώτο σκέλος που καλείται κάποιος να αντιμετωπίσει είναι οι μέθοδοι που υπάρχουν για την ανίχνευση των ψευδών ειδήσεων στο διαδίκτυο. Ενώ αντίθετα το δεύτερο μέρος που μπορεί να αντιμετωπίσει κανείς τα fake news είναι η υλοποίηση ενός κατάλληλα διαμορφωμένου συστήματος.

## Μηχανική Μάθηση (Machine Learning)

### Ιστορική αναδρομή

Όπως έχει ήδη αναφερθεί, η Μηχανική Μάθηση – Machine Learning έχει αναπτυχθεί ιδιαίτερα τα τελευταία χρόνια. Ο επιστημονικός κλάδος καθώς και ο εκπαιδευτικός έχουν επικεντρωθεί στον συγκεκριμένο κλάδο. Παρ' όλα αυτά η έννοια της μηχανικής μάθησης πρωτοεμφανίστηκε μόλις το 1700. Ο κλάδος αυτός βασίστηκε τότε στο θεώρημα του Bayes. (20)

Η ιστορία του κλάδου της μηχανικής μάθησης όπως έχει ήδη αναφερθεί ξεκίνησε εδώ και χιλιάδες χρόνια. Αρχικά αποτέλεσε έναν επιστημονικό κλάδο, ο οποίος ήταν αλληλοεξαρτούμενης με άλλους τρεις βασικούς επιστημονικούς κλάδους. Συγκεκριμένα, οι επιστημονικοί αυτοί κλάδοι ήταν οι υπολογιστές, τα μαθηματικά και την φιλοσοφία. Πρόκειται ουσιαστικά για τον κλάδο αυτόν, ο οποίος συγκέντρωνε όλη την ανθρώπινη γνώση και την ευφυΐα. Η αρχή του συγκεκριμένου κλάδου με βάση τη μορφή που γνωρίζουμε και σήμερα, τοποθετείται από τους ειδικούς αμέσως μετά τον τερματισμό του δεύτερου παγκοσμίου πολέμου περίπου στα τέλη του 1945. (12)



## *Εισαγωγή στη Μηχανική Μάθηση*

Ένα χαρακτηριστικό που παρατηρείται στη συμπεριφορά των ανθρώπων στην κοινωνία μας και θεωρείται πολύ σημαντικό είναι η ικανότητά τους στη Μάθηση (Learning). Για το λόγο αυτό, οι επιστήμονες των Τεχνητών Νευρωνικών Δικτύων στηρίχτηκαν στο συγκεκριμένο χαρακτηριστικών των ατόμων στη μάθηση. Πιο συγκεκριμένα, οι επιστήμονες κατασκευάζουν υπολογιστικά συστήματα, με σκοπό να μπορούν να εκπαιδεύονται και να επιτύχουν τους ορισμένους στόχους που τίθενται. Κάτι τέτοιο επιτυγχάνεται κατά κύριο λόγο μέσω της Μηχανικής Μάθησης (Machine Learning). (21)

Είναι σκόπιμο να αναφερθούμε στο σημείο αυτό στον ορισμό της Μηχανικής Μάθησης (Machine Learning). Πρόκειται για ένα φαινόμενο, στο οποίο τα συστήματα αυξάνουν συνεχώς την απόδοση τους εκτελώντας μια διεργασία ενώ δεν είναι απαραίτητο συνεχώς να προγραμματίζονται εκ νέου. (22)

Αντικείμενο μελέτης της Μηχανικής Μάθησης είναι η δημιουργία κατάλληλων προγραμμάτων, τα οποία θα έχουν την ικανότητα να βελτιώνουν την απόδοσή τους αυτόνομα με βάση τις γνώσεις και την εμπειρία που διαθέτουν. Είναι ένα κλάδος των επιστημών, του οποίου η εξέλιξη είναι ραγδαία και μέρα με τη μέρα εξελίσσεται περαιτέρω.

Το αντικείμενο με το οποίο ασχολείται ο συγκεκριμένος κλάδος είναι η δημιουργία αλγορίθμων βελτιστοποίησης. Αξιοσημείωτο είναι το γεγονός ότι αυτός ο επιστημονικός κλάδος έχει τη δυνατότητα να αλληλεπιδράει με το γνωστικό πεδίο των υπόλοιπων επιστημονικών κλάδων. Με τη χρήση των αλγορίθμων μάθησης επιτυγχάνεται η κατάλληλη εξαγωγή γενικευμένων κανόνων μάθησης, για να επιλύονται προβλήματα με βάση ένα πεπερασμένο σύνολο πληροφοριών εκπαίδευσης. (21)

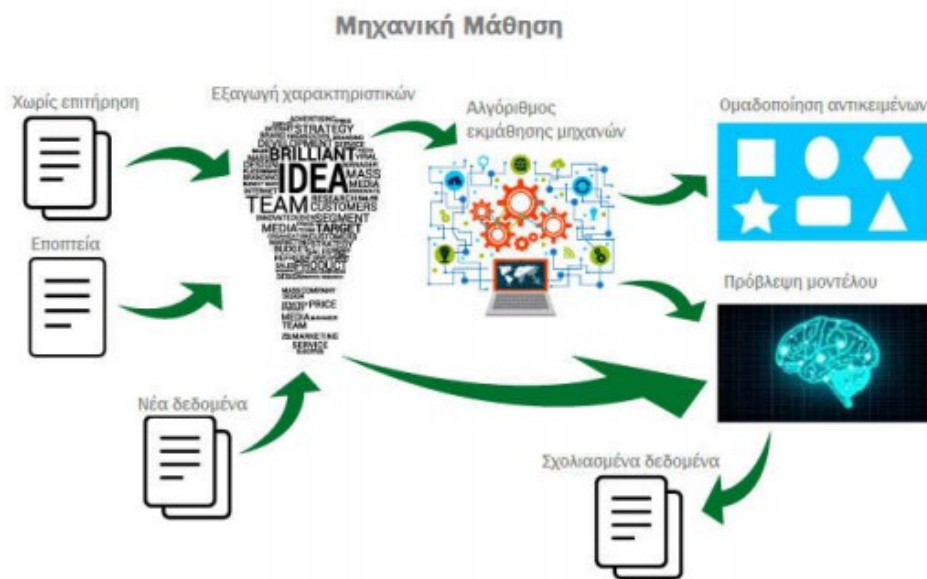
Για τους πολίτες, αλλά και για την ευρύτερη κοινωνία η διαδικασία της συνεχούς αναγνώρισης ψευδών ειδήσεων έχει πλέον πρωταρχική σημασία. Σημαντικότερος στόχος είναι να αποφεύγετε η παραπληροφόρηση ειδικότερα των ανθρώπων με χαμηλότερο δείκτη μόρφωσης αλλά και των μορφωμένων ανθρώπων της κοινωνίας. Καθώς τα άτομα αυτά δεν διαθέτουν γνώσεις σχετικά με τομείς όπως είναι λόγου χάριν τα ιατρικά θέματα στα οποία είναι πολύ δύσκολο να αναγνωριστούν άμεσα οι ψευδείς ειδήσεις από το ευρύτερο κοινό.

Μέχρι σήμερα έχουν δημιουργηθεί και αναπτυχθεί διάφορες τεχνικές μηχανικής μάθησης, οι οποίες στοχεύουν στην άμεση αντιμετώπιση του προβλήματος που αναφέρθηκε προηγουμένως. Πιο συγκεκριμένα, υπάρχουν διαφορετικοί αλγόριθμοι μηχανικής μάθησης.

Οι αλγόριθμοι αυτοί έχουν ως στόχο τους να αναγνωρίσουν μέσω ταξινόμησης τις ψευδείς ειδήσεις στα μέσα κοινωνικής δικτύωσης.

Οι μέθοδοι αυτοί στηρίζονται στις Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines), στην Παλινδρόμηση (Regression), τα Νευρωνικά Δίκτυα (Neural Networks), τη θεωρία Bayes καθώς και στα Δέντρα Αποφάσεων (Decision Trees). (23)

Για να γίνει περισσότερο κατανοητή η διαδικασία της Μηχανικής Μάθησης αρκεί να παρατηρήσει κανείς την εικόνα που βρίσκεται στη συνέχεια.



Εικόνα 3 - Διεργασίες που εφαρμόζονται στη Μηχανική Μάθηση

Διακρίνονται τρεις διαφορετικές προσεγγίσεις για την ανίχνευση των ψευδών ειδήσεων με τη βοήθεια της Μηχανικής Μάθησης. Στη συνέχεια αναφέρονται και περιγράφονται περαιτέρω οι συγκεκριμένες προσεγγίσεις.

### ➤ Εποπτευόμενες - supervised based approaches

Η εποπτευόμενη προσέγγιση είναι γνωστή και ως επιβλεπόμενη. Επίσης, μπορεί κανείς να τη συναντήσει και με τον όρο μάθηση με επίβλεψη. Πρόκειται για τη διαδικασία κατά την οποία εξάγεται ένα συγκεκριμένο μοντέλο που χαρτογραφεί τη σχέση μεταξύ ενός συνόλου περιγραφικών γνωρισμάτων εισόδου και ενός γνωρίσματος εξόδου. (12)

Η εποπτευόμενη μάθηση μπορεί να χωριστεί σε δύο ξεχωριστές κατηγορίες με βάση το είδος της εξόδου. Συγκεκριμένα, εάν η έξοδος είναι συνεχής τότε αναφερόμαστε σε παλινδρόμηση. Εν αντιθέσει, εάν υπάρχουν διακριτές εξοδοί αναφερόμαστε σε κατηγοριοποίηση.

Ο βασικός στόχος της μάθησης με επίβλεψη είναι να κατασκευαστεί ένα μοντέλο, το οποίο με την κατάλληλη εκπαίδευση να είναι ικανό να προβλέπει την έξοδο με κάθε είσοδο.

Η διαδικασία ταξινόμησης των κειμένων πριν τη Μηχανική Μάθηση ήταν μια αρκετά χρονοβόρα διαδικασία, καθώς απαιτούσε μεγάλη παρέμβαση από τον ανθρώπινο παράγοντα. Με την εξέλιξη της Μηχανικής Μάθησης δημιουργήθηκε η λεγόμενη ταξινόμηση εγγράφων με χρήση αλγορίθμων εποπτευόμενης μάθησης. (13) Στην περίπτωση αυτή είναι απαραίτητη η ύπαρξη ενός συνόλου δεδομένων προκειμένου να πραγματοποιηθεί η μάθηση του αλγορίθμου ταξινόμησης. Τα συγκεκριμένα δεδομένα από το σύνολο δεδομένων που προαναφέρθηκε διαθέτουν κατάλληλες ετικέτες. Πρόκειται για μια διαδικασία ταξινόμηση σαφώς λιγότερο χρονοβόρα από ότι παλαιότερα.

Ένα πολύ σημαντικό πλεονέκτημα που παρατηρείται είναι το γεγονός ότι ο συγκεκριμένος αλγόριθμος είναι γενικευμένος, με αποτέλεσμα να έχει τη δυνατότητα να επαναχρησιμοποιείται για άλλες περιπτώσεις. Η χρήση του αλγορίθμου εκ νέου μπορεί να πραγματοποιηθεί ακόμα και στην περίπτωση όπου τα δεδομένα που εισάγονται στον αλγόριθμο να έχουν υποστεί ορισμένες μεταβολές.

Αξιολογείται ως προς την αξιοπιστία του κάθε στοιχείο των κειμένων που τέθηκε προς διερεύνηση. Εν συνεχεία, εξάγονται από τον ταξινομητή ορισμένα χαρακτηριστικά. Ο κύριος στόχος είναι να μπορεί ο ταξινομητής να προσδιορίσει και να αξιολογήσει αυτόματα μια είδηση ως αληθής ή ψευδής.

### ➤ Μη εποπτευόμενες - Unsupervised based approaches

Στη μη επιβλεπόμενη μάθηση (unsupervised learning) η έξοδος του δικτύου βασίζεται στην ικανότητά του να αυτό-οργανώνεται σύμφωνα με τα διανύσματα εισόδου καθώς δεν υπάρχουν αντίστοιχα διανύσματα εξόδου. Σε αυτή τη μέθοδο εκπαίδευσης ως είσοδος δίνεται μια πληροφορία αλλά δεν υπάρχει οποιοδήποτε σύστημα εξωτερικού εκπαιδευτή. (14) Πιο συγκεκριμένα, στη μη επιβλεπόμενη ή μη εποπτευόμενη μάθηση ή μάθηση χωρίς επίβλεψη δεν υπάρχει η επιθυμητή απόκριση καθώς και η επίβλεψη της διαδικασίας με σκοπό να διορθώνονται τυχόν λάθη.

Ως δείγματα για τη μη επιβλεπόμενη μάθηση διαθέτουμε μόνο τα διανύσματα εισόδου και παραλείπονται τα διανύσματα επιθυμητής εξόδου. Η συγκεκριμένη εκπαίδευση πραγματοποιείται με τη χρήση ενός εσωτερικού ελέγχου. Ο κύριος σκοπός του είναι να φτάσει στο σημείο μέσω της εκπαίδευσης η έξοδος να έχει όμοια χαρακτηριστικά με τα διανύσματα εισόδου. Πρόκειται ουσιαστικά για μια εσωτερική μάθηση όπου το δίκτυο εκπαιδεύεται μόνο του και διορθώνει τα λάθη που τυχόν να υπάρχουν στα δεδομένα εισόδου. Όταν δεν υπάρχουν πλέον οι απαιτούμενες διορθώσεις, τότε η διαδικασία της εκπαίδευσης έχει τερματίσει. (14)

### ➤ **Ενισχυτική Μάθηση - Reinforcement Learning**

Η Ενισχυτική Μάθηση - Reinforcement Learning αποτελεί την τρίτη μέθοδο μηχανικής μάθησης. Στην περίπτωση αυτή, ο αλγόριθμος που χρησιμοποιείται εκπαιδεύεται για μια συγκεκριμένη στρατηγική λειτουργιών. Η εκμάθηση πραγματοποιείται μέσω της αλληλεπίδρασης του εκάστοτε συστήματος με το περιβάλλον. (15) Η μέθοδος αυτή χρησιμοποιείται κατά κύριο λόγο σε ότι αφορά προβλήματα Σχεδιασμού - Planning.

Ανάλογα με το εκάστοτε πρόβλημα που καλείται κανείς να βρει λύση μέσω της Μηχανικής Μάθησης μπορεί να χρησιμοποιηθεί η πλέον κατάλληλη μέθοδος από τις τρεις που αναφέρθηκαν και αναλύθηκαν προηγουμένως.

### *Μέθοδοι μηχανικής μάθησης*

Όπως έχει ήδη αναφερθεί, προκειμένου να ξεκινήσει η διαδικασία για την ανίχνευση ψευδών ειδήσεων στα μέσα κοινωνικής δικτύωσης κρίνεται απαραίτητο να πραγματοποιηθεί αρχικά η διαδικασία της κατηγοριοποίησης των εγγράφων - document classification. Κατά τη διαδικασία αυτή κάθε έγγραφο κατατάσσεται σε μια από τις κατηγορίες ενός συνόλου συγκεκριμένων κατηγοριών. Κάθε έγγραφο μπορεί να κατηγοριοποιηθεί σε περισσότερες από μια κατηγορίες του συγκεκριμένου συνόλου. Στην περίπτωση που μελετάμε στη συγκεκριμένη πτυχιακή εργασία θα πρέπει να κατηγοριοποιούνται οι ειδήσεις είτε ως ψευδής είτε ως αληθής με τη μορφή κειμένου.

Στη συνέχεια της ενότητας αυτής θα αναφερθούν και θα αναλυθούν περαιτέρω οι λειτουργίες της Μηχανικής Μάθησης. Πιο συγκεκριμένα, θα αναφερθούμε στον αλγόριθμο Naive Bayes,

στις Μηχανές διανυσμάτων υποστήριξης – SVM, στα τεχνητά νευρωνικά δίκτυα – ANN καθώς και στη Βαθεία Μηχανική Μάθηση – Deep Learning.

### ➤ **Αλγόριθμος Naive Bayes**

Ο αλγόριθμος που ονομάστηκε Naive Bayes είναι ένας από τους πιο γνωστούς και πολύ χρησιμοποιούμενος αλγόριθμος όσον αφορά προβλήματα κατηγοριοποίησης. Προκειμένου να πραγματοποιηθεί η κατηγοριοποίηση, ο αλγόριθμος Naive Bayes κατασκευάζει ένα στατιστικό μοντέλο με τη βοήθεια του θεωρήματος του Bayes. Με τον τρόπο αυτό μέσω του μοντέλου ο αλγόριθμος πραγματοποιεί την πρόβλεψη της κατηγορίας που πρέπει να καταταχθεί κάθε είδηση. (20)

Ο συγκεκριμένος αλγόριθμος για τη διαδικασία της κατηγοριοποίησης θεωρείται ένας από τους πιο απλούς τρόπους. Αξιοσημείωτο είναι το γεγονός ότι ο αλγόριθμος Naive Bayes συγκριτικά με άλλες μεθόδους κατηγοριοποίησης σε πολλές περιπτώσεις έχει επιφέρει πολύ μεγάλα ποσοστά ακρίβειας.

Έχουν παρατηρηθεί πολλά πλεονεκτήματα της συγκεκριμένης μεθόδου κατηγοριοποίησης. Είναι σκόπιμο να αναφερθούμε στο πλέον σημαντικό πλεονέκτημα της. Αναλυτικότερα, έχει τη δυνατότητα να προσεγγίσει τις ορθότερες παραμέτρους που είναι απαραίτητες για το μοντέλο πρόβλεψης. Αυτό ισχύει ακόμα και στην περίπτωση όπου το σύνολο δεδομένων που χρησιμοποιείται για την εκπαίδευση είναι ένα αρκετά μικρό σύνολο.

Στον συγκεκριμένο αλγόριθμο έχουν παρατηρηθεί δύο διαφορετικά μοντέλα όσον αφορά την κατηγοριοποίηση. Πιο συγκεκριμένα, αναφερόμαστε στα Multi – variate Bernouli και Multinomial. Και τα δύο μοντέλα που αναφέρθηκαν προηγουμένως βασίζονται στην κατανομή των λέξεων στα κείμενα που καλούμαστε να κατηγοριοποιήσουμε. Μελετώντας το μοντέλο του Multi – variate Bernouli μπορούμε να αναφέρουμε ότι το εκάστοτε έγγραφο αναπαριστάνεται με τη μορφή ενός διανύσματος αποτελούμενο από διάφορα γνωρίσματα τα οποία έχουν δυαδικές τιμές. (20) Οι τιμές αυτές υποδηλώνουν εάν υπάρχει μια λέξη στο κείμενο ή εάν δεν υπάρχει. Η συγκεκριμένη μέθοδος δεν υπολογίζει εάν μια λέξη έχει εμφανιστεί περισσότερες από μια φορές μέσα στο κείμενο. Εν αντιθέσει, η μέθοδος Multinomial λαμβάνει υπόψη την συχνότητα με την οποία εμφανίζονται οι λέξεις μέσα στο κείμενο. (24)

Όταν πρόκειται για την επίλυση ενός προβλήματος μέσω της ταξινόμησης, το θεώρημα Bayes αναφέρεται στην πιθανότητα που υπάρχει ένα κείμενο να κατατάσσεται σε μια

κατηγορία η οποία συμβολίζεται ως  $\omega_j$ . Το κείμενο αυτό αποτελείται από ένα σύνολο γνωρισμάτων σε μορφή διανύσματος και συμβολίζεται ως  $x_i$ . Ο τύπος που υπολογίζει την συγκεκριμένη πιθανότητα βρίσκεται στη συνέχεια.

$$P(\omega_j | x_i) =$$

Υφίσταται ένα σύνολο το οποίο αποτελείται από όλες τις κατηγορίες στις οποίες μπορεί να ανήκει ένα κείμενο και συμβολίζεται ως  $\{\omega_1, \omega_2, \dots, \omega_n\}$ . Στον παραπάνω τύπο συμβολίζουμε ως  $\omega_j$  την κατηγορία όπου ανήκει το εκάστοτε κείμενο. Με το σύμβολο  $x_i$  που βρίσκεται παραπάνω αναφερόμαστε στο διάνυσμα των γνωρισμάτων. Η πιθανότητα εμφάνισης της κατηγορίας  $\omega_j$  μέσα σε ένα κείμενο συμβολίζεται ως  $P(\omega_j)$ . Ενώ αντίθετα, η πιθανότητα του  $x_i$  γνωρίζοντας την κατηγορία  $\omega_j$  δίνεται από τον συμβολισμό  $P(x_i | \omega_j)$ .

Η συγκεκριμένη πιθανότητα  $P(x_i | \omega_j)$  αναφέρετε συχνά και ως το γινόμενο της κάθε πιθανότητας όλων των γνωρισμάτων. Αυτό μπορεί να συμβεί λόγω της ανεξαρτησίας που παρατηρείται ανάμεσα στα διάφορα γνωρίσματα. Πιο συγκεκριμένα, η πιθανότητα  $P(x_i | \omega_j)$  μπορεί να υπολογιστεί με τον ακόλουθο τύπο:

$$P(x_i | \omega_j) = P(x_1 | \omega_j) \cdot P(x_2 | \omega_j) \cdot P(x_3 | \omega_j) \cdot \dots \cdot P(x_d | \omega_j)$$

$$P(x_i | \omega_j) =$$

Ο τύπος που υπολογίζει την πιθανότητα εμφάνισης ξεχωριστά κάθε ένα από τα γνωρίσματα βρίσκεται στη συνέχεια. Ο συγκεκριμένος τύπος βασίζεται στην κανόνα maximum – likelihood.

$$P(x_i | \omega_j) =$$

Στον παραπάνω τύπο συμβολίζουμε ως τον αριθμό των εμφανίσεων ενός γνωρίσματος το οποίο συμβολίζουμε ως  $x_i$  μέσα σε ένα κείμενο το οποίο με τη σειρά του κατατάσσετε σε μια κατηγορία  $\omega_j$ . Στον παρονομαστή του τύπου αυτού υπάρχει το σύμβολο  $n_j$ , το οποίο εκφράζει το πλήθος των γνωρισμάτων που υπάρχουν στην κατηγορία  $\omega_j$ .

Με τη χρήση του αλγορίθμου Naive Bayes στοχεύουμε στην αύξηση της τιμής της πιθανότητας  $P(\omega_j | x_i)$  στο μέγιστο σημείο στο οποίο μπορεί να φτάσει. Αυτό μπορεί να πραγματοποιηθεί με τη βοήθεια των δεδομένων εκπαίδευσης. Πιο συγκεκριμένα, στοχεύει στην πρόβλεψη της ορθής κατηγοριοποίησης.

### ➤ Μηχανές διανυσμάτων υποστήριξης – SVM

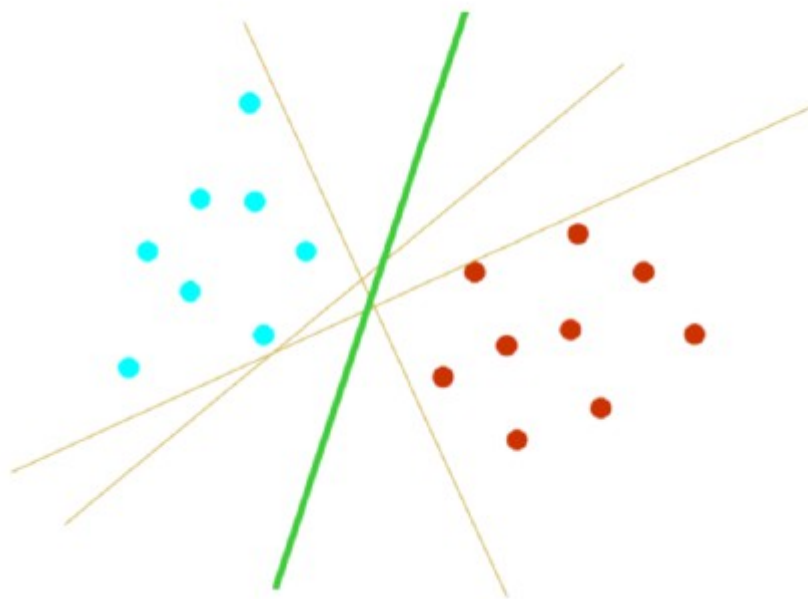
Ο δεύτερος τύπος μηχανικής μάθησης που θα μελετηθεί στα πλαίσια της συγκεκριμένης πτυχιακής εργασίας είναι αυτός των μηχανών διανυσμάτων υποστήριξης, γνωστό και ως Support Vector Machines – SVM. Και αυτός ο αλγόριθμος όπως ακριβώς και ο πρώτος που μελετήθηκε κατατάσσονται στους αλγορίθμους επιβλεπόμενης μάθησης.

Ο αλγόριθμος αυτός πρώτο εμφανίστηκε το 1992 από τρεις επιστήμονες του κλάδου. Συγκεκριμένα, οι Bernhard Boser, Vladimir Vapnik και Isabelle Guyon δημιούργησαν τον αλγόριθμο αυτό. Στηρίχθηκαν στην αρχική θεωρία των Alexey Chervonenkis και Vladimir Vapnik, την οποία κατέγραψαν περίπου το 1963. Οι τρεις αυτοί ερευνητές ανέφεραν τον αλγόριθμο του Support Vector Machines ως μια ιδιαίτερα χρήσιμη μέθοδος για τη κατασκευή τόσο γραμμικών όσο και μη γραμμικών αλγορίθμων ταξινόμησης. (25)

Τόσο οι μηχανές διανυσμάτων υποστήριξης – SVM όσο και οι διάφορες παραλλαγές τους που υπήρξαν κατά καιρούς είναι εξίσου αποτελεσματικοί όσον αφορά την διαδικασία της ταξινόμησης. Θεωρούνται από τους πλέον διαδεδομένους και γνωστούς αλγορίθμους όσον αφορά την απόδοσή τους.

Όσον αφορά τις μηχανές διανυσμάτων υποστήριξης μπορούμε να αναφέρουμε ότι πρόκειται για έναν γραμμικό ταξινομητή. Στην είσοδο του, ο ταξινομητής δέχεται τα γνωρίσματα σε μορφή διανύσματος. Σκοπός του γραμμικού ταξινομητή είναι να πραγματοποιήσει μια πρόβλεψη όσον αφορά την κατηγορία στην οποία πρέπει να καταταχθεί κάθε έγγραφο. Προκειμένου να γίνει αυτό, αρκεί να βρεθεί μέσω υπολογισμών το βέλτιστο δυνατό υπερεπίπεδο. Μέσω του υπερεπιπέδου ξεχωρίζουν τα εκάστοτε δεδομένα στην αντίστοιχη κατηγορία. Πρόκειται ουσιαστικά για μια ευθεία γραμμή, η οποία έχει τη δυνατότητα να ξεχωρίζει τα στοιχεία στις αντίστοιχες κατηγορίες στις οποίες πρέπει να καταταχθούν. (26)

Προκειμένου να γίνει πλήρως κατανοητή η έννοια του υπερεπιπέδου, δηλαδή της ευθείας αυτής γραμμής που ξεχωρίζει τα δεδομένα στις αντίστοιχες κατηγορίες αρκεί να παρατηρήσει κανείς την εικόνα που βρίσκεται στη συνέχεια. (27)



Εικόνα 4 - Παράδειγμα υπερεπιπέδου διαχωρισμού

Στο παραπάνω παράδειγμα πραγματοποιείται ο διαχωρισμός των δεδομένων σε δύο μόνο κατηγορίες. Είναι εύκολο να παρατηρήσει κανείς ότι προκειμένου να διαχωριστούν τα δεδομένα αυτά στην κατηγορία στην οποία ανήκουν δημιουργήθηκαν αρκετές τέτοιου είδους ευθείες.

Κύριος σκοπός της ύπαρξης των μηχανών διανυσμάτων υποστήριξης SVM είναι να βρεθεί η βέλτιστη ευθεία. Πιο συγκεκριμένα, θα πρέπει ο αλγόριθμος αυτός να διαχωρίσει τα δεδομένα και να αυξήσει την απόσταση των δεδομένων αυτών στις κατηγορίες τους.

Ο τύπος που υπολογίζει το υπερεπίπεδο διαχωρισμού των δεδομένων στις αντίστοιχες κατηγορίες δίνεται στη συνέχεια.

$$w \cdot x + b = 0$$

Στον παραπάνω τύπο συμβολίζουμε με  $w$  ένα διάνυσμα βαρών. Το συγκεκριμένο διάνυσμα βρίσκεται κάθετα στο υπερεπίπεδο. Ενώ με το σύμβολο  $b$  αναφερόμαστε στην τιμή κατωφλιού.

Στο παράδειγμα που αναφέρθηκε προηγουμένως προκύπτουν τα εξής. Εάν ένα δείγμα κειμένου ανήκει στη πρώτη κατηγορία ισχύει ο τύπος :

$$w \cdot x_1 + b > 0$$



Ενώ εάν ένα δείγμα κειμένου ανήκει στη δεύτερη κατηγορία ισχύει το εξής:

$$\cdot x_1 + b < 0$$

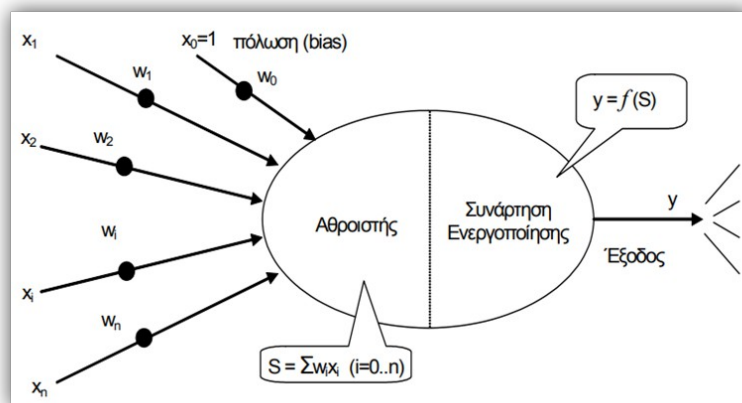
### ➤ Τεχνητά Νευρωνικά Δίκτυα

Πρόκειται για ένα κύκλωμα το οποίο αποτελείται από ένα σύνολο νευρώνων συνδεδεμένοι με κατάλληλο τρόπο προκειμένου να επιτύχει το σκοπό του. Τα τεχνητά νευρωνικά δίκτυα δημιουργήθηκαν από την προσπάθεια των επιστημόνων του κλάδου προκειμένου να προσομοιώσουν όσο το δυνατόν καλύτερα το φυσικό νευρωνικό δίκτυο το οποίο είναι ο ανθρώπινος εγκέφαλος.

Ένας ορισμός που μπορεί να δοθεί για τα Τεχνητά Νευρωνικά Δίκτυα είναι ο ακόλουθος:

Σύμφωνα με τον Haykin ένα νευρωνικό δίκτυο είναι ουσιαστικά ένας επεξεργαστής. Ο επεξεργαστής αυτός αποτελείται από ένα σύνολο απλών μονάδων που στοχεύουν στην αποθήκευση της γνώσης. Επίσης, οι μονάδες αυτές έχουν τη δυνατότητα να υλοποιούν ένα σύνολο διαφόρων χρήσιμων λειτουργιών. (15)

Όπως έχουμε ήδη αναφέρει κάθε νευρώνας αποτελείται από ένα πλήθος νευρώνων. Οι νευρώνες αποτελούν το βασικότερο τμήμα ενός τεχνητού νευρώνα. Είναι σημαντικό να αναφερθούμε περαιτέρω στους τεχνητούς νευρώνες. Πρόκειται για μια μονάδα, η οποία επεξεργάζεται τις εκάστοτε πληροφορίες και τα δεδομένα. Κάθε νευρώνας επεξεργάζεται εξ ολοκλήρου τα τοπικά του δεδομένα. Δέχονται τα δεδομένα τους ως εισόδους. Όπως μπορεί να παρατηρήσει από την εικόνα που ακολουθεί, ένας τεχνητός νευρώνας αποτελείται από τρία διαφορετικά δομικά στοιχεία. Πιο συγκεκριμένα, κάθε νευρώνας αποτελείται από τον αθροιστή τη συνάρτηση ενεργοποίησης καθώς και τις συνάψεις. Οι νευρώνες που υπάρχουν σε ένα νευρωνικό δίκτυο αλληλεπιδρούν συνεχώς μεταξύ τους και είναι συνδεδεμένοι με τις συνάψεις.

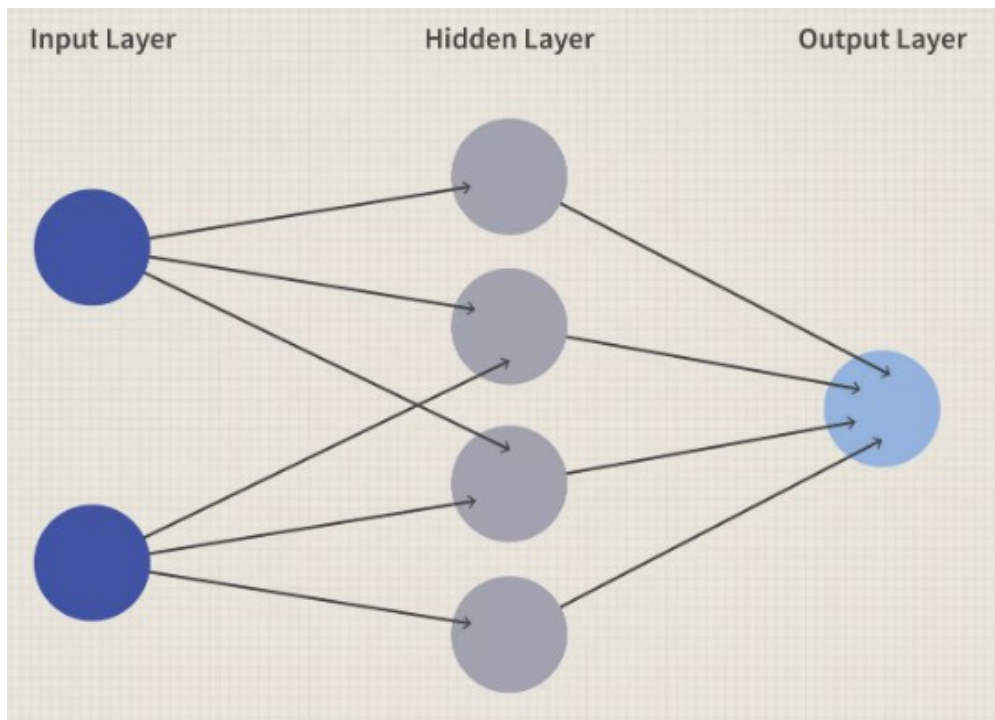


Εικόνα 5 - Τεχνητός Νευρώνας

Υπάρχουν τρεις διαφορετικοί τύποι νευρώνων σε ένα νευρωνικό δίκτυο. Πιο συγκεκριμένα, υπάρχει ο νευρώνας εισόδου, οι κρυμμένοι νευρώνες καθώς και ο νευρώνας εξόδου. Κάθε νευρώνας δέχεται ως είσοδο ένα πλήθος δεδομένων από διάφορες πηγές, όπως λόγου χάριν από προηγούμενους νευρώνες του δικτύου είτε από το εξωτερικό περιβάλλον. Εν συνεχεία, πραγματοποιεί έναν υπολογισμό με τη χρήση των δεδομένων που εισήχθησαν σε αυτόν και τέλος παράγει μια έξοδο. Η συγκεκριμένη έξοδος ενός νευρώνα είτε αποτελεί είσοδο ενός άλλου νευρώνα μέσα στο νευρωνικό δίκτυο είτε πηγαίνει στο εξωτερικό περιβάλλον. (28)

Τα νευρωνικά δίκτυα εκπαιδεύονται προκειμένου να παρουσιάσουν την καλύτερη γενικευτική ιδιότητα, όταν καλούνται να αναγνωρίσουν νέα δεδομένα. Γεγονός που αποτελεί τον βασικότερο σκοπό της κατασκευής των συγκεκριμένων δικτύων. Πιο αναλυτικά, στοχεύουν στην κατασκευή της εσωτερικής τους δομής με τον πλέον κατάλληλο τρόπο προκειμένου να αναγνωρίζουν καινούργια πρότυπα που ταυτίζονται με γνωστά πρότυπα τα οποία έχουν ήδη εκπαιδευτεί. (23)

Στην εικόνα που ακολουθεί αναπαριστάνεται η μορφή ενός τεχνητού νευρωνικού δικτύου. Όπως μπορεί κανείς να παρατηρήσει από την εικόνα 6 υπάρχουν δύο νευρώνες εισόδου, τέσσερις κρυμμένοι νευρώνες καθώς και ένας νευρώνας εξόδου.



Εικόνα 6 - Τεχνητό Νευρωνικό Δίκτυο

Για να γίνει περισσότερο κατανοητή η έννοια των Τεχνητών Νευρωνικών Δικτύων είναι χρήσιμο να αναφερθούμε σε ορισμένες από τις εφαρμογές τους στις μέρες μας. Σήμερα υπάρχουν νευρωνικά δίκτυα που χαρακτηρίζονται ως έτοιμα προϊόντα σε διάφορους τομείς.

Αξίζει να αναφερθούμε στο πρόβλημα της μετατροπής της φωνής σε γραπτό κείμενο και το αντίστροφο. Υπάρχει το πρόγραμμα NETtalk, το οποίο αναλαμβάνει να εκπαιδεύσει το δίκτυο να μπορεί να διαβάζει κείμενα και εν συνεχεία να τα απαγγέλει. Το συγκεκριμένο δίκτυο αποτελείται συνολικά από 309 νευρώνες, τοποθετημένα σε τρία επίπεδα. Η είσοδος διαθέτει επτά ομάδες νευρώνων, ενώ η εκάστοτε ομάδα διαθέτει 29 νευρώνες. Οι 29 αυτοί νευρώνες αναπαριστούν τα 26 γράμματα της αλφαβήτας, την τελεία, το κενό καθώς και το κόμμα. Το μεσαίο επίπεδο του δικτύου αποτελείται συνολικά από 80 νευρώνες, ενώ η έξοδος τους διαθέτει 26 νευρώνες. (29)

Υπάρχει ένα παράθυρο αποτελούμε από επτά χαρακτήρες τους οποίους και καλείται να εξετάσει κάθε φορά το συγκεκριμένο πρόγραμμα. Το παράθυρο μετατοπίζεται κατά ένα χαρακτήρα και έχει τη δυνατότητα να αναγνωρίζει και να διορθώνει τα σφάλματα. Ύστερα από την εκπαίδευση του νευρωνικού δικτύου έχει τη δυνατότητα να εντοπίσει τους κανόνες που υπάρχουν και τέλος αλλάζει αναλόγως και την τιμή των βαρών.

Σε αρχικό στάδιο, η απαγγελία του προγράμματος αποτελούσε μια ακατανόητη διαδικασία. Ύστερα από την περαιτέρω επεξεργασία του νευρωνικού δικτύου, η απαγγελία έφτασε σε νηπιακή μορφή. Ενώ στις μέρες μας, έχει φτάσει σε αναγνωσιμότητα και κατανόηση το μεγάλο ποσοστό του 95%. (29)

### ➤ **Βαθεία Μηχανική Μάθηση**

Η βαθειά μηχανική μάθηση είναι μια υποκατηγορία του κλάδου της Μηχανικής μάθησης. Μπορεί κανείς να την συναντήσει και με την αγγλική της ερμηνεία η οποία είναι Deep Learning. Πρόκειται για μια διαδικασία η οποία συνδυάζει τόσο την διαδικασία των Μηχανών διανυσμάτων υποστήριξης SVM όσο και τα τεχνητά νευρωνικά δίκτυα – ANN. Με τη διαφορά ότι χρησιμοποιούνται σε πολλαπλά κρυφά επίπεδα – hidden layers.

Θεωρείται μια ιδιαίτερα χρήσιμη και συχνά χρησιμοποιούμενη μέθοδος, ειδικότερα στα μέσα κοινωνικής δικτύωσης. Ο λόγος είναι διότι ο τρόπος αυτός εξυπηρετεί όταν οι δημοσιεύσεις είναι μικρές σε περιεχόμενο. (30)

### *Εξόρυξη Δεδομένων & Μηχανική Μάθηση*

Η Μηχανική Μάθηση και η Εξόρυξη δεδομένων πολλές φορές χρησιμοποιούν τις ίδιες μεθόδους για να υλοποιήσουν το στόχο τους. Για το λόγο αυτό, δεν είναι λίγες οι φορές που η μια έννοια επικαλύπτει την άλλη. Πολλοί είναι αυτοί που συγχέουν αυτές τις δύο έννοιες.

Παρατηρείται, όμως μια σημαντική διαφορά ανάμεσα στη μηχανική μάθηση – machine learning και της εξόρυξης δεδομένων – data mining. Πιο συγκεκριμένα η μηχανική μάθηση βασίζεται στην πρόβλεψη. Η πρόβλεψη καθορίζεται από ήδη γνωστές ιδιότητες και χαρακτηριστικά από ένα σύνολο μάθησης. Ενώ αντίθετα, η εξόρυξη δεδομένων βασίζεται στην εύρεση άγνωστων μέχρι στιγμής ιδιοτήτων. Πρόκειται για τη διαδικασία ανακάλυψης γνώσεις από διάφορες βάσεις δεδομένων – Database.

### 3 ΚΕΦΑΛΑΙΟ - ΤΕΧΝΟΛΟΓΙΕΣ ΥΛΟΠΟΙΗΣΗΣ

Το αντικείμενο με το οποίο ασχολείται το τρίτο κεφάλαιο της συγκεκριμένης πτυχιακής εργασίας είναι η αναφορά και η ανάλυση των τεχνολογιών υλοποίησης της διαδικασίας ανίχνευσης ψευδών ειδήσεων. Συγκεκριμένα, θα πραγματοποιηθεί η περιγραφή των υπολογιστικών εργαλείων που συνεισφέρουν στην ανίχνευση των ψευδών ειδήσεων στο διαδίκτυο. Εν συνεχεία, θα πραγματοποιηθεί περαιτέρω ανάλυση των μεθόδων της μηχανικής μάθησης. Τέλος, θα αναφερθούμε στις μετρικές εκτίμησης της απόδοσης των αλγορίθμων ταξινόμησης που χρησιμοποιούνται στη συγκεκριμένη διαδικασία.

#### Υπολογιστικά εργαλεία

Αναμφίβολα οι επιστήμονες που ασχολούνται με την αντιμετώπιση του φαινομένου των ψευδών ειδήσεων έδωσαν ιδιαίτερη έμφαση στην Μηχανική Μάθηση και στις μεθόδους που αυτή χρησιμοποιεί.

Στις μέρες μας οι ερευνητές τόσο στην Ελλάδα όσο και στις υπόλοιπες χώρες παγκοσμίως ασχολούνται με τους τρόπους αντιμετώπισης και ανίχνευσης των ψευδών ειδήσεων στα μέσα κοινωνικής δικτύωσης και γενικότερα στο διαδίκτυο. Η αντιμετώπιση αυτή επιτυγχάνετε συνήθως με τη βοήθεια ειδικών εργαλείων πληροφορικής.

Αξιοσημείωτο είναι το γεγονός ότι οι Έλληνες πολίτες συγκριτικά με τους υπόλοιπες χώρες με βάση το Ευρωβαρόμετρο του Μαρτίου έχουν έρθει σε επαφή συχνότερα με φαινόμενα παραπληροφόρησης. Οι ερευνητές χρησιμοποιούν κατά καιρούς διάφορα προγράμματα για την αντιμετώπιση του φαινομένου. Παραδείγματος χάριν, οι ερευνητές του Ινστιτούτου Τεχνολογιών Πληροφορικής και Επικοινωνιών του Εθνικού Κέντρου Ανάπτυξης και έρευνας σε συνεργασία με ορισμένους ερευνητές της Γερμανίας και άλλων ευρωπαϊκών χωρών στην διαδικασία της λύσης του φαινομένου της παραπληροφόρησης και των ψευδών ειδήσεων χρησιμοποιούν το πρόγραμμα ονόματι InVID. (31)

#### Εξόρυξη δεδομένων (Data mining)

Με τον όρο εξόρυξη δεδομένων γνωστή και ως Data Mining αναφερόμαστε στην εξαγωγή ορισμένων σημαντικών δεδομένων και πληροφοριών από μεγάλες βάσεις δεδομένων. Το Data Mining έχει ολοένα και μεγαλύτερη απήχηση στις μέρες μας. Πρόκειται για την τέχνη της εξόρυξης σημαντικών δεδομένων από ένα μεγάλο πλήθος δεδομένων. (5)

Πρόκειται για την τεχνική εξεύρεσης πληροφοριών από μια μεγάλη ποικιλία από βάσεις δεδομένων χρησιμοποιώντας αλγόριθμους. Η συγκεκριμένη τεχνική χρησιμοποιείται κατά

κύριο λόγο για να εντοπιστούν καινούργιες σχέσεις μεταξύ μεταβλητών. (6) Οι σχέσεις αυτές που δημιουργούνται κατά τη διαδικασία αυτή μπορούν να είναι εμπορικά χρήσιμες.

Η λειτουργία της εξόρυξης δεδομένων αποτελεί μια πολύ χρήσιμη διαδικασία σε αρκετούς τομείς. Πρόκειται για μεθόδους, που αποσκοπούν στην άντληση χρήσιμων πληροφοριών και γνώσεων από ένα σύνολο δεδομένων. Αυτές οι μέθοδοι αποτελούνται από ένα πλήθος μοντέλων που βοηθούν στην αναγνώριση προτύπων, τεχνικών εξόρυξης δεδομένων καθώς και μηχανικής μάθησης. Οι μέθοδοι αυτοί έχουν σκοπό να παράγουν πολύτιμες πληροφορίες, τις οποίες θα χειριστούν κατάλληλα οι αρμόδιοι ώστε να πάρουν χρήσιμες και σωστές αποφάσεις. (32)

Η διαδικασία που ακολουθεί η ανακάλυψη γνώσης από βάσεις δεδομένων απαρτίζεται από πέντε βασικά στάδια. Στο πρωταρχικό στάδιο βρίσκεται η διαδικασία της συλλογής των δεδομένων και εν συνεχεία η προ επεξεργασία αυτών. Ύστερα, ακολουθεί η διαδικασία του μετασχηματισμού και της εξόρυξης των δεδομένων. Και τέλος, πραγματοποιείται η ερμηνεία και η αξιολόγηση τους.

## Μετρικές απόδοσης

Όταν αναφερόμαστε στην απόδοση ενός αλγορίθμου, αναφερόμαστε στην εγκυρότητα των προβλέψεων του. Πιο συγκεκριμένα, προκειμένου να μετρήσουμε την απόδοση των αλγορίθμων για την ταξινόμηση ενός συνόλου, θα πρέπει να χρησιμοποιήσουμε τις τεχνικές μετρικής εκτίμησης της απόδοσης. Οι συγκεκριμένες μετρικές εκτίμησης της απόδοσης δεν είναι ίδιες για κάθε τύπο και για κάθε διαφορετικό αλγόριθμο.

Για τους αλγορίθμους ταξινόμησης κειμένου, οι μετρικές εκτίμησης της απόδοσης είναι οι ακόλουθες:

- Ορθότητα
- Ανάκληση
- Ακρίβεια
- F1-Score

Στη συνέχεια αυτής της ενότητας θα πραγματοποιηθεί περαιτέρω ανάλυση των συγκεκριμένων μεθόδων μετρικής απόδοσης.

## Πίνακας σύγχυσης (confusion matrix)

Στο σημείο αυτό είναι σκόπιμο να αναφερθούμε στον πίνακα σύγχυσης (confusion matrix). Πρόκειται για ένα ιδιαίτερα χρήσιμο εργαλείο για την απεικόνιση όσον αφορά την απόδοση ενός αλγορίθμου που σχετίζεται με την κατηγοριοποίηση.

Ο πίνακας σύγχυσης είναι ένας πίνακας δύο διαστάσεων. Αποτελείται από δύο γραμμές καθώς και δύο στήλες. Στις συγκεκριμένες γραμμές και στήλες αναπαριστάνεται το πλήθος των προβλέψεων που έχουν προκύψει από την διαδικασία της ταξινόμησης μέσω ενός κατάλληλου ταξινομητή. Οι γραμμές του διδιάστατου αυτού πίνακα απεικονίζουν τις προβλέψεις του ταξινομητή, ενώ αντίστοιχα οι στήλες του πίνακα σύγχυσης απεικονίζουν τις ήδη γνωστές εγγραφές. (32)

Οι προβλέψεις, οι οποίες απεικονίζονται σε έναν πίνακα σύγχυσης – confusion matrix μπορούν να διαχωριστούν σε τέσσερις διαφορετικές κατηγορίες, οι οποίες αναφέρονται στη συνέχεια.

### ➤ **True Positives (TP)**

Στην κατηγορία αυτή κατατάσσονται οι εγγραφές που η κατηγοριοποίηση τους πραγματοποιήθηκε επιτυχώς στην κατηγορία “Positive”.

### ➤ **True Negatives (TN)**

Στην κατηγορία αυτή κατατάσσονται οι εγγραφές που σωστά βρέθηκαν να μην αντιστοιχούν στην κατηγορία “Positive”.

### ➤ **False Positives (FP)**

Στην κατηγορία αυτή κατατάσσονται οι εγγραφές που λανθασμένα βρέθηκαν να αντιστοιχούν στην κατηγορία “Positive”.

### ➤ **False Negative (FN):**

Στην κατηγορία αυτή κατατάσσονται οι εγγραφές που λανθασμένα βρέθηκαν να αντιστοιχούν στην κατηγορία “Negative”.

Στη συνέχεια θα αναφερθούμε σε ένα τυχαίο παράδειγμα προκειμένου να γίνει πλήρως κατανοητός ο τρόπος λειτουργίας του πίνακα σύγχυσης που αναφέρθηκε προηγουμένως. Αρχικά, θα υποθέσουμε ότι μια αληθής είδηση μπαίνει στην μεταβλητή όπου ονομάσαμε “Positive”, ενώ αντίθετα μια ψευδής πηγαίνει στην “Negative”. (32) Οι τέσσερις κατηγορίες που αναφέρθηκαν προηγουμένως, έχουν πλέον την ακόλουθη μορφή.

➤ **True Positives (TP)**

Στην κατηγορία αυτή κατατάσσονται οι αληθής ειδήσεις “Positive”, όπου ύστερα από την ταξινόμηση, ο ταξινομητής την τοποθέτησε σωστά στην κατηγορία “Positive”.

➤ **True Negatives (TN)**

Στην κατηγορία αυτή κατατάσσονται οι ψευδής ειδήσεις “Negative”, όπου ύστερα από την ταξινόμηση, ο ταξινομητής την τοποθέτησε σωστά στην κατηγορία “Negative”.

➤ **False Positives (FP)**

Στην κατηγορία αυτή κατατάσσονται οι ψευδής ειδήσεις “Negative”, που λανθασμένα ο ταξινομητής τις κατατάσσει στην κατηγορία “Positive”.

➤ **False Negative (FN):**

Στην κατηγορία αυτή κατατάσσονται οι αληθής ειδήσεις “Positive”, που λανθασμένα ο ταξινομητής τις κατατάσσει ως ψευδής στην κατηγορία “Negative”.

Στη συνέχεια παραθέτουμε τον πίνακα σύγχυσης όπως αυτός προέκυψε από το παράδειγμα που αναφέρθηκε προηγουμένως.

		Actual	
		Positive	Negative
Prediction	Positive	TP	FP
	Negative	FN	TN

Εικόνα 7 - Παράδειγμα πίνακα σύγχυσης

Προηγουμένως έγινε αναφορά στις τέσσερις μετρικές εκτίμησης της απόδοσης των αλγορίθμων ταξινόμησης. Στο σημείο αυτό, θα πραγματοποιηθεί η περαιτέρω ανάλυση αυτών.

### Ακρίβεια (Accuracy)

Όσον αφορά την Ακρίβεια -Accuracy αναφερόμαστε στον λόγο των προβλέψεων που προέκυψαν από τη διαδικασία του ταξινομητή και θεωρήθηκαν σωστές με το συνολικό αριθμό προβλέψεων. (33)

Ο μαθηματικός τύπος, ο οποίος μπορεί να υπολογίσει την ακρίβεια – accuracy είναι ο ακόλουθος:



$$\text{Accuracy} = \quad (3.1)$$

Στο σημείο αυτό είναι σκόπιμο να αναφέρουμε και την τιμή του σφάλματος. Ο τύπος που υπολογίζει την τιμή του σφάλματος είναι ο εξής:

$$\text{Error Rate} = 1 - \text{Accuracy} \quad (3.2)$$

### Ανάκληση (Recall)

Όσον αφορά την Ανάκληση – Recall είναι ικανή να μας δώσει χρήσιμες πληροφορίες. Αυτή η μετρική εκτίμησης μας δίνει το ποσοστό των κειμένων τα οποία έχουν καταταχθεί στην κατηγορία πραγματικά ψευδής. (33)

Ο τύπος (3.3) είναι αυτός που μπορεί να υπολογίσει την ανάκληση και βρίσκεται στη συνέχεια:

$$\text{Recall} = \quad (3.3)$$

Αν παρατηρήσει κανείς τον τύπο αυτόν που υπολογίζει τη συγκεκριμένη μετρική, θα αντιληφθεί ότι στον παρονομαστή υπάρχει και η τιμή FN . Ο λόγος που συμβαίνει αυτό είναι διότι για να υπολογιστεί σωστά ο συνολικός αριθμός των ψευδών ειδήσεων θα πρέπει να συμπεριληφθούν τόσο οι πραγματικά ψευδής ειδήσεις όσο και οι ειδήσεις που λανθασμένα χαρακτηρίστηκαν ως αληθής (FN).

### Ακρίβεια προσέγγισης (Precision)

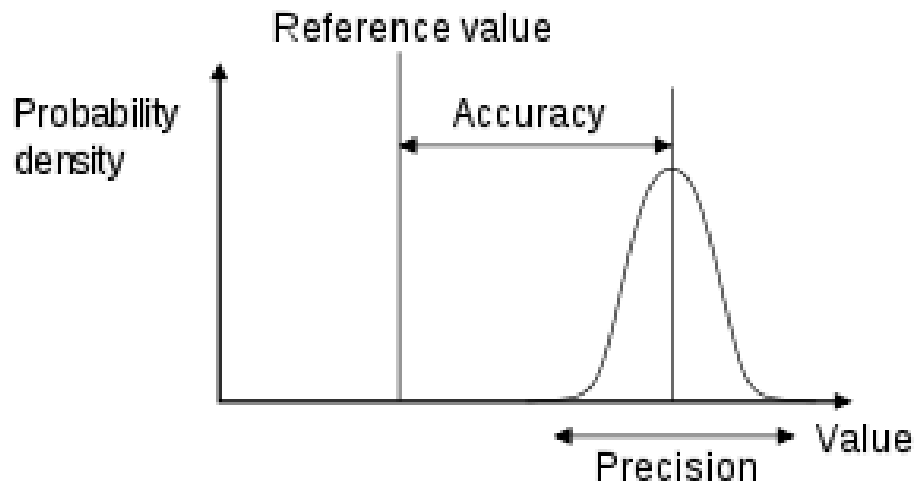
Με την ακρίβεια – precision αναφερόμαστε στο πλήθος των ειδήσεων το οποίο ο ταξινομητής ορθά κατέταξε στην κατηγορία πραγματικά ψευδής ειδήσεις. Ο μαθηματικός τύπος, ο οποίος υπολογίζει την τιμή της ακρίβειας είναι ο τύπος 3.4 που βρίσκεται στην συνέχεια:

$$\text{Precision} = \quad (3.4)$$

Όπως μπορεί κανείς εύκολα να παρατηρήσει στον αριθμητή του παραπάνω τύπου βρίσκεται το TP, ενώ ο παρονομαστής εκφράζει το σύνολο των κειμένων τα οποία ο εκάστοτε ταξινομητής έχει κατατάξει ως ψευδής.

Τη διαφορά ανάμεσα στις έννοιες Accuracy και Precision μπορεί κανείς να τη διακρίνει εύκολα από την εικόνα που ακολουθεί. Η ακρίβεια – accuracy είναι η ορθότητα των αποτελεσμάτων μέτρησης όσον αφορά πραγματικές μετρήσεις. Εν αντιθέσει, όταν οι

συνθήκες μένουν αναλλοίωτες το precision είναι ο βαθμός κατά τον οποίο οι ίδιες μετρήσεις επιφέρουν τα ίδια αποτελέσματα.



Εικόνα 8 - Επεξήγηση των Accuracy και Precision

### F1- Score

Όσον αφορά την μετρική F1-Score μπορούμε να αναφέρουμε ότι πρόκειται για τον μέσο όρο μεταξύ των μετρικών της ακρίβειας προσέγγισης – Precision και της ανάκλησης Recall. Ο μαθηματικός τύπος που δίνει την τιμή της μετρικής F1- Score βρίσκεται στη συνέχεια:

$$F1_{SCORE} = \quad (3.5)$$

Ιδανικά οι τιμές της ακρίβειας καθώς και της ανάκλησης θα πρέπει να είναι όσο το δυνατόν υψηλότερες. Η τιμή 1 για την μετρική  $F1_{SCORE}$  είναι η καλύτερη τιμή που θα μπορούσε να πάρει, εν αντιθέσει με την τιμή 0 που είναι και η χειρότερη περίπτωση. (20)

Οι μετρικές που αναφέρθηκαν και αναλύθηκαν προηγουμένως χρησιμοποιούνται με σκοπό να μετρήσουν την απόδοση ενός αλγορίθμου ταξινόμησης. Η ιδανικότερη περίπτωση είναι όταν οι τιμές των FP και FN να είναι μηδενικές. Όταν συμβαίνει κάτι τέτοιο, που είναι και πιο ιδανικό σενάριο τότε ο αλγόριθμος είναι κατά εκατό της εκατό επιτυχής. Πρόκειται ουσιαστικά για έναν αλγόριθμο ταξινόμησης ο οποίος έχει την ικανότητα να ταξινομήσει σωστά την κατηγορία στην οποία ανήκει το κάθε κείμενο και η κάθε είδηση στην περίπτωση μας. Ένας τέτοιου είδους αλγόριθμος όμως στην πράξη είναι αδύνατον να δημιουργηθεί. (34)

Προκειμένου να εξάγουμε σωστά και χρήσιμα συμπεράσματα από τις μετρικές θα πρέπει να τις χρησιμοποιούμε κατάλληλα για κάθε ένα διαφορετικό πρόβλημα που καλούμαστε να λύσουμε. Σε κάθε διαφορετική περίπτωση θα πρέπει να κρίνεται σε ποια από τις μετρικές

είναι ωφέλιμο να δοθεί η μεγαλύτερη βαρύτητα. Είναι εφικτό στο εκάστοτε πρόβλημα να πραγματοποιείται προσπάθεια βελτιστοποίησης μιας από τις τιμές FP και FN. (33)

Προκειμένου αυτό να γίνει περισσότερο κατανοητό είναι χρήσιμο να αναφερθούμε σε ένα παράδειγμα. Πιο συγκεκριμένα, εάν η κατηγοριοποίηση στο παράδειγμα μας πραγματοποιείται ανάμεσα σε ασθενής πάσχοντες ή μη, τότε είναι σκόπιμο να μειώσουμε ιδιαίτερα την τιμή της FN. Ο λόγος είναι ότι αν δεν μειωθεί σε μεγάλο βαθμό ο αριθμός αυτός ελλοχεύει ο κίνδυνος οι άνθρωποι να νοσούν από τη συγκεκριμένη νόσο αλλά να καταταχθούν στην κατηγορία μη πάσχοντες.

## 4 ΚΕΦΑΛΑΙΟ - ΥΛΟΠΟΙΗΣΗ

Στην συγκεκριμένη εργασία, επικεντρωθήκαμε στην μελέτη ορισμένων υλοποιήσεων που αφορούν την επεξεργασία φυσικής γλώσσας (Natural Language Processing – NLP) και συγκεκριμένα, στο πρόβλημα του εντοπισμού ψευδών ειδήσεων όσον αφορά την κατάσταση της πανδημίας του COVID-19.

### Πρόβλημα

Για την συγκεκριμένη, υπό μελέτη, υλοποίηση, ως ερέθισμα λειτούργησε ένας παγκόσμιος διαγωνισμός που ξεκίνησε αρχές Οκτωβρίου του 2020 και ολοκληρώθηκε στις 25 Δεκεμβρίου του ίδιου έτους και αφορά την εύρεση τρόπων αντιμετώπισης δημοσιεύσεων που παραπληροφορούν σχετικά με την γενικότερη κατάσταση του COVID-19 τόσο για την Αγγλική γλώσσα όσο και για την Φιλιππινέζικη. Εδώ, λόγω χρήσης, επικεντρωθήκαμε στην μελέτη της υλοποίησης που κέρδισε στον διαγωνισμό, για την πρώτη γλώσσα.

### Δεδομένα

Οι διοργανωτές του συγκεκριμένου διαγωνισμού προσφέρουν οι ίδιοι σετ δεδομένων τόσο για τα στάδια της εκπαίδευσης, του validation όσο και του testing στα οποία περιέχεται μία πληθώρα από δημοσιεύσεις ποικίλων κοινωνικών δικτύων όπως το Facebook, Instagram, Twitter κτλ., τα οποία έχουν κατηγοριοποιηθεί ως fake ή real, για παράδειγμα:

*If you take Crocin thrice a day you are safe – Fake*

*Wearing mask can protect you from the virus – Real*

### Κατηγοριοποίηση του προβλήματος

Με βάση τα μέχρι τώρα δεδομένα του προβλήματος, είναι εύκολο να κατανοήσουμε ότι αναφέρεται σε ένα NLP ζήτημα. Συγκεκριμένα, υπάρχει μία πληθώρα δεδομένων που προέρχονται από την ανθρώπινη γλώσσα (με την μορφή δημοσιεύσεων εδώ) και ζητάμε την βοήθεια της Τεχνητής Νοημοσύνης στην επεξεργασία και την κατηγοριοποίηση μίας πρότασης ως αληθής ή ψευδής σε σύγκριση με τις έως τώρα γνωστές πληροφορίες για την πανδημία. Πριν προχωρήσουμε στην ανάλυση της μεθόδου που ακολούθησαν οι ερευνητές, είναι σημαντικό να επεκταθούμε συνοπτικά σε ορισμένες βασικές έννοιες της μηχανικής μάθησης, των νευρωνικών δικτύων και εν τέλει των NLP διαδικασιών.

## Natural Language Processing

Όπως έχει ήδη ειπωθεί, η διαδικασία που περιλαμβάνει την μάθηση της ανθρώπινης γλώσσας από την τεχνητή νοημοσύνη αποτελεί μία μεγάλη πρόκληση λόγω της σύγκρουσης μεταξύ της φύσης της ανθρώπινης γλώσσας και των δεδομένων που λαμβάνουν οι υπολογιστές. Οι αρχιτεκτονικές μηχανικής μάθησης, απαιτούν οι είσοδοί τους να είναι δομημένοι, ή τουλάχιστον να υπακούν σε ορισμένους βασικούς κανόνες ώστε να μπορούν εν δυνάμει να δομηθούν. Η ανθρώπινη γλώσσα, όμως, αρκετά συχνά περιλαμβάνει υψηλά επίπεδα διαφορούμενων νοημάτων και αμφιβολίας καθώς και χαρακτηρίζεται από μία μη – δομημένη φύση. Αυτή η ασάφεια στον γλωσσικό παράγοντα είναι θεμιτό να αναλυθεί περαιτέρω:

- Λεξική Ασάφεια: Όταν μία λέξη μπορεί να σημαίνει παραπάνω από ένα πράγμα.
- Συντακτική Ασάφεια: Όταν μία φράση, λόγω της σύνταξής της μπορεί να ερμηνευθεί με παραπάνω από έναν τρόπους.
- Σημασιολογική Ασάφεια: Όταν μία φράση, λόγω της εγγενούς σημασίας της μπορεί να ερμηνευθεί με παραπάνω από έναν τρόπους.
- Αναφορική Ασάφεια: Όταν σε ένα σύνολο φράσεων, λέξεις όπως αυτός/αυτή/αυτό μπορούν, εν δυνάμει να αναφέρονται σε περισσότερο από ένα πρόσωπο/αντικείμενο
- Πραγματική Ασάφεια: Όταν η φράση είναι ασαφής εκ φύσεως με τον τρόπο που διατυπώθηκε

Ως πανάκεια στο παραπάνω πρόβλημα, παρουσιάζονται οι NLP διαδικασίες οι οποίες προσπαθούν, μέσω συγκεκριμένων βημάτων – σταδίων να προσδώσουν μία δομή σε τέτοιου είδους δεδομένα, καθιστώντας τα κατάλληλα για επεξεργασία από τις αντίστοιχες αρχιτεκτονικές νευρωνικών δικτύων. Για να μπορέσουμε, όμως, να κατανοήσουμε, τις περισσότερο δυσνόητες έννοιες των NLP, θα πρέπει αρχικά να γνωρίσουμε, σύντομα, τον κόσμο της τεχνητής νοημοσύνης και ιδιαίτερα των νευρωνικών δικτύων.

## Θεωρία Νευρωνικών δικτύων

Αρχικά, είναι σημαντικό να κατανοήσουμε γιατί μας είναι χρήσιμα τα νευρωνικά δίκτυα. Όπως έχει προαναφερθεί, σε αρκετές περιπτώσεις, στη σύγχρονη εποχή, υπάρχουν προβλήματα που αφορούν είτε την πρόβλεψη (regression) ορισμένων τιμών, είτε την κατηγοριοποίηση (classification) σε γνωστές ομάδες (supervised) είτε άγνωστες (unsupervised). Αν θέλαμε, να ομαδοποιήσουμε την μηχανική μάθηση και να αναφέρουμε

λίγο πιο τυποποιημένα τις μεγάλες της κατηγορίες θα μπορούσαμε να έχουμε την εξής απαρίθμηση

- Μάθηση με επιτήρηση (Supervised Learning): Όπως αναφέραμε εδώ οι αλγόριθμοι στέλνουν ως εισόδους τα απαραίτητα δεδομένα και παράγουν αποτελέσματα, οπότε στόχους του υπολογιστικού συστήματος είναι να αντιστοιχεί την κάθε είσοδο με το αντίστοιχο αποτέλεσμα, όντας το αποτέλεσμα ήδη γνωστό. Τα λάθη, δηλαδή οι διαφορές μεταξύ των πραγματικών και θεμιτών εξόδων, ‘μεταδίδονται’ στις διαδικασίες του νευρωνικού δικτύου με αποτέλεσμα την διόρθωση των βαρών ώστε να ελαχιστοποιηθούν. Σημαντικό είναι να αναφερθεί η μεγάλη ποσότητα και καλή ποιότητα δεδομένων που απαιτείται για την εκπαίδευση ενός νευρωνικού δικτύου στις περιπτώσεις εκπαίδευσης με επιτήρηση, καθώς σε αντίθεση περίπτωση θα έχουμε προβλήματα με την επίδοση.
- Μάθηση χωρίς επιτήρηση (Unsupervised Learning): όπως αναφέρθηκε και στο παράδειγμα, οι αλγόριθμοι μάθησης πρέπει να βρουν μόνοι τους κάποια δομή στα υπάρχοντα δεδομένα, χωρίς να έχει δοθεί κάποια τέτοια κατεύθυνση από τον ερευνητή. Έτσι, βασιζόμαστε, ουσιαστικά στο αλγόριθμο να εξάγει συμπεράσματα σχετικά με τις ομαδοποιήσεις και αλληλοσυσχετίσεις. Επί του παρόντος, η χωρίς επίβλεψη μάθηση δεν είναι καλά κατανοητή, λόγω των πολύπλοκων υπολογισμών – αριθμητικών πράξεων που λαμβάνουν χώρο. Το βασικό χαρακτηριστικό αυτών των συστημάτων είναι ότι, αρκετές φορές τα συμπεράσματα που εξάγουν, μπορεί να μην γίνονται κατανοητά από τον ανθρώπινο νου, με την έννοια ότι οι συσχετίσεις που παρουσιάζονται να μην βγάζουν κάποιο νόημα.

Μία επιπλέον, λιγότερο γνωστή διαδικασία, είναι η μάθηση με ενίσχυση (Reinforcement Learning), όπου οι αλγόριθμοι αλληλοεπιδρούν με το περιβάλλον και τα δεδομένα που παίρνουν με σκοπό να φτάσουν στο αποτέλεσμα. Αυτό σημαίνει ότι υπάρχει μία μορφή ανάδρασης στη συγκεκριμένη περίπτωση, σε πολύ πιο ενεργό βαθμό, σε σχέση με πριν.

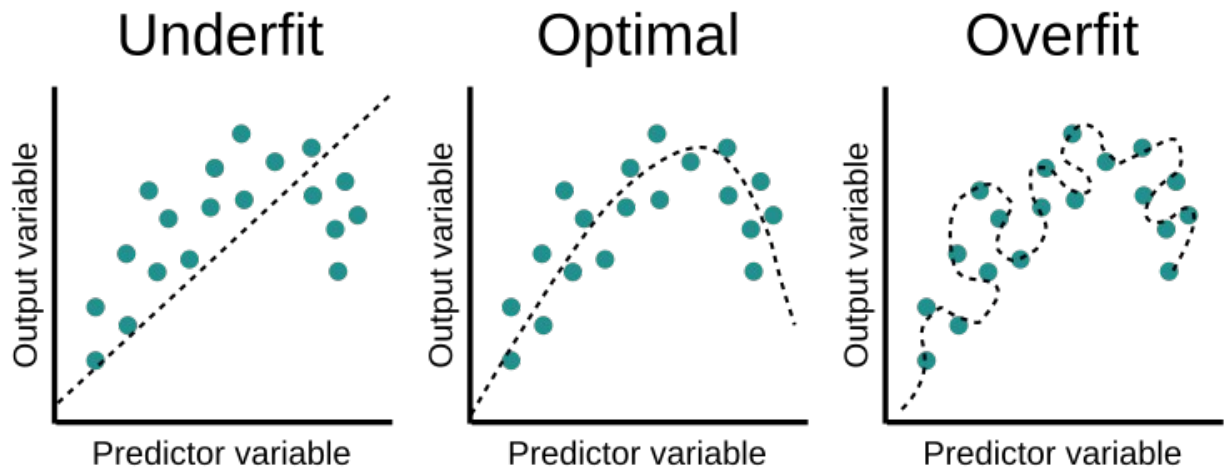
Θα αναρωτηθεί κάποιος, φυσικά, εάν, ανάμεσα στη μάθηση με επιτήρηση και στη μάθηση χωρίς επιτήρηση, προκύπτει και κάποια άλλη κατηγορία. Αυτή υπάρχει και ονομάζεται διαδικασία ημι-επιτήρησης μάθησης (semi - Supervised), όπου οι αλγόριθμοι μάθησης δίνουν σχεδόν μη επαρκείς εισόδους δεδομένων, γεγονός που οδηγεί στο να λείπουν κάποιοι αντίστοιχοι στόχοι.

Επίσης, αναφέραμε και εκτός από την επιτήρηση, τα προβλήματα που μπορούν να επιλύσουν τα προβλήματα μηχανικής μάθησης, αυτά μπορούν αν χωριστούν στα παρακάτω:

- Έχουμε την **Ταξινόμηση** (Classification), όπου εδώ, τα δεδομένα που έρχονται ως είσοδοι χωρίζονται σε δύο ή περισσότερες κατηγορίες, άρα, αποτελεσματικά, το σύστημα πρέπει να αντιστοιχήσει την καινούργια είσοδο σε μία ή περισσότερες κατηγορίες, αντίστοιχα.
- Έχουμε την **Παλινδρόμηση** (Regression), όπου εδώ τα αποτελέσματα που παίρνουμε ανήκουν σε συνεχές εύρος και όχι διακριτό.
- Επίσης, υπάρχει η **Συσταδοποίηση** (Clustering), μία μη επιτηρούμενη μέθοδος όπου τα δεδομένα εισόδων πρέπει να χωριστούν σε άγνωστες κατηγορίες (γεγονός που δικαιολογείται και από το unsupervised)

Σε λιγότερο διάστημα προβλήματα υπάρχουν, εκείνα που πραγματεύονται την εκτίμηση της πυκνότητας πιθανότητας (εύρεση χώρου για τα δεδομένα) αλλά και εκείνα της μείωσης διαστασιμότητας (Dimensionality Reduction), σε περιπτώσεις προβλημάτων μεγάλων δεδομένων (Big Data) όπου ζητάμε την απλοποίηση εισόδων σε μικρότερο χώρο, όχι μόνο για λόγους οικονομίας υλικών αλλά και βελτίωσης της αποδοτικότητας του μοντέλου, ώστε να αποφύγουμε προβλήματα όπως το overfitting.

Αναφερόμενοι στο πρόβλημα της υπερπροσαρμογής (Overfitting), που αποτελεί ένα από τα πιο συχνά προβληματικά φαινόμενα στον τομέα της μηχανικής μάθησης, κατά το ίδιο, το παραγόμενο μοντέλο κατηγοριοποίησης, με το τέλος της εκπαίδευσης, είτε στην αρχή είτε στο τέλος ή και ενδιάμεσα, μπορεί να παρουσιάσει υψηλούς δείκτες σχετικά με την ικανότητα επιτυχημένης πρόβλεψης δεδομένων σε ένα γνωστό σετ, γεγονός που δεν θα μας ανησυχούσε, από μόνο του. Το πρόβλημα αρχίζει, με την ικανότητα που παρουσιάζει στην εκτίμηση νέων τιμών, αφού κατά το overfitting, παρουσιάζει φανερή αδυναμία να πραγματοποιήσει προβλέψεις στις περιπτώσεις όπου έρθουν καινούργια – άγνωστα δεδομένα. Ένα απλοϊκό παράδειγμα παρουσιάζεται στην παρακάτω εικόνα (36), παρατηρούνται τρία διαφορετικά σενάρια ανάλογα με τις περιπτώσεις, με την μεσαία να είναι η καλύτερη δυνατή προσεγγίσεις – γραμμές. Ενώ η τρίτη, στο γνωστό σετ δεδομένων θα παρουσιάσει αυξημένους δείκτες επιτυχίας, δεν θα μπορέσει να γενικεύσει αυτά τα αποτελέσματα στις περιπτώσεις καινούργιων δεδομένων, σε σύγκριση με την δεύτερη καμπύλη.



Εικόνα 9: Underfit/Overfit

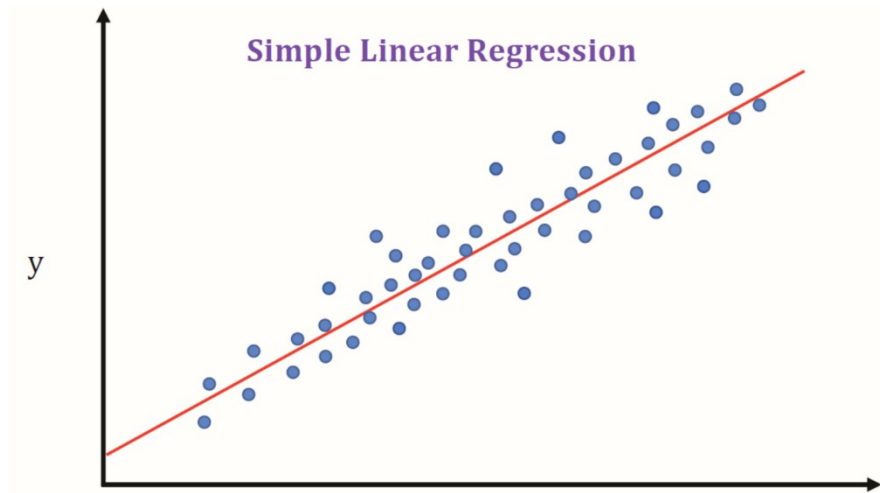
Τέλος, ακριβώς το αντίθετο φαινόμενο εντοπίζεται στη αποτυχία κατασκευής μίας προσεγγιστικής συνάρτησης ταξινόμησης για το δεδομένο σετ, δηλαδή στην πρώτη καμπύλη από πάνω. Είναι ένα φαινόμενο που ονομάζεται *Underfitting* και συχνά, είναι ήδη έκδηλο από τους χαμηλούς δείκτες αξιολόγησης του κατηγοροποιητή, γεγονός που φαίνεται από την αρχή της εκπαίδευσης.

Στη συνέχεια είναι σημαντικό, να προχωρήσουμε σε μία μικρή αναφορά των αλγορίθμων μηχανικής μάθησης, πριν αναφερθούμε στις αρχιτεκτονικές νευρωνικών δικτύων. Οι βασικοί αλγόριθμοι τους οποίους θα αναλύσουμε αποτελούν τα

- Δύο ήδη διαφορετικών παλινδρομήσεων (linear και logistic)
- kNN Neighbors
- Ο Naive Bayes
- Το δένδρο αποφάσεων
- Support Vector Machines.

Αρχικά, για την γραμμική παλινδρόμηση, η ίδια χρησιμοποιείται με σκοπό την εύρεση γραμμικής σχέσης μεταξύ των (37) μεταβλητών που επιδιώκουμε να προβλέψουμε και των εισόδων τους. Υπάρχει η απλή και η πολλαπλή όπου αντίστοιχα συγκρίνουμε πολλές εισόδους εξόδους ή μία.



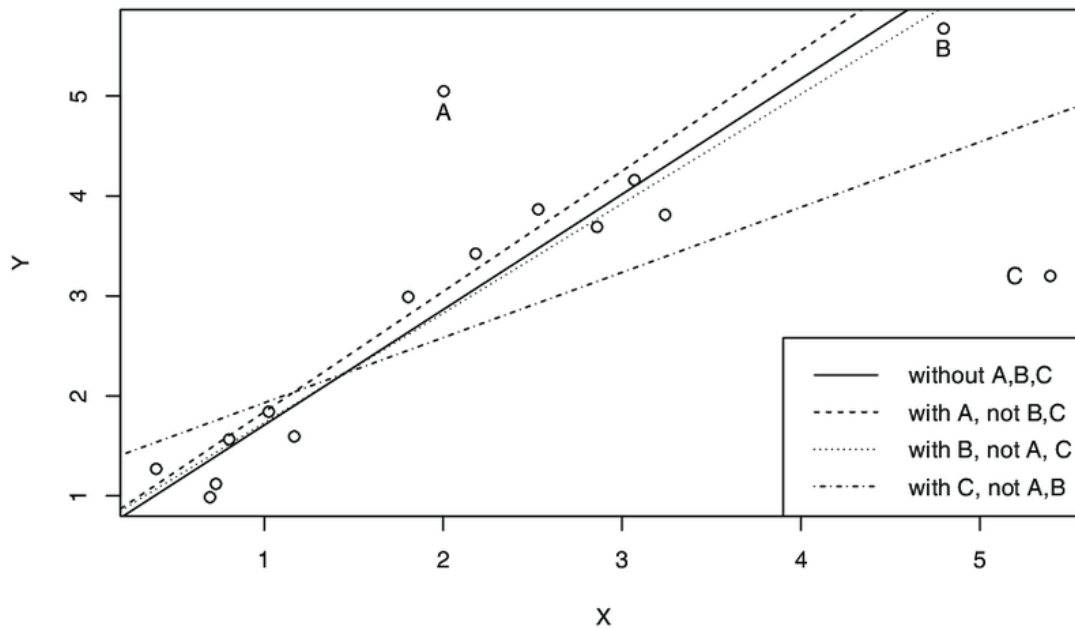


Εικόνα 10: Απλή γραμμική παλινδρόμηση

Στην περίπτωση της απλής, η εξίσωση που περιγράφει την γραμμή που φαίνεται στην εικόνα μπορεί να περιγραφεί με το παρακάτω:

Το επόμενο αρκετά βασικό πρόβλημα είναι πώς τραβιέται η συγκεκριμένη γραμμή, δηλαδή πώς ταιριάζονται σε αυτή τα διαθέσιμα δεδομένα κάθε φορά. Η διαδικασία αυτή ονομάζεται fit. Υπάρχουν διάφοροι αλγόριθμοι που μπορούν να το πραγματοποιήσουν αυτό, με τον πιο γνωστό να σχετίζεται με την μέθοδο των ελαχίστων τετραγώνων (Least Squares), όπου τετραγωνίζουμε τις αποστάσεις κάθε σημείου από τη γραμμή. Πριν προχωρήσουμε στην επόμενη μέθοδο, είναι ορθό να μιλήσουμε γιατί η χρήση της συγκεκριμένης μεθόδου είναι αρκετά περιορισμένη, γεγονός που οφείλεται στην απλότητα της μεθόδου αλλά και το γεγονός ότι είναι αρκετά ευαίσθητες στην ύπαρξη των outlier τιμών.

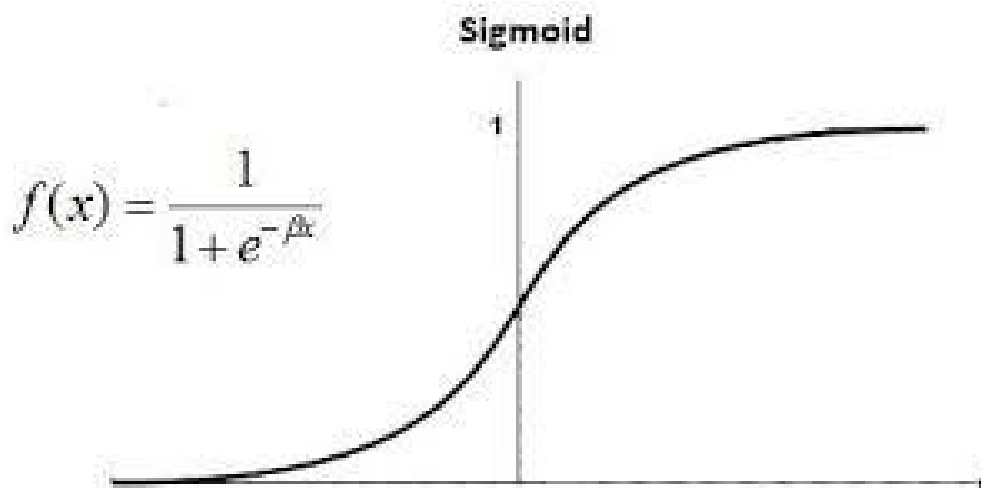
Το παραπάνω φαίνεται και από την εικόνα, δηλαδή, πώς το γεγονός ότι συμπεριλαμβάνουμε αρκετές outlier τιμές, οδηγεί σε μεγάλες διαφορές της τελικής ευθείας.



Εικόνα 11: Επιρροή των outliers

Στη συνέχεια, προχωράμε σε μία παραλλαγή της παλινδρόμησης και ιδιαίτερα στον τι είδους γραμμή κατασκευάζεται.

Εδώ, αναφερόμα (38)στε στην κατηγορία των προβλημάτων παλινδρόμησης κατά τα οποία η εξαρτημένη μεταβλητή μπορεί να λάβει δυαδικές τιμές (0 ή 1, όχι/ναι κτλ). Η ονομασία του, οφείλεται στην μορφή της γραμμής, η οποία είναι η λογιστική (logistic) συνάρτηση, η οποία αλλιώς ονομάζεται και σιγμοειδής, λόγω του σχήματος της:



Εικόνα 12: Σιγμοειδής Συνάρτηση

Ο μαθηματικός τύπος της είναι ο παρακάτω, κατά τον οποίο, για οποιαδήποτε πραγματικό αριθμό, υπάρχει μία αντίστοιχη μεταξύ (39) του 0 και 1. Αντίστοιχα, αν λάβουμε υπόψιν την

γραμμική παλινδρόμηση θα μπορούσαμε να σχηματίσουμε τους δύο επόμενους μαθηματικούς τύπους.

Για τον υπολογισμό των τιμών του  $b$  όρου, από το διαθέσιμο σετ δεδομένων, χρησιμοποιείται, συνήθως ο αλγόριθμος του υπολογισμού της μέγιστης πιθανότητας (maximum likelihood estimation), όπου επιδιώκουμε να επιλέξουμε αυτές τις τιμές  $b$  που θα οδηγήσουν το μοντέλο σε προβλέψεις πολύ κοντά στο 1 για την κατηγορία που επιθυμούμε και, αντίστοιχα, κοντά στο 0 για την άλλη.  $\sigma$

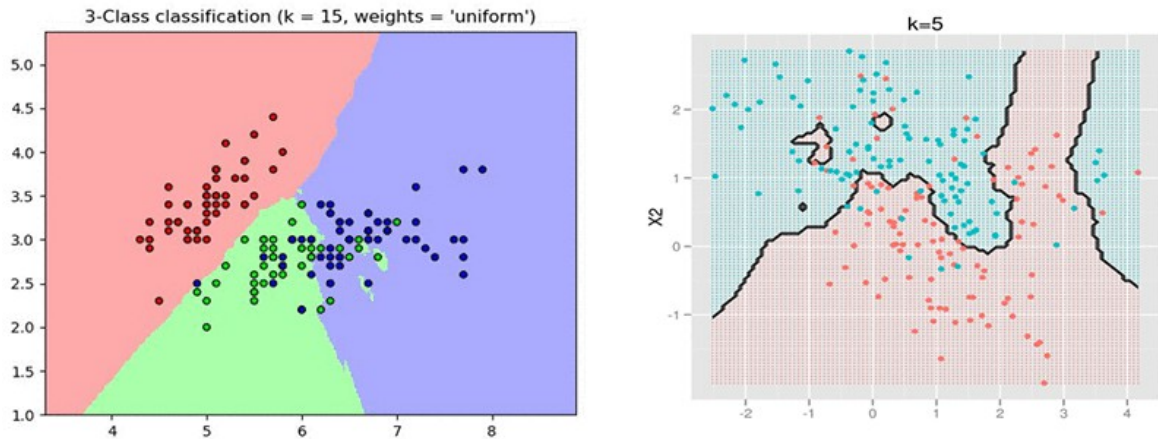
Ο επόμενος αλγόριθμος, που πρέπει να αναλύσουμε είναι ο kNN Neighbors.

Ο αλγόριθμος  $k$  (nearest neighbor) αριθμού πλησιέστερων γειτόνων αποτελεί άλλον έναν αρκετά σημαντικό αλγόριθμο τόσο για διαδικασίες παλινδρόμησης όσο και τις αντίστοιχες κατηγοριοποίησης. Εδώ, το σημαντικό είναι ότι χρησιμοποιούμε της κοντινότερες, από θέμα ομοιότητας, τιμές ώστε να τις ομαδοποιήσουμε σε γκρουπ με τα περισσότερα κοινά χαρακτηριστικά, γεγονός που έχει οδηγήσει στην χρήση της σε ποικιλία προβλημάτων κατηγοριοποίησης βιομηχανία. Η βασική του λογική με την οποία κινείται, βασίζεται στον υπολογισμό της απόστασης (η πιο γνωστή είναι η ευκλείδεια απόσταση) μεταξύ των δεδομένων διαφορετικών κατηγοριών και της, καινούργιας – υπό εκτίμηση, μέτρησης.

Το αποτέλεσμα του παραπάνω, αποτελεί το πρώτο βήμα, το οποίο ακολουθεί η αύξουσα ταξινόμηση των (40) προηγούμενων αποτελεσμάτων και αποφασίζεται με βάση τον αριθμό πλησιέστερων γειτόνων η κατηγορία της νέας τιμής.

Για παράδειγμα, έστω ότι θέλουμε να χρησιμοποιήσουμε αξιοποιήσουμε τους πλησιέστερους  $k=5$  γείτονες για ένα συγκεκριμένο πρόβλημα κατηγοριοποίησης. Στην περίπτωση μας, δηλαδή αναφερόμαστε σε μία ανάλυση 5- Κοντινότερων Γειτόνων (kNN).

Οι περιορισμοί του συγκεκριμένου αλγορίθμου εντοπίζονται στην περίπτωση, ξανά των big data όπως αναφέραμε και στην παλινδρόμηση όπου υπάρχει χρήση μεγάλου αριθμού υπολογιστικών πόρων.



Εικόνα 13: kNN Αλγόριθμος

Η επόμενη τεχνική που θα μελετήσουμε είναι η Naïve Bayes όπου αν και απλή είναι ιδιαίτερα διάσημη, η βάση της οποίας είναι η υπόθεση της ανεξαρτησίας μεταξύ των διάφορων χαρακτηριστικών μέσα στο σύνολο δεδομένων (για αυτό και η ονομασία Naïve – Αφελής.). Αυτό, σημαίνει ότι, (41)σύμφωνα με την υπόθεσή μας, η αλλαγή της τιμής μίας μεταβλητής δεν επηρεάζει τις υπόλοιπες. Λαμβάνουμε υπόψιν για τον υπολογισμό τον γνωστό τύπο του Bayes για την υπο-συνθήκη ή δεσμευμένη πιθανότητα:

$$\Pr [A | B] = \frac{\Pr [A \cap B]}{\Pr [B]}.$$

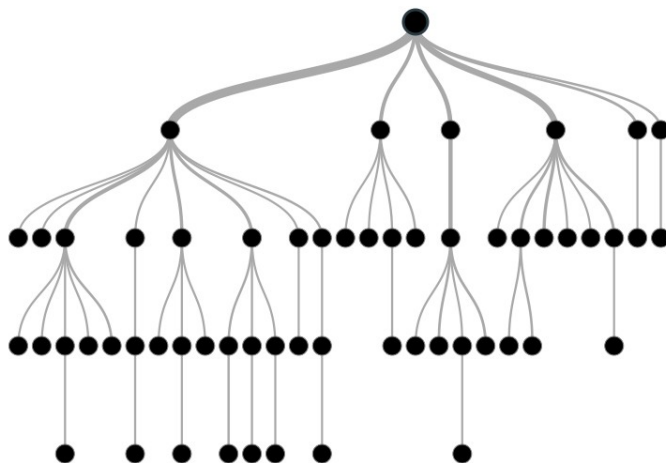
Σε σχέση με την κλασσική στατιστική όμως, και ειδικά στην περίπτωση που έχουμε περισσότερες μεταβλητές στην είσοδο, η υπόθεση της ανεξαρτησίας μας επιτρέπει να προχωρήσουμε στον πολλαπλασιασμό των πιθανοτήτων για κάθε μεταβλητή με την αντίστοιχη πιθανότητα να επιλέξουμε τυχαία την εκάστοτε οντότητα και το αποτέλεσμα αυτό να το διαιρέσουμε με το γινόμενο της συνολικής πιθανότητας να εμφανιστεί το εκάστοτε χαρακτηριστικό.

Έτσι, μπορούμε, να προχωρήσουμε στην πρόβλεψη με βάση τα χαρακτηριστικά, για το πορτοκάλι, αντικαθιστώντας στην εξίσωση του Bayes τον αριθμητή και τον παρονομαστή με τις παραπάνω τιμές:

Στην συγκεκριμένη μέθοδο μπορούμε να προχωρήσουμε σε βελτιώσεις όπως το να υπολογίσουμε την λογαριθμική πιθανότητα αλλά και να υποθέσουμε διαφορετικές διασπορές στα σύνολα δεδομένων μας.

Στη συνέχεια, ο επόμενος, με βάση την σειρά αναφοράς, αλγόριθμος είναι εκείνος που αφορά τις αποφάσεις μέσω δένδρων (δεντρικών αποφάσεων – decision trees). Η ονομασία του προέρχεται κυρίως από τον τρόπο ανάλυσης – αναπαράστασης που έχει την μορφή δένδρου. Το ίδιο, περιλαμβάνει ένα σύνολο κανόνων-εντολών, σύμφωνα με τις οποίες γίνεται προσπάθεια να λυθεί το ζητούμενο πρόβλημα κατηγοριοποίησης/παλινδρόμησης που έχουμε να αντιμετωπίσουμε.

Εδώ, ο στόχος, όπως αναφέραμε, είναι η κατασκευή ενός μοντέλου, το οποίο με την βοήθεια συγκεκριμένων κανόνων θα είναι σε θέση να παρέχει συγκεκριμένη πρόβλεψη για την εκάστοτε υπό μελέτη, μεταβλητή – δείγμα της μεταβλητής.

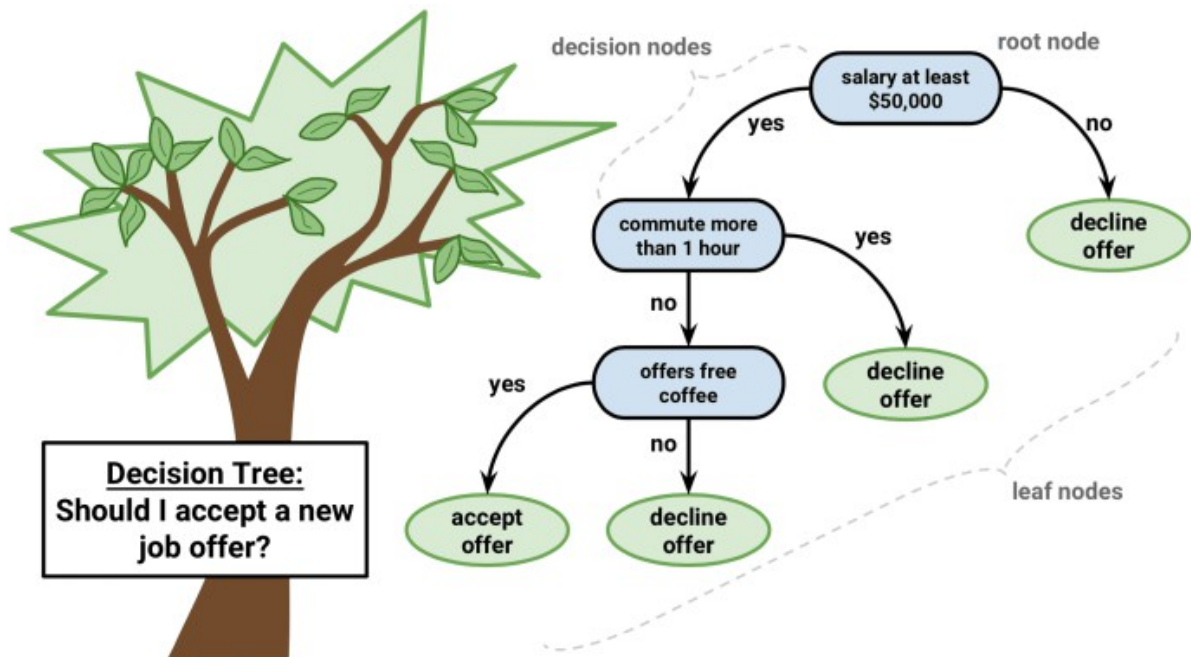


Εικόνα 14: Reversed Tree

Η μορφή, όπως προ είπαμε ομοιάζει αυτή του δένδρου (ανεστραμμένου συγκεκριμένα), όπου ένας κόμβος αντιπροσωπεύει κάποιο χαρακτηριστικό ενώ αντίστοιχα το κάθε φύλλο μία από τις κατηγορίες υπό πρόβλεψη για την εκάστοτε μεταβλητή. Τέλος, υπάρχει η ρίζα, δηλαδή ο αρχικός κόμβος. (42)

Λόγω της δομής του συγκεκριμένου αλγόριθμου, είναι εύκολο με τον ανθρώπινο τρόπο σκέψης να κατανοήσουμε (σε σχέση με πιο σύνθετους αλγόριθμους που θα δούμε στη συνέχεια) πώς οδηγούμαστε σε κάθε περίπτωση, στο να εκτιμήσουμε την κατηγορία για την

κάθε άγνωστη μεταβλητή, καθώς οδηγούμαστε σε κάθε κόμβο/φύλλο με βάση την θετική/αρνητική απάντηση σε ένα ναι ή όχι όπως φαίνεται παρακάτω:



Εικόνα 15: Decision Tree

Ο αλγόριθμος της κατασκευής του δένδρου, ξεκινάει με την εύρεση/επιλογή του βέλτιστου χαρακτηριστικού ως αρχικό κόμβο, μέσω της χρήσης κάποιου κριτηρίου. Έπειτα, προχωράμε σε διαχωρισμό του αρχικού σετ σε μικρότερα υποσύνολα και επαναλαμβάνουμε την διαδικασία μέχρι να ισχύουν μία από τις παρακάτω συνθήκες:

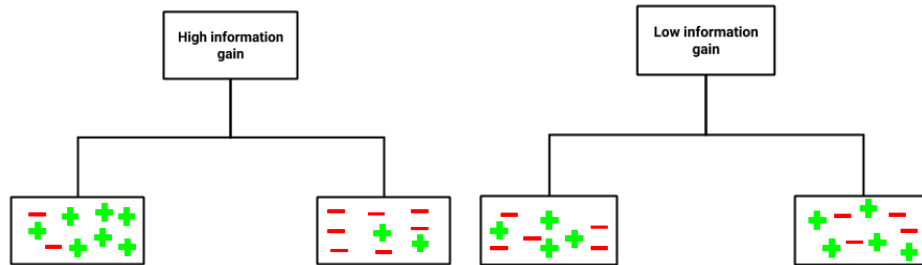
1. Δεν έχουμε άλλα δεδομένα
2. Δεν υπάρχουν άλλα διαθέσιμα attributes

Τέλος, πρέπει να αναφερθούμε στα κριτήρια με τα οποία θα πραγματοποιηθούν τα παραπάνω βήματα. Τα πιο διάσημα, είναι τα παρακάτω:

- Κέρδος Πληροφορίας - Information Gain
- Gini Index

Κέρδος Πληροφορίας (Information Gain) : Εδώ, μας ενδιαφέρει η πληροφορία σχετικά με το πόσο αποτελεσματικά το συγκεκριμένο χαρακτηριστικό είναι σε θέση να προχωρήσει σε διάκριση των διαφορετικών δεδομένων, υπό το πρίσμα της εκτίμησης μίας συγκεκριμένης κατηγορίας/τιμής.

Παρακάτω φαίνονται, σε ποιοτικό βαθμό πώς θα ήταν τα δύο τελείως αντιδιαμετρικά συστήματα.



Εικόνα 16: Information gain

Σημαντική έννοια εδώ είναι και η **Εντροπία Πληροφορίας**, η οποία αποτελεί το μέτρο της αβεβαιότητας μίας τυχαίας μεταβλητής ή αλλιώς, την περιγραφή της τυχειότητας ενός συστήματος. Όταν έχουμε δείγματα από μία μόνο κατηγορία σε σημαντικό ποσοστό, θέλουμε μικρές τιμές πληροφορίες. Σε πλήρως ίδιο σύνολο, έχουμε μηδέν.

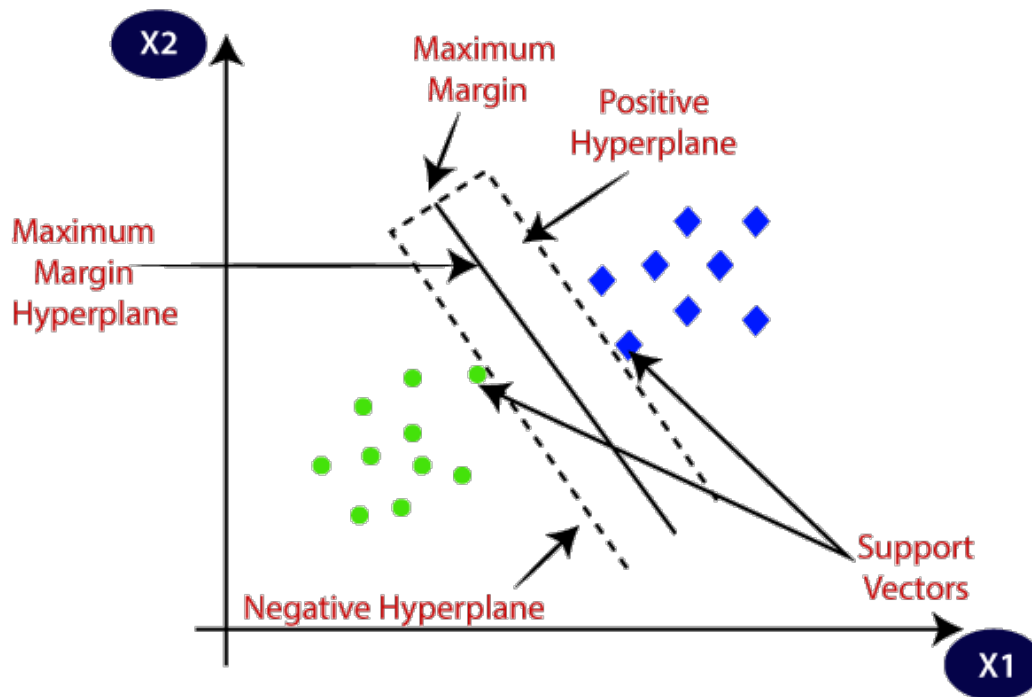
Το δεύτερο κριτήριο που αναφέραμε είναι το **Gini Index**, μία πιθανότητα που εκφράζει την ικανότητα να επιλέξουμε τυχαία αντικείμενα ίδια τάξης από έναν πληθυσμό. Σε περίπτωση όπου λάβουμε αυτό το κριτήριο υπόψιν στον σχηματισμό των δένδρων αποφάσεων, θα πρέπει αρχικά να υπολογίσουμε τους αντίστοιχους δείκτες για τους υπο-κόμβους αθροίζοντας το τετράγωνο της πιθανότητας επιτυχίας/αποτυχίας και στη συνέχεια, συγκρίνοντας τις τιμές ώστε να συμπεράνουμε ποιο χαρακτηριστικό είναι εκείνο με τον μεγαλύτερο δείκτη Gini.

Τελευταίος αλγόριθμος προς μελέτη αποτελούν τα Support Vector Machines, όπου αρκετά συνοπτικά, επιδιώκεται η εύρεση ενός βέλτιστου υπερ-επιπέδου/hyperplane το οποίο θα είναι σε θέση να διακριτοποιήσει, με μέγιστο δυνατό περιθώριο τα δεδομένα του, υπό μελέτη, συνόλου, στην ίδια λογική με τον αλγόριθμο k κοντινότερων γειτόνων

Αρχικά, πρέπει να αναλυθεί η έννοια του **υπερ-επιπέδου**.

Το ίδιο (43) είναι ένα όριο απόφασης του οποίου η γεωμετρική μορφή εξαρτάται από τις διαστάσεις των χαρακτηριστικών του συνόλου δεδομένων. Το ίδιο είναι υπεύθυνο για την κατηγοριοποίηση των διαφορετικών σημείων και για αυτό το λόγο η εύρεση του βέλτιστου αποτελεί το σημαντικότερο ζήτημα σε μία SVM ανάλυση. Συγκεκριμένα, λαμβάνοντας

υπόψιν την θέση και την μορφή του είμαστε σε θέση απευθείας να ορίσουμε την κατηγορία της καινούργιας-άγνωστης μεταβλητής



Εικόνα 17: SVM

Το βασικό πρόβλημα, όπως είναι λογικό εντοπίζεται στο πώς θα μπορέσουμε να ορίσουμε το συγκεκριμένο επίπεδο καθώς και τον τρόπο που θα βρούμε το βέλτιστο δυνατό. Για ένα τυχαίο υπερ επίπεδο, όπως στο παραπάνω σχήμα, το μέτρο της τιμής του περιθωρίου, που προαναφέραμε, ορίζεται ως η απόσταση μεταξύ του υπερ επιπέδου και του κοντινότερου σημείου του συνόλου δεδομένων.

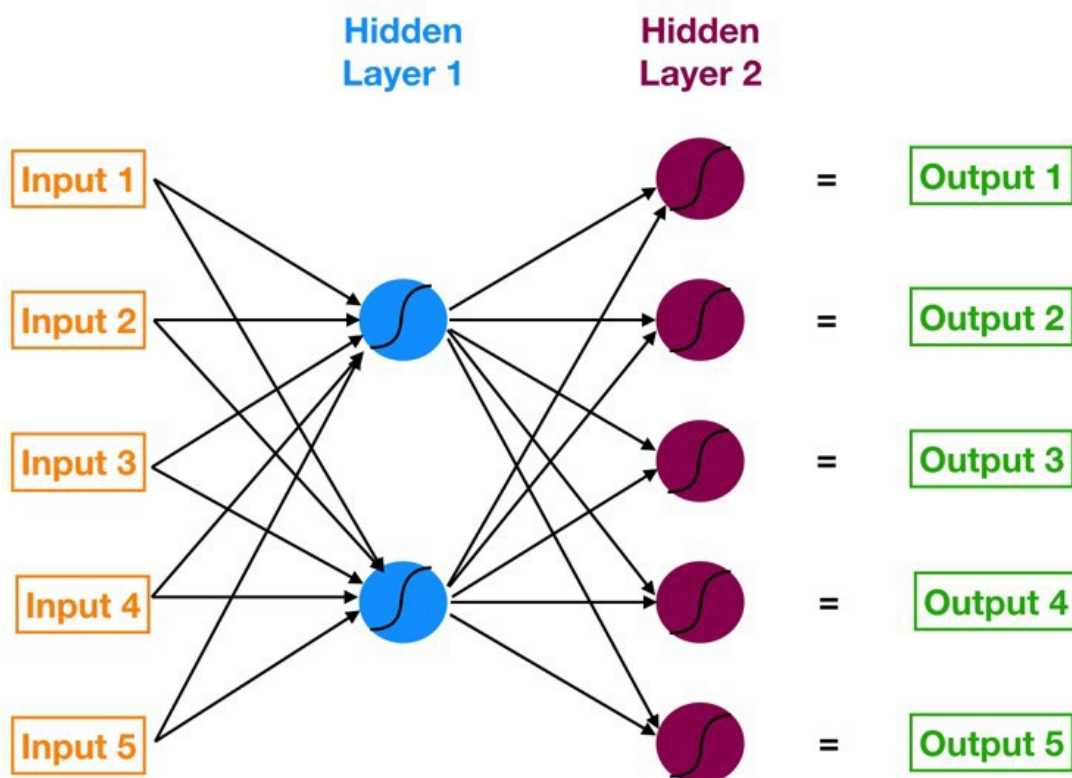
Εκτελώντας διανυσματικές πράξεις μεταξύ των πλησιέστερων σημείων (που ονομάζονται και Διανύσματα Υποστήριξης, και από εκεί βγαίνει και το όνομα) και έχοντας ως οδηγό τις τιμές των αποστάσεων μεταξύ αυτών και του υπερεπιπέδου, είμαστε σε θέση να βρούμε την βέλτιστη δυνατή μορφή που παρέχει το μέγιστο δυνατό περιθώριο. Η χρήση της συγκεκριμένης τεχνικής, είναι δυνατόν να μας παρέχει σημαντική εκτιμητική ικανότητα στις περιπτώσεις όπου διαθέτουμε προβλήματα κατηγοριοποίησης με πολλά χαρακτηριστικά ενώ, ταυτόχρονα παρουσιάζεται αποδοτική αξιοποίηση των υπολογιστών πόρων .

Με την ολοκλήρωση της ανάλυσης ορισμένων βασικών αλγορίθμων, μηχανικής μάθησης, είμαστε σε θέση να ξεκινήσουμε να αναφερόμαστε στην βαθιά μάθηση και στα νευρωνικά δίκτυα.



Στην σημερινή εποχή, κάποιο είδος νευρωνικού δικτύου αξιοποιείται, τουλάχιστον σε έναν βαθμό, σχεδόν καθημερινά από όλους μας. Εργασίες όπως αναγνώριση ομιλίας, μεταφράσεις κειμένου και πολλές άλλες, εκτελούνται σχεδόν αποκλειστικά από αυτές τις δομές.

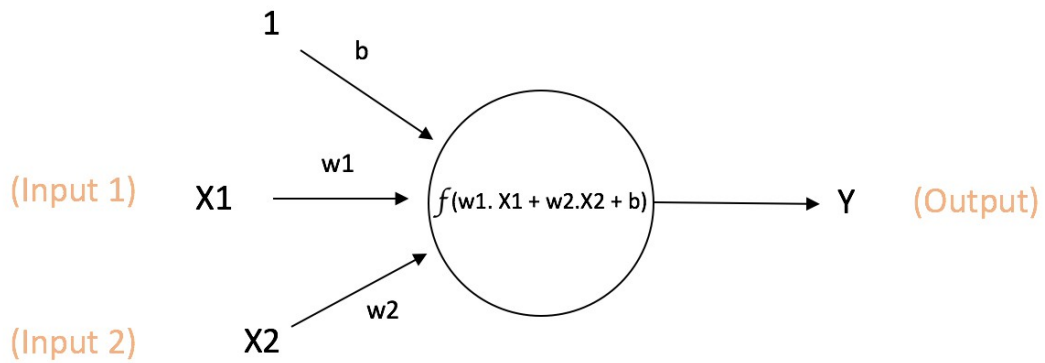
Ένα νευρωνικό δίκτυο (39), συνήθως χαρακτηρίζεται από τρία διαφορετικά επίπεδα, το επίπεδο – layer εισόδου (input layer), το/τα κρυφό/κρυφά (hidden) και τέλος, την έξοδο/ (output layer) όπως φαίνεται και στην παρακάτω εικόνα:



Εικόνα 18: Neural Network

Το σημαντικότερο τους γνώρισμα εντοπίζεται στο ότι, λόγω του τρόπου λειτουργίας και εκπαίδευσής τους, έχουν την δυνατότητα να εξάγουν συμπεράσματα μέσα από τεράστια ποσότητα πολύπλοκων και ασαφών δεδομένων, σε σύγκριση με τους αλγορίθμους μηχανικής μάθησης που είδαμε σε προηγούμενα κεφάλαια ή τον ανθρώπινο παράγοντα.

Σχετικά με τις συναρτήσεις ενεργοποίησης, μία χαρακτηριστική είναι η ReLU που θα αναφερθεί παρακάτω, γεγονός που την κάνει αρκετά σημαντική και αυτή αλλά και το πώς λειτουργεί ένας νευρώνας, στο δίκτυο.

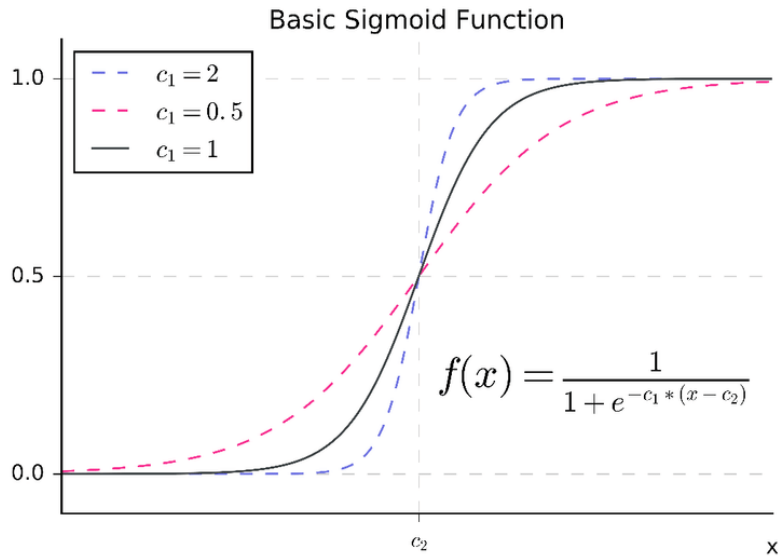


$$\text{Output of neuron} = Y = f(w1.X1 + w2.X2 + b)$$

**Εικόνα 19: Neuron**

Για να μπορέσουμε (44) να κατανοήσουμε την λειτουργία του, πρέπει να γυρίσουμε στον perceptron/shallow neural network που αναφέραμε στην εισαγωγή, όπου ο ίδιος αξιοποιεί μία συνάρτηση με σκοπό να αντιστοιχίσει ένα διάνυσμα δυαδικών τιμών σε μία μόνη. Το τελικό βήμα (μετά τον πολλαπλασιασμό βαρών, τα οποία ρυθμίζονται κατά την διάρκεια της εκπαίδευσης) είναι η χρήση μίας συνάρτησης ενεργοποίησης (Activation Function) η οποία ουσιαστικά (στην περίπτωση για παράδειγμα μίας δυαδικής απόφασης – Ναι/Όχι) μας απαντάει αν η προηγούμενη τιμή είναι μεγαλύτερη ενός threshold, ώστε να μπορέσουμε να δώσουμε την σωστή πρόβλεψη.

Η πιο συνηθισμένη συνάρτηση ενεργοποίησης είναι η σιγμοειδής, την οποία έχουμε δει και στην λογιστική παλινδρόμηση πιο πάνω.



Εικόνα 20: Σιγμοειδή Συνάρτηση

Όπως φαίνεται από τον μαθηματικό τύπο, η υπολογισμένη ποσότητα δίνεται ως είσοδος στην τιμή  $z$  και παράγεται το αντίστοιχο αποτέλεσμα. Παρόλα αυτά λόγω του τύπου, το ίδιο δημιουργεί προβλήματα στον **backpropagation** αλγόριθμο, τον βασικό δηλαδή μηχανισμό επιστροφής των σφαλμάτων στην είσοδο για την διόρθωσή τους και το tweaking των βαρών η οποία αποτελεί μπορεί να αναλυθεί στα παρακάτω βήματα:

- Αρχικοποίηση των τιμών των βαρών
- Εκκίνηση νέας εποχής εκπαίδευσης
- Μεταφορά της πληροφορίας προς τα εμπρός
- Προς τα πίσω διάδοση των λαθών
- Προσαρμογή των βαρών

Από ιστορική σκοπιά, η μέθοδος αυτή έχει προταθεί αρκετά παλιά, αλλά είχε αγνοηθεί σε σημαντικό βαθμό λόγω του μεγάλου υπολογιστικού κόστους, όπως και αντίστοιχα η γενικότερη λύση των αρχιτεκτονικών των νευρωνικών δικτύων, ενώ συνέχεια ‘ανακαλύφθηκε’ ξανά γύρω στα μέσα της δεκαετίας του 80 και ουσιαστικά ξεκίνησε να αποτελεί ένα γενικά αποδεκτό εργαλείο για την εκπαίδευση του πολυεπίπεδο perceptron.

Αρχικά, όλα τα βάρη του δικτύου, αρχικοποιούνται με τυχαίες τιμές, το εύρος των οποίων κατανέμεται ομοιόμορφα και σε μικρή έκταση. Στην περίπτωση όπου οι τιμές αυτές προσεγγίζουν το 0 οι αντίστοιχες κλίσεις που θα υπολογιστούν θα τείνουν και αυτές στο 0,

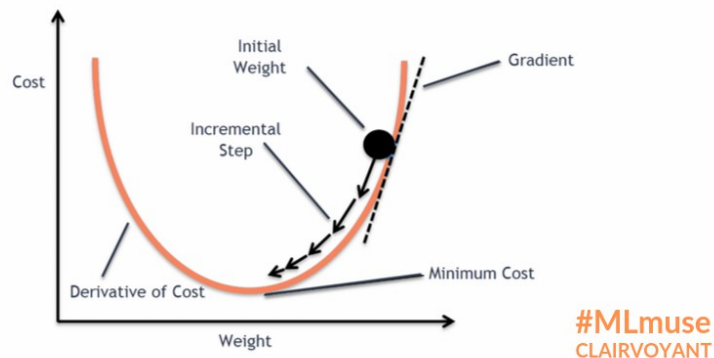
στην περίπτωση όπου έχουμε κρυφά επίπεδα, γεγονός που θα οδηγήσει στην μη εκπαίδευση του δικτύου. Περισσότερο υποδεικνύονται προσπάθειες εκπαίδευσης, με διαφορετικά αρχικά βάρη, για να βρεθεί η καλύτερη τιμή για τη συνάρτηση κόστους (ελάχιστο σφάλμα).

Αντίθετα, εάν οι αρχικές τιμές είναι μεγάλες, τότε αυτές τείνουν να κορεστούν με αποτέλεσμα, η σιγμοειδής συνάρτηση ενεργοποίησης που συνήθως χρησιμοποιείται να είναι πολύ μικρή. Λόγω του γεγονός ότι θα λειτουργήσει ως πολλαπλασιαστικός παράγοντας κατά τη διάρκεια της μαθησιακής διαδικασίας η παραπάνω τιμή, οι κορεσμένες μονάδες θα είναι σχεδόν μπλοκαρισμένες, η οποία καθιστά τη μάθηση πολύ αργή.

Σε μια νέα εποχή εκπαίδευσης, παρουσιάζει όλα τα παραδείγματα του σει προπόνησης ο αλγόριθμος στο νευρωνικό δίκτυο. Στις περισσότερες περιπτώσεις, η εκπαίδευση του δικτύου περιλαμβάνει περισσότερες από μία εποχές. Κατά την εκπαίδευση, τα βάρη ρυθμίζονται μόνο μετά από όλα τα διανύσματα δοκιμής που θα εφαρμοστούν στο δίκτυο. Επομένως, για τις κλίσεις των βαρών, οι ίδιες πρέπει να απομνημονεύονται και να προσαρμόζονται μετά από κάθε μοντέλο στο αντίστοιχο σει προπόνησης.

Σαν αποτέλεσμα τα βάρη θα αλλάξουν μόνο μία φορά (υπάρχει μια online παραλλαγή, πιο απλό, στο οποίο ενημερώνονται τα βάρη άμεσα, σε αυτήν την περίπτωση, τη σειρά με την οποία τα διανύσματα του δικτύου που παρουσιάζονται μπορεί να έχει σημασία). Όλες οι βαθμίδες των βαρών και το τρέχον σφάλμα αρχικοποιούνται με 0. Εδώ είναι σημαντικό να αναφερθούμε στην έννοια της *Gradient Descent*. Η ίδια η έννοια είναι στενός συνδεδεμένη με τον αλγόριθμο *backpropagation* και την γενικότερη ιδέα της βελτιστοποίησης του συστήματος.

Ουσιαστικά αφορά την ελαχιστοποίηση του σφάλματος της παραγώγου που αναφέρθηκε παραπάνω. Για να την κατανοήσουμε καλύτερα, θα πρέπει να γνωρίζουμε και ορισμένες έννοιες όπως *minima* και *maxima* (ολικά/τοπικά μέγιστα – ελάχιστα), συναρτήσεις σύγκλισης, συνθήκη παύσης και τον ρυθμό εκπαίδευσης, τις οποίες εδώ τις θεωρούμε γνωστές, χάριν συντομίας

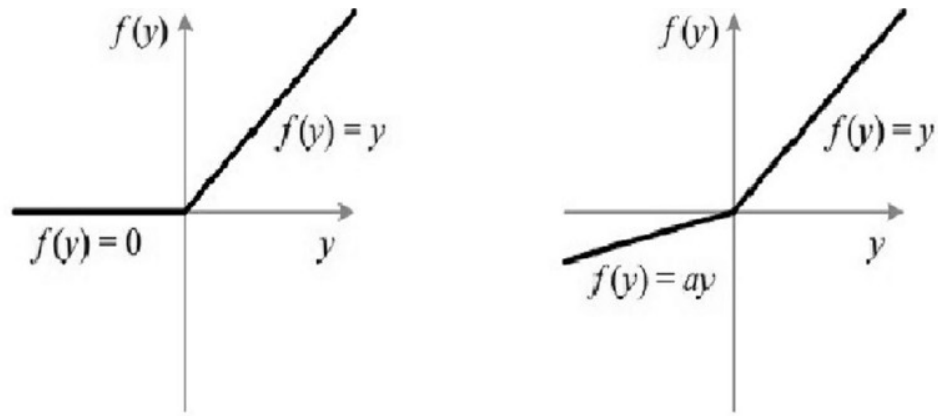


Εικόνα 21: Gradient Descent

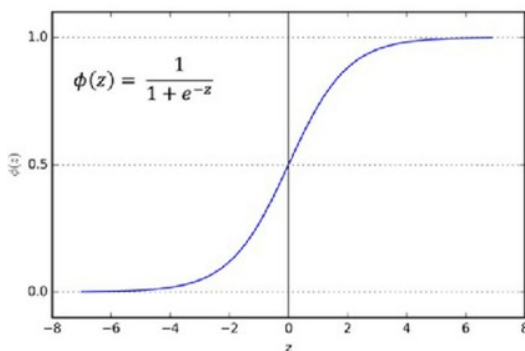
Η Rectified Linear Unity (ReLU) αντίστοιχα, αποτελεί την default συνάρτηση ενεργοποίησης στα βαθιά νευρωνικά δίκτυα.

Αυτό συμβαίνει καθώς η ίδια, σε αντίθεση με την σιγμοειδή δεν υποχρεώνει το δίκτυο να πραγματοποιήσει εκθετικές πράξεις, ενώ δεν υποφέρει από το φαινόμενο της παραγώγου που τείνει να εξαφανιστεί (vanishing gradient), όπως στην σιγμοειδή, λόγω της γραμμικής της φύσης. Αυτό μπορεί να φανεί από το γεγονός ότι η ίδια απλά επιλέγει την μέγιστη τιμή μεταξύ του 0 και της τιμής  $x$  (ή  $z$  όπως την ορίσαμε στην σιγμοειδή). Παρόλα αυτά και η ίδια φέρει κινδύνους, καθώς σε περίπτωση, λόγω της φύσης της, που πολλές τιμές είναι 0, οι νευρώνες απλά δεν θα ενεργοποιηθούν, σε αντίθεση με την σιγμοειδή.

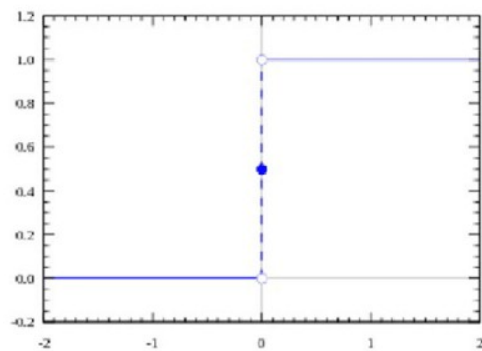
Για την ReLU, υπάρχει μία μικρή τροποποιημένη μορφή της, όπου επιτρέπει αντί να μηδενίζει τις μη θετικές τιμές, να υπάρχουν αρνητικές τιμές, αυξάνοντας το εύρος των αποτελεσμάτων που μπορεί να δώσει η συνάρτηση (Leaky ReLU)



a)



b)



c)

Εικόνα 22: ReLU

Οι εφαρμογές (45) των νευρωνικών δικτύων είναι πολλαπλές, γεγονός που μπορεί να εντοπιστεί σε μία ποικιλία κατηγοριών όπως έχουμε ήδη δει για τους αλγορίθμους μηχανικής μάθησης όπως, χρήση σε προβλήματα παλινδρόμησης (Regression) είτε για την πρόβλεψη σε ζητήματα χρονοσειρών είτε καθαρά για λόγους μοντελοποίησης καθώς και την διευκόλυνση πραγματοποίησης ορισμένων ενεργειών όπως απάντηση σε κλήσεις ή έλεγχος του συστήματος αναπαραγωγής μουσικής με συγκεκριμένες κινήσεις των χεριών. Διαχείριση τεχνολογιών όπως e-books και πλοήγηση σε ιστοσελίδες με αντίστοιχες χειρονομίες.

Σημαντικό κομμάτι αποτελεί και η επεξεργασία δεδομένων μέσω ενεργειών φιλτραρίσματος, διαχωρισμού σήματος και ενεργειών συμπίεσης. Άλλες εφαρμογές είναι τα Συστήματα αναγνώρισης και ελέγχου, όπως πχ στο σύστημα μεταφορών, καθώς βοηθούν στην υλοποίηση αυτο-ελεγχόμενων οχημάτων, αυτά που αφορούν ιατρική βοήθεια τόσο όσον αφορά την πρόβλεψη ασθενειών όσο και κατά την διάγνωση ιατρικών εικόνων όπως πχ ακτινογραφίες, παιχνίδια όπου η λήψη αποφάσεων διαδραματίζει σημαντικό ρόλο (τάβλι, σκάκι), όπου ήδη βαθιά νευρωνικά δίκτυα είναι σε θέση να νικήσουν τους καλύτερους

επαγγελματίες. Εφαρμογές στην οικονομία και στατιστική, όπως εξόρυξη δεδομένων ή ανακάλυψη τυχών trends μεταξύ παραμέτρων, σε φρενήρεις ρυθμούς όπως αυτούς τους χρηματιστηρίου αλλά και τέλος στην ψυχαγωγία όπως αλγόριθμοι στην πλατφόρμα YouTube, συνδυάζουν τέτοιες τεχνολογίες με σκοπό να παρέχουν, σε πραγματικό χρόνο, το καλύτερο δυνατό προτεινόμενο βίντεο με βάση τον εκάστοτε χρήστη.

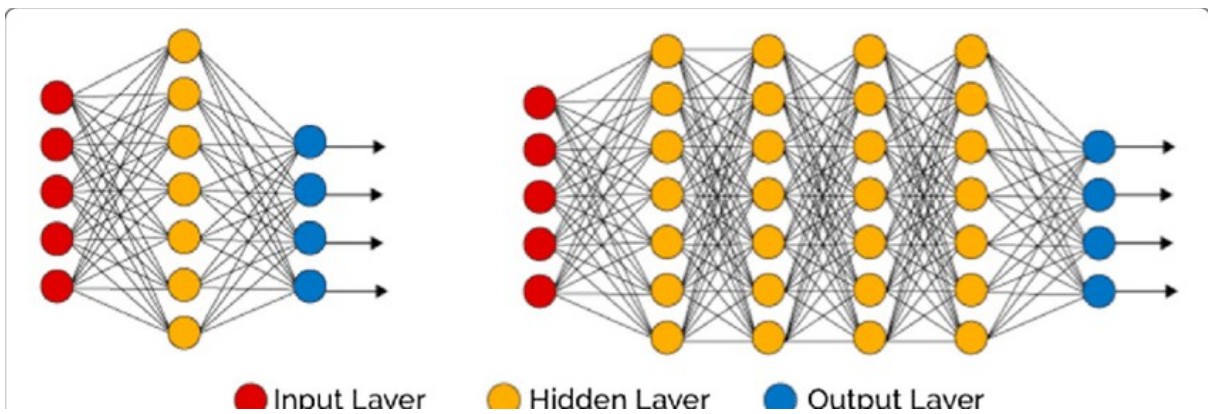
Η χρήση τους αυτή μπορεί να δικαιολογηθεί σε ένα σημαντικό βαθμό από τα πλεονεκτήματά τους σχετικά με τους αλγορίθμους μηχανικής μάθησης, συγκεκριμένα:

- Λόγω της ικανότητας τους για προσαρμοσμένη μάθηση, είναι δυνατόν να μάθουν να εκτελούν καθήκοντα μόνο παρέχοντας εισόδους για δεδομένα προς εκπαίδευση, είτε μέσω, στο ίδιο πλαίσιο, της αρχικής εμπειρίας από τον ανθρώπινο παράγοντα.
- *Αυτο-οργάνωση*: Όπως προ είπαμε, ένα νευρωνικό δίκτυο μπορεί να δημιουργήσει τον δικό του τρόπο οργάνωσης ανεξαρτήτως της πολυπλοκότητας της εισόδου που του παρέχουμε και να παρέχει, κάποια εκπροσώπηση των πληροφοριών κατά την διάρκεια της εκπαίδευσης.
- Υπό τις κατάλληλες συνθήκες, είναι δυνατόν οι λειτουργίες του νευρωνικού δικτύου να γίνονται σε πραγματικό χρόνο, παρέχοντας έτσι ένα πολύ σημαντικό εργαλείο για μία πληθώρα εφαρμογών.
- Το βασικό χαρακτηριστικό αυτών των συστημάτων είναι ότι η εκπαίδευση δεν περιλαμβάνει διαδικασίες προγραμματισμού. Αυτό σημαίνει, ότι ακολουθώντας έναν αλγόριθμο βελτιστοποίησης και αλλάζοντας ορισμένες παραμέτρους, ο δημιουργός του συστήματος είναι σε θέση να παρατηρεί αλλαγές χωρίς να εμπλέκεται στο να προγραμματίσει κάποιο βήμα εξαρχής.
- Συνδυάζοντας τα παραπάνω, καταλαβαίνουμε ότι τα ίδια τα δίκτυα παρουσιάζουν, εφόσον χρειαστεί, μεγάλη ευελιξία στις πραγματικές συνθήκες ως προς την αλλαγή των παραμέτρων που τα εκπαιδεύουν, γεγονός που τα καθιστά κατάλληλα για την συνεχώς μεταβαλλόμενη πραγματικότητά μας.
- Σε περιπτώσεις όπου παραδοσιακές μέθοδοι όπως, ο προγραμματισμός και η χρήση της στατιστικής, αποτυγχάνουν να παρέχουν αποτελέσματα ή να εξάγουν συμπεράσματα, η φύση των παραπάνω μοντέλων είναι σε θέση να παρέχει, πιθανές λύσεις.

- Τέλος, σε συνδυασμό με το παραπάνω, αρκετές φορές οι λύσεις που παρέχονται από τις συγκεκριμένες διατάξεις, υπό κατάλληλες προϋποθέσεις φυσικά, είναι σε θέση να ξεπεράσουν, σε εκτιμητική ικανότητα τις παραδοσιακές μεθόδους.

Για την βαθιά μηχανική μάθηση, η ίδια προσφέρει ποικίλους τρόπους στην εκπαίδευση μοντέλων και την ταξινόμηση δεδομένων. Έχει όμως και σημαντικές διαφορές που την κάνουν να ξεχωρίζει από τις προηγούμενες. Η μεταφορά της αναγνώρισης και της εξοικείωσης του ανθρώπινου περιβάλλοντος σε μηχανικό σύστημα είναι ο πρωτεύον στόχος και των δύο αυτών πεδίων.

Η εκμάθηση και η αναγνώριση παραδειγμάτων σε άριστο ποσοστό επιτυχίας είναι το ζητούμενο που ερευνάται στα δύο πεδία αλλά ο τρόπος που διεξάγεται στο καθένα είναι ξεχωριστός.



Εικόνα 23: Deep Neural Network

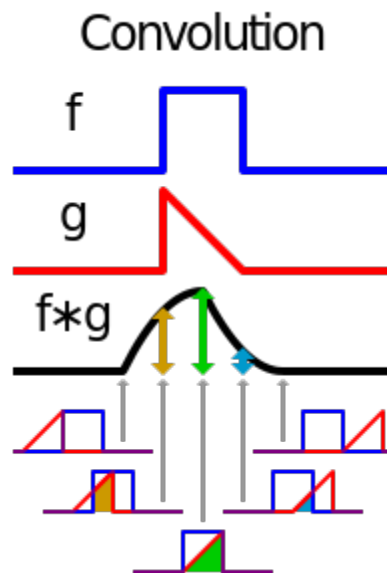
Χρησιμοποιώντας έναν υπολογιστή για ταξινόμηση εικόνων με την προσέγγιση της μηχανικής μάθησης, η επιλογή των βασικών χαρακτηριστικών, όπως των άκρων και των γωνιών του αντικειμένου που υπάρχει στην εικόνα, γίνεται με τη βοήθεια ανθρώπινου παράγοντα. Αυτό αποσκοπεί στην εκμάθηση του μοντέλου με βάση τη μηχανική μάθηση. Έπειτα, το συγκεκριμένο μοντέλο κάνει μια αναφορά αναλύοντας και ταξινομώντας τα νέα αντικείμενα. Αυτή η τεχνική γίνεται επίσης για την αναγνώριση σκηνής ή τοπίου και την ανίχνευση αντικειμένων. Η λύση προβλημάτων με τη μέθοδο της μηχανικής μάθησης έχει συγκεκριμένη ροή ενεργειών. Παρατηρώντας την εικόνα, γίνεται εξαγωγή των βασικών στοιχείων που τη διέπουν και στη συνέχεια δημιουργείται ένα μοντέλο το οποίο περιγράφει ή υποθέτει το αντικείμενο.

Αντιθέτως, στη βαθιά μάθηση παραλείπεται ο ανθρώπινος παράγοντας ώστε να γίνει η εξαγωγή των χαρακτηριστικών των εικόνων. Αντ' αυτού, γίνεται άμεση εισαγωγή της



εικόνας στον αλγόριθμο της βαθιάς μάθησης που αποτελείται το σύστημα, και τότε γίνεται η πρόβλεψη του αντικειμένου. Έτσι, η βαθιά μάθηση κηρύσσεται ως ένας υπό τύπος της μηχανικής μάθησης, διότι επεξεργάζεται απευθείας τις εικόνες και συνήθως είναι πιο περίπλοκη από τη μηχανική μάθηση, λόγω των χαρακτηριστικών που την απαρτίζουν.

Εδώ, πρέπει να αναφερθούμε και στη συνέλιξη σαν διαδικασία, η οποία συνέλιξη σαν μαθηματική πράξη, είναι αρκετά σημαντική και εντοπίζεται σε μία πληθώρα κατηγοριών όπως πχ η επεξεργασία σήματος, όπου κατέχει ιδιαίτερα σημαντική θέση. Μπορεί να εκφραστεί ως το ολοκλήρωμα δύο συναρτήσεων όταν η μία αντιστρέφεται και της εφαρμόζουμε ένα shift στον χρόνο ενώ την κάνουμε slide στην άλλη. Πρακτικά, δηλώνει πώς η μία συνάρτηση, αλλάζει το σχήμα της άλλης



Εικόνα 24: Συνέλιξη

Αντίστοιχα με την επεξεργασία σήματος, αυτό το slide στην περίπτωση των νευρωνικών δικτύων αφορά το sliding συγκεκριμένων φίλτρων στις εικόνες εισόδου με σκοπό να αναγνωρίσουμε συγκεκριμένα χαρακτηριστικά τους.

Η βαθιά μάθηση απαιτεί **μεγάλο όγκο επεξεργασίας** δεδομένων όπως επίσης και ονοματοποιημένα αρχεία σε αντίθεση με την μηχανική μάθηση. Αυτό συμβαίνει γιατί χρησιμοποιώντας τα μοντέλα της βαθιάς μάθησης, τα οποία είναι πιο περίπλοκα, πρέπει να υπάρχει ένας ογκώδης χώρος αποθήκευσης χιλιάδων εικόνων έτσι ώστε να είναι επιτυχημένα και αληθή τα αποτελέσματα. Χρησιμοποιώντας λοιπόν μεγάλο όγκο δεδομένων πρέπει να υπάρχει ο κατάλληλος εξοπλισμός επεξεργασίας γραφικών. Κάρτες γραφικών τελευταίας

τεχνολογίας είναι απαραίτητες, αφού μειώνουν αισθητά το χρόνο επεξεργασίας αυτών των δεδομένων.

Η μέθοδος επίλυσης των προβλημάτων με τη βοήθεια της μηχανικής μάθησης όμως δίνει την επιλογή εκμάθησης του μοντέλου σε περισσότερους ταξινομητές. Γίνεται επίσης γνωστή η εξαγωγή των συγκεκριμένων χαρακτηριστικών που θα χρησιμοποιηθούν για το καλύτερο αποτέλεσμα. Επιπλέον με τη μηχανική όραση δίνεται η δυνατότητα επιλογής διαφόρων συνδυασμών των προσεγγίσεων που μπορούν να χρησιμοποιηθούν. Διαφορετικοί ταξινομητές και χαρακτηριστικά μπορούν επίσης να χρησιμοποιηθούν για να δουλέψουν συνδυαστικά με τα δεδομένα.

Συνοψίζοντας λοιπόν, με τη μηχανική μάθηση χρειάζονται λιγότερα δεδομένα απ' ό τι με τη βαθιά μάθηση και η εκμάθηση του μοντέλου γίνεται γρηγορότερα. Παράλληλα, με τη βαθιά μάθηση, που γίνεται όλο και πιο δημοφιλής, τα αποτελέσματα είναι ακριβέστερα και δε χρειάζεται να υπάρξει ανθρώπινη παρεμβολή ώστε να παρθούν αποφάσεις για τα κατάλληλα χαρακτηριστικά εκπροσώπησης του αντικειμένου, διότι μαθαίνονται αυτόματα από το μοντέλο. Έτσι, η ευθύνη της επιλογής των ταξινομητών είναι ελάχιστη σε σύγκριση με αυτής στη μηχανική μάθηση. Στη βαθιά μάθηση χρειαζόμαστε μεγαλύτερη ποσότητα δεδομένων, το οποίο σημαίνει περισσότερος χρόνος εκμάθησης του μοντέλου, όπως και καλύτερη ποιότητα υλικού. Γενικότερα, η επιλογή μεταξύ μηχανικής και βαθιάς μάθησης εξαρτάται από τα δεδομένα που υπάρχουν στην κατοχή του χρήστη, αλλά και το πρόβλημα το οποίο τίθεται να λύσει.

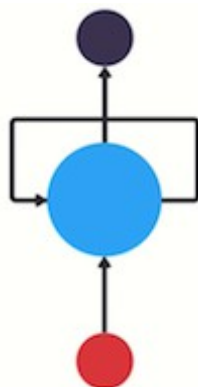
Στη συνέχεια, θα προχωρήσουμε στην σύντομη ανάλυση των παρακάτω αρχιτεκτονικών των νευρωνικών δικτύων:

- Συνελκτικά Νευρωνικά Δίκτυα ( Convolutional Neural Networks – CNNs)
- Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNNs)

Στην δική μας περίπτωση, μας ενδιαφέρει σημαντικά η RNN κατηγορία.

Η πρώτη έννοια που πρέπει να γίνει κατανοητή είναι αυτή της ακολουθίας (sequence). Σύμφωνα με αυτήν, ακολουθία ενός αντικειμένου ή φαινομένου αποτελεί μία συγκεκριμένη σειρά κατά την οποία μίας κατάσταση A προηγείται μίας B. Παραδείγματα ακολουθιών είναι ο ήχος όπως πχ ένα μουσικό τραγούδι ή μία ηχογράφιση αλλά και, όπως είναι λογικό για να αναφέρεται και εδώ, το κείμενο.

Σε αντίθεση με τα Feed Forward νευρωνικά δίκτυα που είδαμε πιο πριν, όπου δηλαδή η πληροφορία κινείται μόνο προς μία κατεύθυνση εδώ υπάρχει η έννοια της μνήμης, της ικανότητάς, με άλλα λόγια, μίας προηγούμενης πληροφορίας να επηρεάζει την επόμενη. Αυτό στα RNNs γίνεται με την εισαγωγή ενός βρόγχου στα κρυφά επίπεδα όπως φαίνεται απλοϊκά παρακάτω:



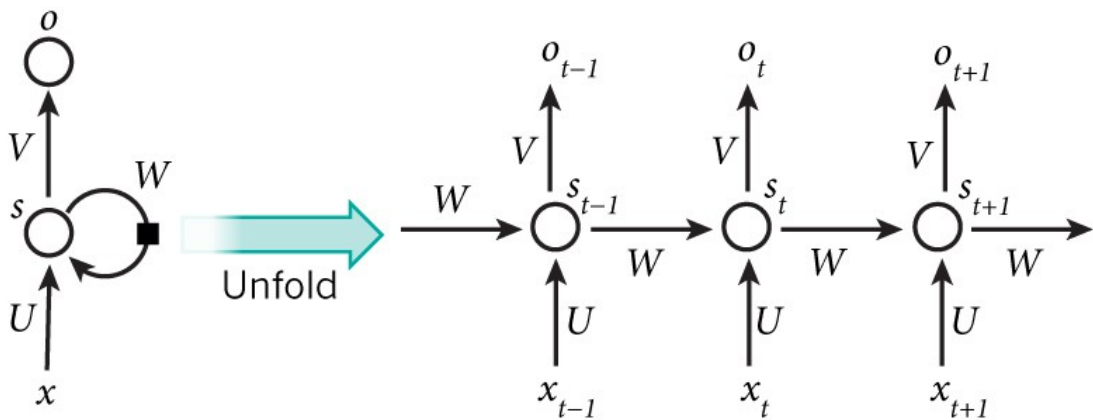
Εικόνα 25: Αναπαράσταση RNN

Το πρώτο σημαντικό ζήτημα που πρέπει να αναφερθεί σχετίζεται με τον λόγο ύπαρξης της συγκεκριμένης κατηγορίας των νευρωνικών δικτύων. Συγκεκριμένα ποιος ήταν ο περιορισμός των συνελκτικών νευρωνικών δικτύων, που αναφέρθηκαν προηγουμένως, που οδήγησε στην κατασκευή μίας εντελώς καινούργιας κατηγορίας;

Το βασικό γνώρισμα των συνελκτικών δικτύων είναι ότι έχουν ανάγκη από εισόδους fixed, δηλαδή συγκεκριμένων διαστάσεων, με σκοπό να παράγουν αντίστοιχα, συγκεκριμένων διαστάσεων εξόδους. Παρόλα αυτά, όπως είναι εύκολο να φανταστεί κανείς, εκτός από τον τομέα επεξεργασίας εικόνας, υπάρχουν τομείς όπως η μετάφραση ή γενικότερα η επεξεργασία κειμένου, όπου η είσοδος δεν είναι σχεδόν ποτέ από πριν γνωστή.

Εδώ, λοιπόν έρχονται τα αναδρομικά νευρωνικά δίκτυα, τα οποία επιτρέπουν το μήκος της εισόδου (αλλά και της εξόδου) να είναι μεταβλητό. Για να πραγματοποιηθεί αυτό, περιλαμβάνουν την ‘ανάμειξη’ στην είσοδο το αποτέλεσμα προηγούμενων εξόδων, εισάγοντας την έννοια της μνήμης στο δίκτυο. Αυτό μπορούμε να το κατανοήσουμε, πώς είναι αρκετά σημαντικό, όταν για παράδειγμα θέλουμε να αναλύσουμε μία πρόταση, καθώς η μία λέξη με την άλλη, υπάρχει μεγάλη πιθανότητα, να έχουν στενή σημασιολογική σύνδεση, σε αντίθεση με την περίπτωση ενός συνελκτικού νευρωνικού δικτύου όπου η εικόνας προς επεξεργασία, δεν έχει σχέση, υπό την έννοια της τελικής πρόβλεψης, με προηγούμενες εισόδους που έχουν υπάρξει στο δίκτυο.

Σχηματικά, η διαδικασία που ακολουθείται σε ένα κόμβο του αναδρομικού δικτύου είναι η παρακάτω:



Εικόνα 26: RNN

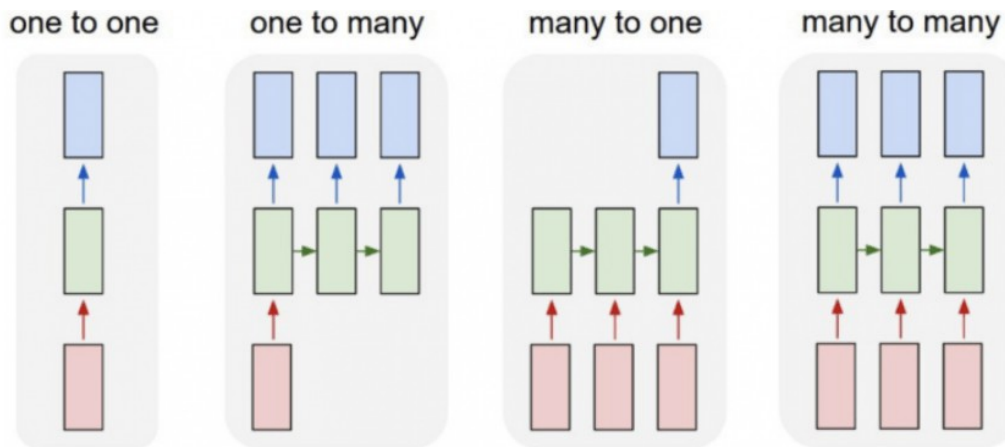
Όπου έστω  $x_t$  είναι η ‘τωρινή’ τιμή που μελετάμε και  $O_t$  το αντίστοιχο της output. Η έννοια της μνήμης εδώ εντοπίζεται στο  $S_{t-1}$  το οποίο, ουσιαστικά, αναπαριστά την τιμή του βάρους που έχει δοθεί από την προηγούμενη είσοδο και το οποίο διατηρείται στα υπόλοιπα όλων των επίπεδων, σε αντίθεση με μία κλασσική αρχιτεκτονική όπου η παράμετρος αυτή αλλάζει ανά επίπεδο

Βέβαια, λόγω των διαφορετικών τιμών που μπορεί να λάβει η τιμή της εισόδου, υπάρχουν και οι αντίστοιχες αρχιτεκτονικές για αυτό που λέγεται sequence learning (εκμάθηση σειράς – αλληλουχίας). Αυτές είναι:

- Μία προς μία (One to One): Εδώ, είναι ευνόητο ότι αυτή η δομή ομοιάζει σημαντικά των δικτύων εμπρόσθιας ανατροφοδότησης και θα μπορούσε να χρησιμοποιηθεί για τις εφαρμογές που αναφέραμε ήδη για τα συνελκτικά δίκτυα, όπως επεξεργασία εικόνας.
- Πολλές προς μίας (Many to One): Εδώ, δεχόμαστε ένα σύνολο εισόδων, ‘πχ μία αλληλουχία λέξεων, και πρέπει να παράγουμε μόνο μία έξοδο, πχ αν ένα σχόλιο είναι θετικό ή όχι, ψευδές ή αληθινό (pseudo news detection) κτλ.
- Μία προς πολλές (One to Many): Στην συγκεκριμένη εφαρμογή, που μπορεί να βρει πεδίο στην παραγωγή νέων δεδομένων από προηγούμενη, αξιοποιούμε την έξοδο της μίας εισόδου που παρέχεται στο σύστημα ώστε να μπορέσουμε να δημιουργήσουμε πολλαπλές εξόδους

- Πολλές προς πολλές (Many to Many): Σε συνδυασμό με ότι είδαμε παραπάνω, εδώ επιδιώκουμε μέσα από τις εισόδους να δημιουργήσουμε νέα εξόδους, οι οποίες όμως να σχετίζονται μεταξύ τους

Παρακάτω, μπορεί κάποιος να δει, την βασική δομή για την κάθε κατηγορία RNN:



Εικόνα 27: Διαφορετικοί τύποι των RNNs

Πριν προχωρήσουμε στην τελευταία κατηγορία, είναι σημαντικό να τονίσουμε ότι εκτός από την επεξεργασία γλώσσας, τα RNNs συμμετέχουν ενεργά και στην παραγωγή νέων εισόδων.

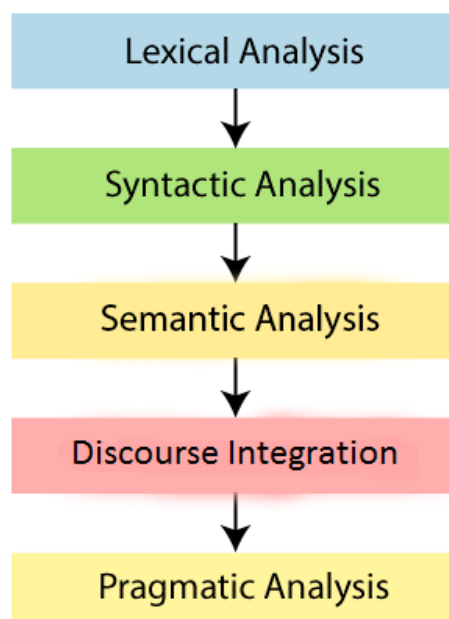
Αυτές, πάνω – κάτω, παραμένουν, στην ουσία τους όμοιες σε όλες τις NLP διαδικασίες, με μικρές αλλαγές/προσαρμογές με βάση το εκάστοτε μοντέλο μηχανικής μάθησης που αξιοποιείται αλλά και την φύση του προβλήματος που καλείται ο μηχανικός να λύσει.

Όπως φαίνεται και στην παρακάτω εικόνα, ονομαστικά τα βήματα που ακολουθούνται είναι τα εξής:

1. Λεξική/Μορφολογική Ανάλυση (Lexical Analysis): Εδώ, γίνεται προσπάθεια ‘σπασίματος’ της λέξης στις ρίζες της – τις μικρότερες δυνατές υπό λέξεις. Η διαδικασία αυτή επεκτείνεται σε μεγαλύτερα λεξικά σύνολα με την διάσπασή τους σε προτάσεις κτλ.
2. Συντακτική Ανάλυση (Syntactic Analysis): Σε αυτό το στάδιο πραγματοποιείται έλεγχος της πρότασης για να διαπιστωθεί αν είναι ορθώς διατυπωμένη με σκοπό να την διασπάσει σε μία δομή η οποία θα περιλαμβάνει την συντακτική σχέση μεταξύ των διαφορετικών λέξεων μέσα στην πρόταση.

3. Σημασιολογική Ανάλυση (Semantic Analysis): Κατά την διάρκεια αυτής της ανάλυσης, πραγματοποιείται αναζήτηση νοήματος στην πρόταση που δίνεται ως είσοδος.
4. Ενσωμάτωση Συζήτησης (ελεύθερη μετάφραση)/Discourse Integration - Πραγματιστική Ανάλυση (Pragmatic Analysis): Εδώ, τοποθετούνται τα αντικείμενα/πρόσωπα στις προτάσεις, με βάση την σημασιολογική ανάλυση που πραγματοποιήθηκε στο προηγούμενο στάδιο.

Είναι βέβαια κατανοητό, ότι τα παραπάνω στάδια παρουσιάζουν με μη συγκεκριμένο τρόπο ένα σύνολο διεργασιών που λαμβάνουν χώρα κατά την Επεξεργασία Φυσικής Γλώσσας.



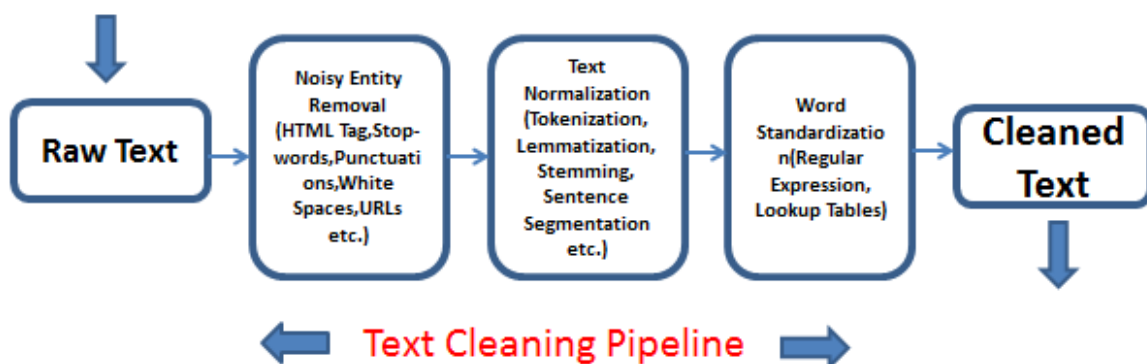
Εικόνα 28: Τα βασικά στάδια μίας NLP διαδικασίας

Για να μπορέσουμε, όπως να εξηγήσουμε καλύτερα την τεχνική που ακολούθησαν οι ερευνητές θα πρέπει να αναλύσουμε, έστω και συνοπτικά τις βασικές NLP λειτουργίες:

1. Tokenization (Διαχωρισμός Συμβόλων): Με τον όρο token, εννοούμε μία μοναδική οντότητα που λειτουργεί ως θεμέλιο για την κατασκευή προτάσεων ή παραγράφων. Η ίδια αποτελεί την μικρότερη ‘κατανοητή’ μονάδα από τον υπολογιστή. Το tokenization μπορεί να υλοποιηθεί, όπως εννοήθηκε παραπάνω, είτε για να εξάγουμε το σύνολο των λέξεων από μία παράγραφο ή/και το σύνολο των προτάσεων (με το να εντοπίζουμε τις τελείες).
2. Stopwords Removal (Διαγραφή των Stopwords). Η έννοια Stopword στις NLP διαδικασίες αντιμετωπίζεται ως θόρυβος και αφορά λέξεις όπως is, am, are,the,to,and

ή τα a/an, υπό το σκεπτικό ότι δεν προσθέτουν κάποιο νόημα στην πρόταση ενώ ταυτόχρονα σπαταλούν πόρους. Ο έλεγχος τους, μπορεί να γίνει εύκολα εφόσον ορίσουμε το σύνολο των λέξεων που θεωρούμε ως stop words, καθώς δεν υπάρχει κάποιο κοινό συμφωνημένο σετ.

3. **Stemming**: Εδώ, πραγματοποιείται η αφαίρεση των affixes (των λεξικολογικών πρόσθετων στην ρίζα της λέξης) είτε στην αρχή είτε στο τέλος των λέξεων. Βέβαια εδώ πρέπει να σημειωθεί ότι τέτοιες οντότητες μπορούν να δημιουργούν νέες λέξεις και οπότε η περικοπή τους θα δημιουργούσε πρόβλημα στο σετ δεδομένων. Για τον λόγο αυτό, πρέπει και εδώ να γίνεται προσεκτική ομαδοποίηση των λέξεων προς stemming.
4. **Lemmatization**: Σε αυτή την λειτουργία, τα tokens τροποποιούνται στην βασική τους μορφή καθώς και διαφορετικές μορφές της ίδιας λέξης (πχ παρελθοντικοί χρόνοι) ταυτίζονται σε μία. Παρόλο το γεγονός ότι ο σκοπός της ομοιάζει αυτού του Stemming, εδώ αξιοποιούμε διαφορετικά εργαλεία. Συγκεκριμένα, στο Lemmatization επιδιώκεται να βρεθεί η οντότητα όπως ορίζεται σε λεξικό, γεγονός που καθιστά απαραίτητη την ανάγνωση/αποθήκευση λεξικών στον αλγόριθμό. Πρέπει βέβαια να σημειωθεί ότι λόγω του ότι η συγκεκριμένη διαδικασία επιλύει και προβλήματα που περιέχουν διφορούμενης σημασίας νοήματα, είναι αρκετά πιο απαιτητική από θέμα υπολογιστικών πόρων σε σχέση με το Stemming.



Εικόνα 29: Βασικό Pipeline Επεξεργασίας

Στη συνέχεια, είναι σημαντικό να υπογραμμίσουμε ορισμένες βασικές αρχιτεκτονικές μηχανικές μάθησης που αξιοποιούνται για NLP διαδικασίες, τα Μοντέλα Γλώσσας – Language Models.

Ένα μοντέλο γλώσσας – Language model ουσιαστικά εκφράζει την κατανομή της πιθανότητας σε μία λέξη ή ένα σύνολο λέξεων, προσπαθώντας να εκφράσει την εγκυρότητα ή μη μίας ακολουθίας προτάσεων, υπό το πλαίσιο του πώς πρέπει να εκφραστεί και όχι από γραμματικής άποψης. Οι βασικοί τύποι των μοντέλων γλώσσας είναι οι παρακάτω: Πιθανοτικοί (**probabilistic**) μέθοδοι και language models που είναι βασισμένα σε αρχιτεκτονικές νευρωνικών δικτύων (**neural network based**):

### Πιθανοτικά Μοντέλα

Για τα πιθανοτικά μοντέλα, η κατασκευή τους στηρίζεται στον υπολογισμό των n-gram πιθανοτήτων (όπου n-gram συμβολίζει μία ακολουθία n λέξεων με n ακέραιο θετικό αριθμό), αξιοποιώντας, στην ουσία, την υπόθεση του Markov – Κρυφή Αλυσίδα του Markov (hidden Markov chain). Περιγραφικά, θα μπορούσαμε να πούμε ότι η ιδιότητα του Markov αναφέρει ότι:

*Έστω ότι έχουμε μία τυχαία διαδικασία (διαδικασία που αφορά τυχαίες μεταβλητές) και μία κατάσταση που συμβαίνει σε ένα δεδομένο χρόνο  $t$ . Η πρόβλεψη για μία μελλοντική κατάσταση δεν εξαρτάται από παρελθοντικές καταστάσεις (απουσία μνήμης).*

Υπό αυτό το πλαίσιο, για το δικό μας πρόβλημα, καταλαβαίνουμε ότι έχοντας ως δεδομένο το n-1 gram που αποτελεί την τωρινή κατάσταση, οι n πιθανότητες δεν εξαρτώνται από τα n-2 , n-3 που αποτελούν, αντίστοιχα τα παρελθοντικά grams/καταστάσεις σύμφωνα και με την Markov ιδιότητα.

Η συγκεκριμένη τεχνική, ενώ στην ουσία της είναι απλή, παρουσιάζει σοβαρά προβλήματα κυρίως λόγω της φύσης του λεξικολογικού περιεχομένου των προτάσεων. Αναλυτικά, αρκετές φορές το πραγματικό νόημα μίας πρότασης καθώς και η επιλογή λέξεων μπορούν και επηρεάζονται από την χρήση των προηγούμενων (πρόβλημα περιεχομένου – context). Επιπλέον, εδώ είναι σημαντικό να σημειωθεί η πιθανότητα σημαντικού υπολογιστικού μεγέθους των μελλοντικών καταστάσεων οι οποίες θα πρέπει να υπολογιστούν. Για τον λόγο αυτό, περισσότερο κυρίαρχα σε αυτή την κατηγορία είναι τα γλωσσικά μοντέλα που σχετίζονται με την εκμετάλλευση των αρχιτεκτονικών των νευρωνικών δικτύων (neural network based).



## Μοντέλα βασισμένα σε Νευρωνικά Δίκτυα

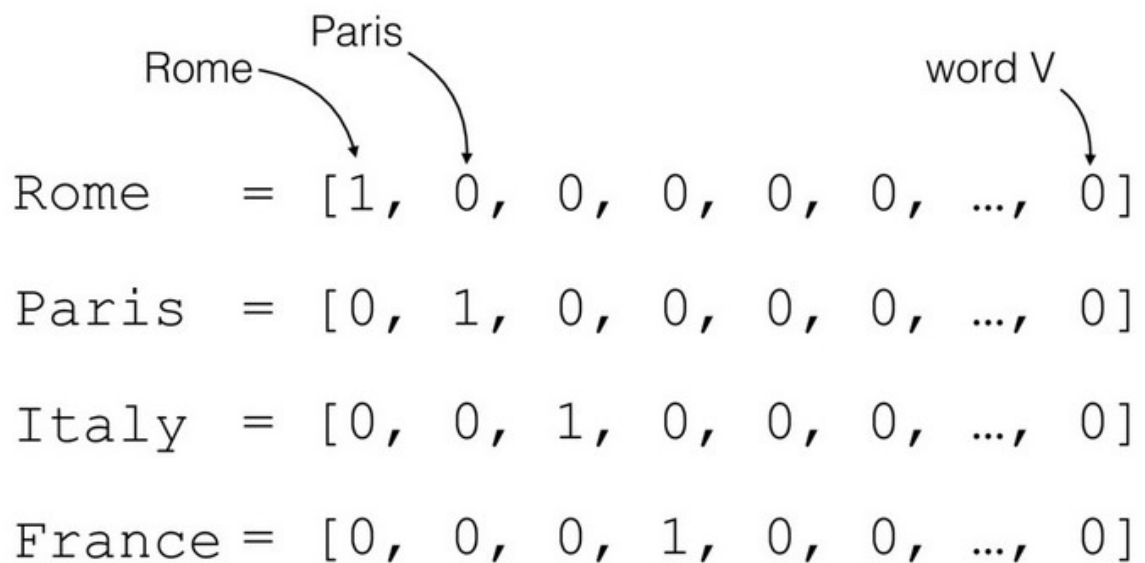
Τα γλωσσικά μοντέλα που στηρίζονται σε αρχιτεκτονικές νευρωνικών δικτύων, αρχικά αντιμετωπίζουν -κωδικοποιούν με διαφορετικό τρόπο την είσοδο. Η διαδικασία της κωδικοποίησης (encoding) είναι ιδιαίτερα γνωστή πρακτική προ επεξεργασίας στο κομμάτι της τεχνητής νοημοσύνης καθώς επιτρέπει τον μετασχηματισμό λεκτικών μεταβλητών σε μορφές που μπορούν να γίνουν κατανοητές και να τεθούν υπό επεξεργασία από τα νευρωνικά δίκτυα όπως για παράδειγμα το one – hot encoding, όπου μπορούμε να πάρουμε μία μεταβλητή με γνωστές τιμές όπως πχ το φύλο (Αρσενικό – Θηλυκό) και να την μετασχηματίσουμε σε μορφή που κατανοεί η αρχιτεκτονική μας (0 ή 1) διατηρώντας όμως την ουσία του διαχωρισμού, με την έννοια ότι δεν θα προχωρήσουμε σε αριθμητικές συγκρίσεις ή πράξεις μεταξύ τους (αριθμητική τιμή του 1 μεγαλύτερη του 0 άρα 1 καλύτερο του μηδέν κτλ).

Εδώ, λοιπόν επιδιώκουμε όπως και στις αντίστοιχες περιπτώσεις του φύλου ή άλλες μεταβλητές όπως του καιρού ή του συναισθήματος, τον μετασχηματισμό των λέξεων σε διανύσματα πραγματικών αριθμών. Ονομαστικά, οι βασικές μέθοδοι είναι οι παρακάτω:

- One – hot κωδικοποίηση
- TF – IDF μετασχηματισμός
- Word2Vec
- GloVe
- FastText
- Poincare κωδικοποίηση

### One Hot Κωδικοποίηση – Count Vectorization

Όπως ειπώθηκε πριν, η ίδια αποτελεί μία απλή μέθοδο κωδικοποίησης των λέξεων προς μετασχηματισμό που βασίζεται στην καταμέτρηση των διαφορετικών τιμών που μπορούν να πάρουν (πχ για το φύλο δύο τιμές 0 -1, για τα αισθήματα χαρά, λύπη, εκνευρισμός κτλ). Στις NLP διαδικασίες, το σετ από το οποίο θα προκύψει η κωδικοποίηση μπορεί να είναι λέξεις, προτάσεις, παράγραφοι ακόμη και ολόκληρα άρθρα. Η φράση count vectorization προσδιορίζει ακριβώς αυτή την διαδικασία της μετατροπής της καταμέτρησης σε αριθμητικό διάνυσμα.



Εικόνα 30: One Hot Encoding

Με την ίδια λογική μία άλλη μέθοδος, η χρήση του πίνακα συνύπαρξης (Co-Occurrence Matrix) όπου, όπως ακριβώς αναφέρει και η ίδια η ονομασία, καταγράφεται η συνύπαρξη δύο λέξεων με 1. Αντίθετα μπαίνει η τιμή 0. Ουσιαστικά, σε περίπτωση όπου η ερώτηση: Αυτές οι δύο λέξεις έρχονται μαζί; Έχει θετική απάντηση τότε μπαίνει η τιμή 1.

Ένα παράδειγμα αυτή της μεθόδου είναι το παρακάτω όπου έστω ότι έχω 3 προτάσεις:

1. I enjoy flying
2. I like NLP
3. I like deep learning

Ως αποτέλεσμα το πίνακας θα είναι ο παρακάτω:

$$X = \begin{matrix} & I & like & enjoy & deep & learning & NLP & flying & . \\ \begin{matrix} I \\ like \\ enjoy \\ deep \\ learning \\ NLP \\ flying \\ . \end{matrix} & \left[ \begin{array}{cccccccc} 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{array} \right] \end{matrix}$$

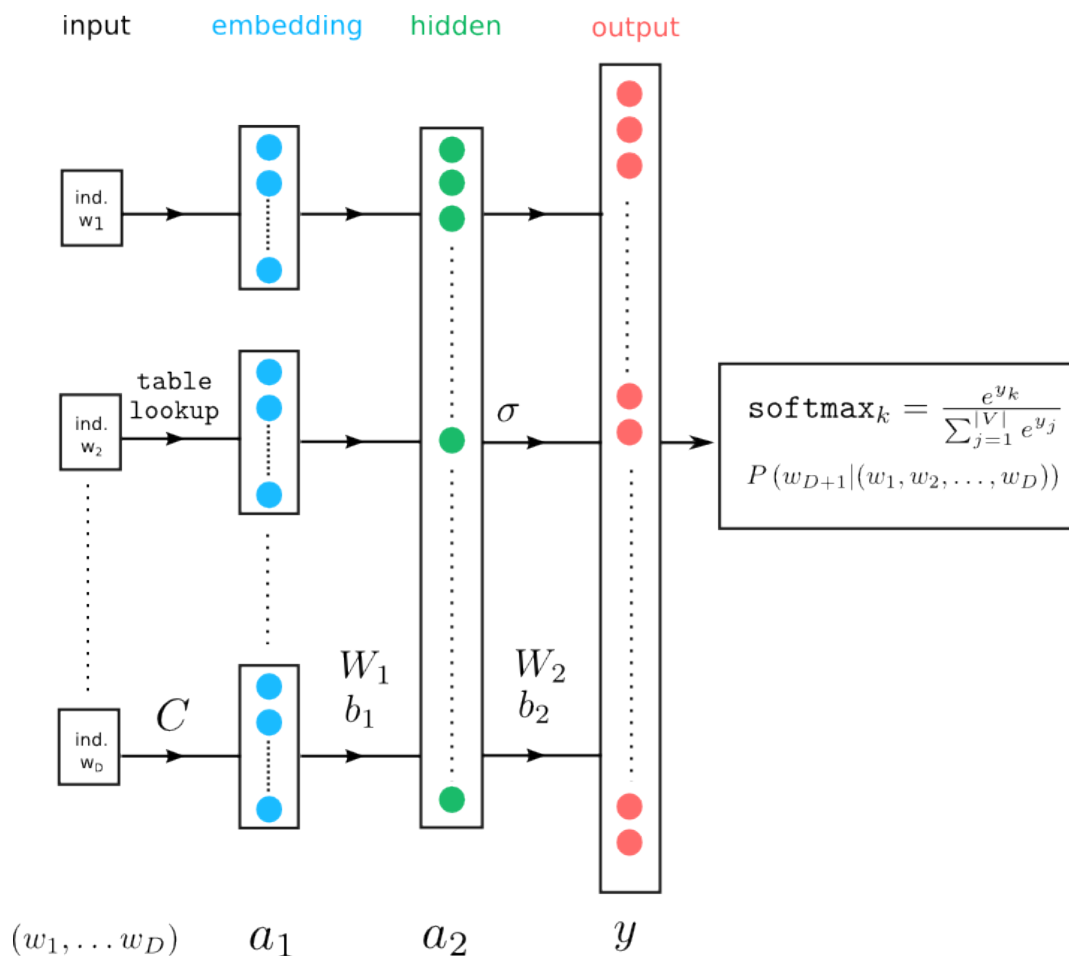
Εικόνα 31: Co-Occurrence Πίνακας

Γίνεται γρήγορα κατανοητό ότι η συγκεκριμένη κατηγορία μεθόδων για τις NLP διαδικασίες είναι σχεδόν απαγορευτική καθώς το βασικό χαρακτηριστικό της σημασιολογικής συνάφειας μεταξύ των λέξεων αποτελεί το μεγαλύτερο τροχοπέδη της συγκεκριμένης υλοποίησης που πρέπει να αξιοποιείται μόνο ως τελευταία λύση, έχοντας ως αποτέλεσμα την δημιουργία αραιών (sparse) πινάκων – μητρών (matrix) υπό την έννοια της ύπαρξης πολλών μηδενικών.

### Νεύρο – Πιθανοτικό Μοντέλο (Neuro Probabilistic)

Πριν προχωρήσουμε σε άλλες κατηγορίες, είναι σημαντικό να αναφερθεί και άλλη μία τεχνική η χρήση ενός νεύρο-πιθανοτικού μοντέλου, το οποίο συνδυάζει τις ικανότητες των νευρωνικών δικτύων με την απλή one-hot encoding διαδικασία.

Με αυτόν τον τρόπο είναι σε θέση να παράγει ναι μεν και πάλι one hot διανύσματα αλλά μέσω της χρήσης κάποιας αρχιτεκτονικής νευρωνικού δικτύου ώστε αυτή η αναπαράσταση να είναι βέλτιστη σε αντίθεση με την απλή one hot κωδικοποίηση ή τον πίνακα συνύπαρξης.



Εικόνα 32: Παράδειγμα Νεύρο-Πιθανοτικού Μοντέλου

### TF – IDF Μετασχηματισμός

Εδώ, αν και υπάρχει συνάφεια με την προηγούμενη μέθοδο, καθώς πάλι λαμβάνει χώρα αναπαράσταση λέξεων (ή παραγράφων κτλ.) που υπάρχουν στο κείμενο και μετασχηματισμός τους σε αριθμητικά διανύσματα. Η βασική διαφορά εντοπίζεται στο γεγονός ότι εδώ οι λέξεις αναπαρίστανται από την συχνότητα που χρησιμοποιούνται σαν όρος πολλαπλασιασμένη με την αντίστροφη συχνότητα που εμφανίζονται στο κείμενο. Στην ουσία, επιδιώκεται να μειωθεί σημαντικά η βαρύτητα που δίνεται σε λέξεις που εμφανίζονται υπερβολικά συχνά και παντού στο κείμενο όπως πχ άρθρα, αντωνυμίες κτλ συνδυαστικές λέξεις (και, ή, επίσης, κτλ) καθώς κατά ένα πολύ σημαντικό ποσοστό δεν προσφέρουν περισσότερες πληροφορίες.

Αντίθετα, λέξεις που εμφανίζονται ελάχιστα ή συχνά αλλά σε συγκεκριμένα σημεία, έχουν μεγαλύτερη βαρύτητα καθώς η πιθανότητά τους να φέρουν σημαντικό νόημα είναι μεγαλύτερη.

Εδώ βέβαια, το πρόβλημα της διαστατικότητας συνεχίζει να υφίσταται όπως και στις προηγούμενες μεθόδους. Με τον όρο πρόβλημα διαστατικότητας – Dimensionality Reduction, εννοούμε τα προβλήματα μηχανικής μάθησης που έχουν μεγάλο αριθμό ανεξάρτητων μεταβλητών – διαστάσεων. Είναι κατανοητό ότι όσες περισσότερες οι μεταβλητές που πρέπει να ληφθούν υπόψιν καθόλη την διαδ (40)ικασία της μηχανικής μάθησης τόσο περισσότεροι υπολογιστικοί πόροι θα καταναλωθούν καθώς και η ίδια η απόδοση του νευρωνικού δικτύου κινδυνεύει είτε από έννοιες όπως το overfitting είτε το underfitting.

## Word2Vec

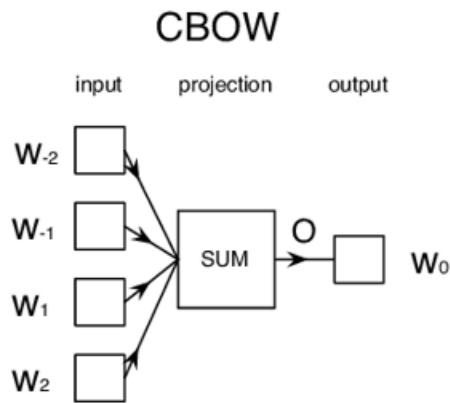
Η word2vec υλοποίηση στηρίζεται αμιγώς σε αρχιτεκτονικές νευρωνικών δικτύων και μπορεί να χωριστεί σε 3 μεγάλες κατηγορίες:

1. One-word περιεχόμενο
2. Multi-word περιεχόμενο
3. Skip-gram μοντέλο

### *One – Word Context*

Στην πρώτη κατηγορία έχουμε μία αρχιτεκτονική δικτύου με είσοδο ένα one-hot κωδικοποιημένο διάνυσμα μήκους  $V \times 1$  ή αλλιώς, όπως είναι πιο γνωστό το CBOW (Continuous Bag-of-Words) μοντέλο. Εδώ, επιδιώκουμε να προβλέψουμε την επιθυμητή (target) λέξη, προσπαθώντας να κατανοήσουμε τις γύρω λέξεις που την περιβάλλουν. Το σύνολο των λέξεων που θα διερευνηθούν ορίζεται από τον ίδιο τον χρήστη και ονομάζεται παράθυρο, ορίζοντας αντίστοιχα και τον μέγεθος του διανύσματος εισόδου στο νευρωνικό δίκτυο.

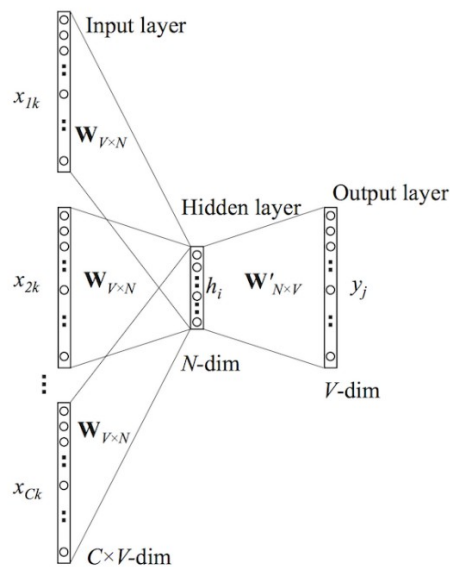
Σε σχέση με το skip gram μοντέλο που θα αναφέρουμε παρακάτω, η εκπαίδευση είναι ταχύτερη στο CBOW και με καλύτερα αποτελέσματα, υπό την προϋπόθεση όμως, ότι οι λέξεις παρουσιάζονται συχνά μέσα στο κείμενο



Εικόνα 33: High Level απεικόνιση του CBOW

*Multi – Word context*

Στην συγκεκριμένη μέθοδο, η ουσία (45) δεν αλλάζει σε σύγκριση με το CBOW, καθώς και πάλι πραγματοποιείται διερεύνηση στις γειτονικές λέξεις με σκοπό την πρόβλεψη του νοήματος. Εδώ η αλλαγή εντοπίζεται στην παραγωγή πολλαπλών πολυωνυμικών κατανομών με βάση πολλές διαφορετικές λέξεις περιεχομένου (για αυτό και η ονομασία multi word) και το άθροισμα των αντίστοιχων διανυσμάτων που παράγουν.

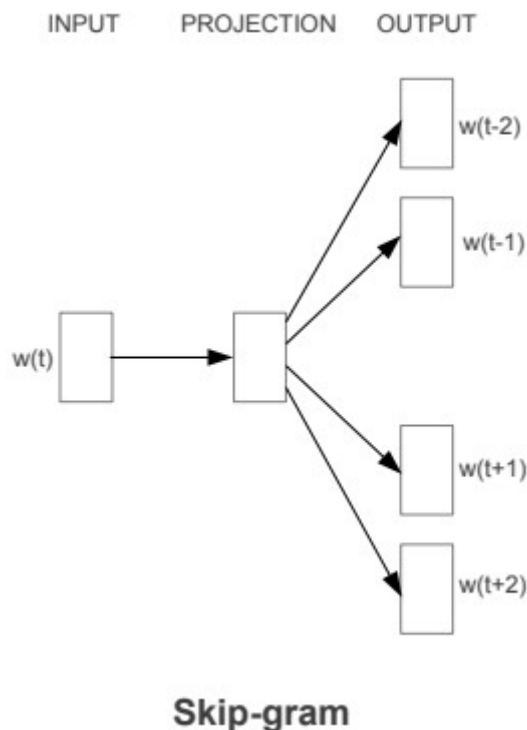


Εικόνα 34: Multi Word

### *Skip – Gram Μοντέλο*

Στην περίπτωση του Skip – Gram μοντέλου, σε αντίθεση με πριν, σαν είσοδο παρέχεται μία λέξη και πρέπει το νευρωνικό δίκτυο να προβλέψει τις αντίστοιχες γειτονικές του, πάλι με την χρήση ενός παραθύρου λέξεων. Γίνεται εύκολα κατανοητό ότι το αντίστοιχο διάνυσμα εισόδου εδώ θα είναι αντεστραμμένο ( $1 \times V$ ). Στη συνέχεια μέσω του υπολογισμού των λαθών (καθώς εμείς θα γνωρίζουμε ποιες λέξεις είναι πράγματι κοντά και ποιες όχι) και μέσω του κλασσικού αλγορίθμου backpropagation θα μπορέσουμε να καταλήξουμε και πάλι στον πίνακα του κρυφού επιπέδου που είναι ο ζητούμενος και περιλαμβάνει τα βάρη των γειτονικών λέξεων. Κατά το backpropagation μπορούμε να διακρίνουμε τα εξής βήματα:

1. Αρχικοποίηση των τιμών των βαρών
2. Εκκίνηση νέας εποχής (epoch) εκπαίδευσης
3. Μεταφορά της πληροφορίας προς τα εμπρός
4. Προς τα πίσω διάδοση των λαθών
5. Προσαρμογή των βαρών



Εικόνα 35: Skip Gram μοντέλο

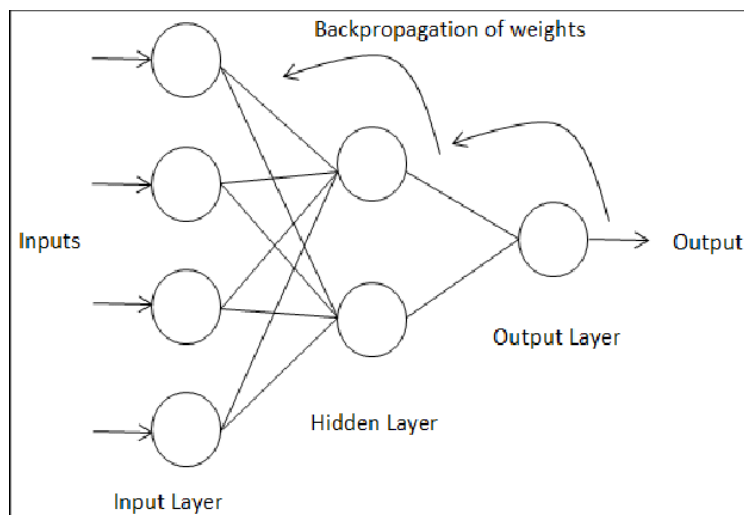
Εδώ, η μάθηση βασίζεται συνήθως στην ελαχιστοποίηση των σφαλμάτων μέτρησης μεταξύ έξοδων δικτύου και επιθυμητών εξόδων. Αυτό συνεπάγεται, πίσω διάδοση μέσω ενός δικτύου παρόμοιου ή και ακριβώς ίδιου με αυτό που γίνεται η εκπαίδευση. Για αυτόν τον λόγο και το όνομα του αλγορίθμου είναι οπίσθια διάδοσης - backpropagation. Από ιστορική σκοπιά, η μέθοδος αυτή έχει προταθεί αρκετά παλιά, αλλά είχε αγνοηθεί σε σημαντικό βαθμό λόγω του μεγάλου υπολογιστικού κόστους, ενώ συνέχεια 'ανακαλύφθηκε' ξανά γύρω στα μέσα της δεκαετίας του 80 και ουσιαστικά ξεκίνησε να αποτελεί ένα γενικά αποδεκτό εργαλείο για την εκπαίδευση του πολυεπίπεδο perceptron. Η ιδέα είναι να βρείτε το συνάρτηση ελάχιστου σφάλματος  $e(w)$  σε σχέση με τα βάρη συνδέσεων.

Μία σύντομη αναφορά για τα γενικά βήματα ενός πολυεπίπεδο perceptron πριν την ανάλυση του backpropagation είναι η παρακάτω:

Αρχικά, όλα τα βάρη του δικτύου, (46) αρχικοποιούνται με τυχαίες τιμές, το εύρος των οποίων κατανέμεται ομοιόμορφα και σε μικρή έκταση. Στην περίπτωση όπου οι τιμές αυτές προσεγγίζουν το 0 οι αντίστοιχες κλίσεις που θα υπολογιστούν θα τείνουν και αυτές στο 0, στην περίπτωση όπου έχουμε κρυφά επίπεδα, γεγονός που θα οδηγήσει στην μη εκπαίδευση του δικτύου. Περισσότερο υποδεικνύονται προσπάθειες εκπαίδευσης, με διαφορετικά αρχικά βάρη, για να βρεθεί η καλύτερη τιμή για τη συνάρτηση κόστους (ελάχιστο σφάλμα). Αντίθετα, εάν οι αρχικές τιμές είναι μεγάλες, τότε αυτές τείνουν να κορεστούν με αποτέλεσμα, η σιγμοειδής συνάρτηση ενεργοποίησης που συνήθως χρησιμοποιείται να είναι πολύ μικρή. Λόγω του γεγονός ότι θα λειτουργήσει ως πολλαπλασιαστικός παράγοντας κατά τη διάρκεια της μαθησιακής διαδικασίας η παραπάνω τιμή, οι κορεσμένες μονάδες θα είναι σχεδόν μπλοκαρισμένες, η οποία καθιστά τη μάθηση πολύ αργή. Σε μια νέα εποχή εκπαίδευσης, παρουσιάζει όλα τα παραδείγματα του σετ προπόνησης ο αλγόριθμος στο νευρωνικό δίκτυο. Στις περισσότερες περιπτώσεις, η εκπαίδευση του δικτύου περιλαμβάνει περισσότερες από μία εποχές. Κατά την εκπαίδευση, τα βάρη ρυθμίζονται μόνο μετά από όλα τα διανύσματα δοκιμής που θα εφαρμοστούν στο δίκτυο. Επομένως, για τις κλίσεις των βαρών, οι ίδιες πρέπει να απομνημονεύονται και να προσαρμόζονται μετά από κάθε μοντέλο στο αντίστοιχο σετ προπόνησης.

Σαν αποτέλεσμα τα βάρη θα αλλάξουν μόνο μία φορά (υπάρχει μια online παραλλαγή, πιο απλό, στο οποίο ενημερώνονται τα βάρη άμεσα, σε αυτήν την περίπτωση, τη σειρά με την οποία τα διανύσματα του δικτύου που παρουσιάζονται μπορεί να έχει σημασία. Όλες οι βαθμίδες των βαρών και το τρέχον σφάλμα αρχικοποιούνται με 0.



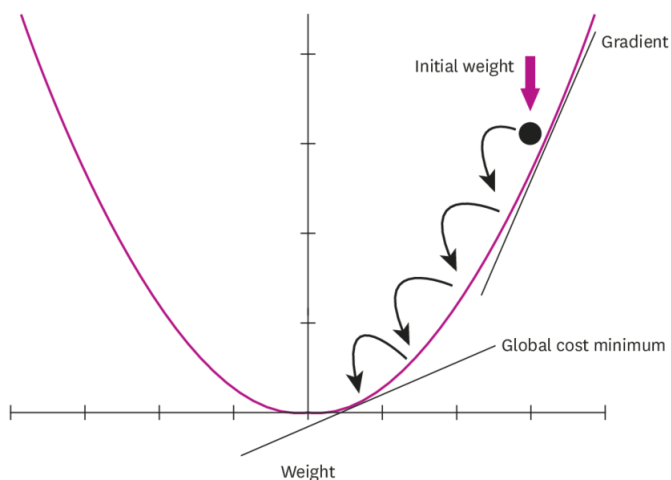


Εικόνα 36: BackPropagation

### GloVe (Global Vector for Word Representation)

Η GloVe μέθοδος αποτελεί επέκταση του word2vec και ουσιαστικά βοήθησε στην απομάκρυνση του παραθύρου λέξεων που αναφέρθηκε προηγουμένως. Μέσα από το Co occurrence πίνακα που έχει ήδη αναφερθεί επιδιώκεται να υπολογιστούν οι αντίστοιχες πιθανότητες των λέξεων που θα προσδίδουν την σημασιολογική ομοιότητα τους. Για λόγους συντομίας δεν θα προχωρήσουμε σε περαιτέρω ανάλυση της συγκεκριμένης μεθόδου παρά θα αναφερθούμε στην σημασία της συνάρτησης κόστους η οποία παίζει μεγάλη σημασία στην συγκεκριμένη μέθοδο. Η ίδια αφορά τον βασικό τρόπο μέτρησης του σφάλματος μεταξύ της πραγματικής και επιθυμητής εξόδου. Προφανώς, κάθε αλγόριθμος βημάτων βελτιστοποίησης ενός νευρωνικού δικτύου, στοχεύει στην όσο το δυνατόν μεγαλύτερη μείωση της αριθμητικής τιμής αυτής της συνάρτησης, καθώς αυτό θα σημαίνει μικρότερη απόκλιση μεταξύ των δύο εξόδων και, εν τέλει, την παραγωγή εκτιμητών με ισχυρότερες δυνατότητες πρόβλεψης.

Μία αρκετά γνωστή συνάρτηση είναι, αυτήν του μέσου τετραγωνικού κόστους (Mean Squared Error – MSE).



Εικόνα 37: Συνάρτηση κόστους

Σε παρόμοια λογική, χρησιμοποιείται και η λογαριθμική συνάρτηση κόστους ή εντροπίας:

Εδώ, η ίδια εντοπίζεται περισσότερο συχνά σε προβλήματα ταξινόμησης. Ως κλίση – Gradient ονομάζουμε το διάνυσμα, για  $n$  μεταβλητές ενός προβλήματος, το διάνυσμα μήκους  $n$  που αναπαριστά την διεύθυνση σύμφωνα με την οποία το κόστος της συνάρτησης αυξάνεται πιο γρήγορα.

Για παράδειγμα, έστω μία γραμμική εξίσωση:

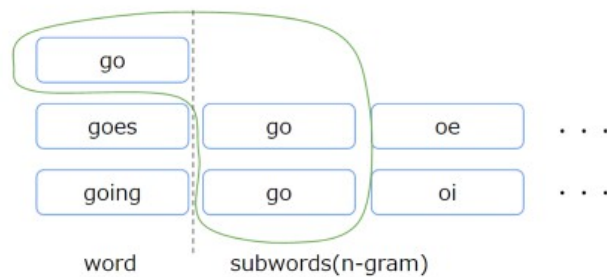
Και έστω ότι θέλουμε σε αυτήν να βρούμε το ελάχιστο σφάλμα οπότε η παραπάνω είναι και η συνάρτηση κόστους. Στην συγκεκριμένη περίπτωση η κλίση θα είναι προφανώς η:

Στην περίπτωση αυτή, παίρνουμε ένα, αρχικό τυχαίο σημείο και κοιτάμε να δούμε πώς μπορούμε να φτάσουμε σε κάποιο ελάχιστο, ελπίζοντας ότι θα είναι ολικό.

### FastText

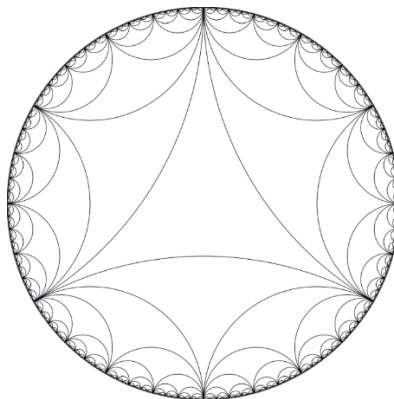
Με την μέθοδο FastText ξεκινάμε να λαμβάνουμε υπόψιν τη μορφολογία της λέξης, αξιοποιώντας ένα skip gram μοντέλο. Αρχικά, το ίδιο σπάει όλες τις λέξεις σε  $n$ -gram

χαρακτήρες (μεγέθους συνήθως μεταξύ 3 και 6) δημιουργώντας ύπο-λέξεις τις οποίες μπορεί στη συνέχεια να τις αξιοποιήσει σε περίπτωση που συναντήσει κάποιο παράγωγο τους σε μία άλλη σύνθετη. Σε αντίθεση με τις υπόλοιπες μεθόδους όπου ένα τέτοιο παράδειγμα θα το ταξινομούσαν στην κατηγορία των άγνωστων λέξεων, στην περίπτωση όπου βρεθεί μία γνωστή, είναι δυνατόν να διαμορφωθεί κάπως το νόημα της λέξης σε αυτή την μέθοδο. Αρκετά σημαντικό είναι το γεγονός ότι η εταιρία Facebook που δημοσίευσε αυτή την μέθοδο έχει ήδη εκπαιδευμένα pretrained FastText διανύσματα σε 294 διαφορετικές γλώσσες .



**Εικόνα 38: FastText**

Τέλος, συνοπτικά πρέπει να αναφερθούμε και στην κωδικοποίηση Poincare με σκοπό να προχωρήσουμε στη συνέχεια στα RNN δίκτυα. Η ίδια, στην βάση της επιδιώκει να βρει ιεραρχικές ιδιότητες των λέξεων και σημασιολογικές συνδέσεις αξιοποιώντας το υπερβολικό επίπεδο σε αντίθεση με την αντίστοιχη ευκλείδεια αναζήτηση που λαμβάνει χώρα στις υπόλοιπες μεθόδους. Η ονομασία της προέρχεται από το Poincare μοντέλο μπάλας που αξιοποιείται στην υπερβολική γεωμετρία όπως φαίνεται και παρακάτω:



**Εικόνα 39: Poincare μοντέλο**

Παρόλες τις παραπάνω μεθόδους όμως, το πρόβλημα της αντίληψης του γενικού περιεχομένου συνεχίζει να είναι έκδηλο καθώς όλες οι παραπάνω μέθοδοι επιδιώκουν την εύρεση γειτονικών λέξεων, είτε ευθέως είτε αντίστροφα. Για τον λόγο αυτό αξιοποιήθηκε ένα

καινούργιο είδος αρχιτεκτονικών νευρωνικών δικτύων, τα Αναδρομικά Νευρωνικά Δίκτυα Recurrent Neural Networks RNNs, τα οποία αναφέραμε παραπάνω.

## Ανάλυση της Προσέγγισης

Στη συνέχεια, προχωράμε στην ανάλυση του υπό μελέτη paper, την αρχιτεκτονική και της μεθόδους του καθώς και στην περαιτέρω περιγραφή των εννοιών που είναι απαραίτητες για την πλήρη κατανόησή τους.

## Προ επεξεργασία Κειμένου

Αρχικά, σύμφωνα με τους ερευνητές, το πρώτο κομμάτι της προ επεξεργασίας, σε αρχικό βαθμό, συγκεντρώθηκε γύρω από την αφαίρεση – απομόνωση ορισμένων χαρακτήρων που σχετίζονται με την πλατφόρμα του Twitter, όπως hashtags, emojis και credentials (username). Για την συγκεκριμένη διαδικασία, οι ίδιοι αξιοποίησαν μία έτοιμη βιβλιοθήκη στην γλώσσα προγραμματισμού python, η οποία πέραν από βασικό καθαρισμό προσφέρει και δυνατότητες tokenization.

Για την συγκεκριμένη περίπτωση, πάντως, παρουσιάζεται για παράδειγμα μία βασική διαδικασία – λειτουργία που προσφέρει ως προς τον καθαρισμό:

Καθώς και μέσα στον κώδικα των ερευνητών, εδώ:

Ενώ παρακάτω, φαίνεται, μετά την προ επεξεργασία και τα πρώτα πέντε δεδομένα που περιλαμβάνονται στο σετ, με την χρήση της εντολής head().

```
df.head()
```

	id	tweet	label
0	1	The CDC currently reports 99031 deaths. In gen...	real
1	2	States reported 1121 deaths a small rise from ...	real
2	3	Politically Correct Woman (Almost) Uses Pandem...	fake
3	4	#IndiaFightsCorona: We have 1524 #COVID testin...	real
4	5	Populous states can generate large case counts...	real

Εικόνα 40: Σετ δεδομένων

Στη συνέχεια προχωράμε στην διαδικασία του Tokenization, όπου οι ίδιοι, με βάση το εκάστοτε μοντέλο που έχει εκπαιδευτεί (BERT, DeBERTa, ROBERTa, XLMROBERTa, XLNet) έχουν χρησιμοποιήσει και την αντίστοιχη tokenization τεχνική.

```
def get_dataloader(network, train_sentences, val_sentences, train_labels, val_labels, test_sentences, train_token_ids, val_token_ids, test_token_ids):
    dic = {"XLNET": "xlnet-base-cased", "ROBERT": "roberta-base", "XLM-ROBERT": "xlm-roberta-base", "DeBERTa": "microsoft/deberta-base", "Electra": "google/electra-base-discriminator"}
    if network=="XLNET":
        tokenizer = XLNetTokenizer.from_pretrained(dic[network], do_lower_case=True)
    elif network=="DeBERTa":
        tokenizer = DebertaTokenizer.from_pretrained(dic[network], do_lower_case=True)
    elif network=="Electra":
        tokenizer = ElectraTokenizer.from_pretrained(dic[network], do_lower_case=True)
    else:
        tokenizer = RobertaTokenizerFast.from_pretrained(dic[network], do_lower_case=True)

    tokenized_train_texts = [tokenizer.tokenize(sent) for sent in train_sentences]
    print ("Tokenize the first sentence:")
    print (tokenized_train_texts[0])
```

Εικόνα 41: Φόρτωμα διαφορετικών tokenizers

Για να κατανοήσουμε την αναγκαιότητα διαχωρισμού της μεθόδου, πρέπει πρώτα να δούμε τα διαφορετικά αυτά μοντέλα.

Αρχικά, μπορούμε να παρατηρήσουμε ότι τα περισσότερα μοντέλα, περιλαμβάνουν – έχουν ως παράγωγο την λέξη BERT μέσα. Η ίδια απορρέει από τις εξής λέξεις:

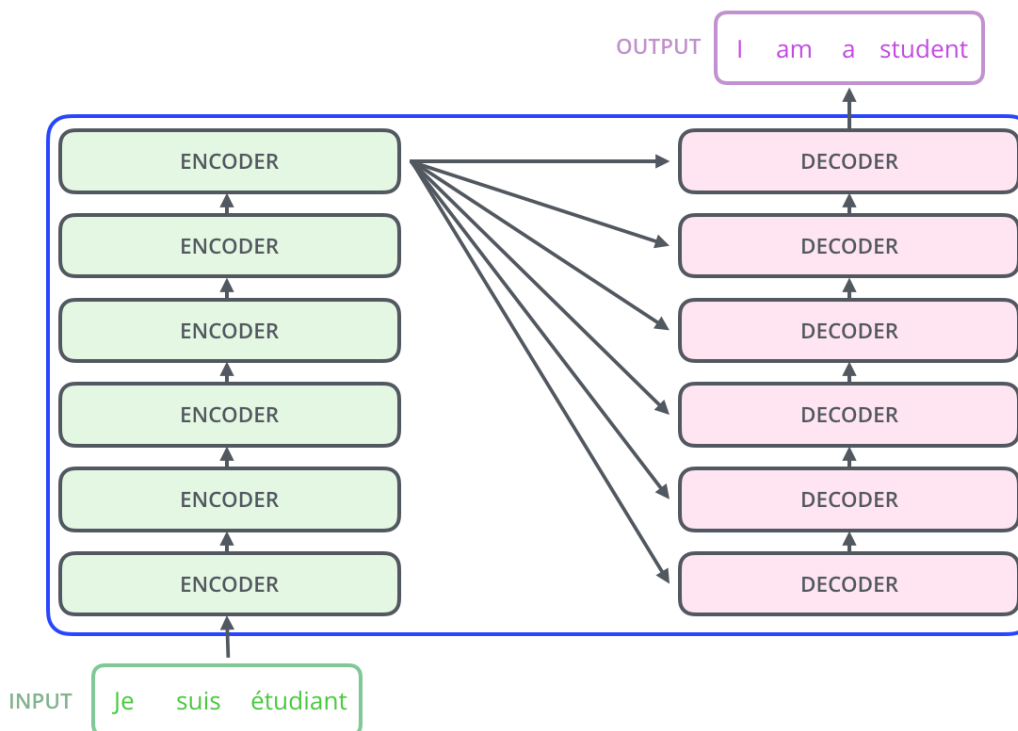
- Bidirectional
- Encoder
- Representations from
- Transformers

Και προέρχεται από ένα αρκετά πρόσφατο paper της ερευνητικής ομάδας της Google γύρω από τον χώρο του NLP (47). Αποτέλεσε, σημαντική καινοτομία στον χώρο του NLP καθώς εισήγαγε, όπως είδαμε και στα RNNs την εκπαίδευση και προς στις δύο κατευθύνσεις (bi directional) σε ενός Transformer, σε αντίθεση με τις υπόλοιπες προσεγγίσεις όπου αντιμετώπιζαν/διάβαζαν την αλληλουχία του κειμένου προς μία κατεύθυνση, είτε αριστερά είτε δεξιά, γεγονός που οδήγησε στην κατασκευή/εκπαίδευση pre trained μοντέλων που μπορούσαν να πετύχουν κατανόηση εις βάθος του κειμένου. Για να μπορέσουμε όμως να κατανοήσουμε την σημασία του πρέπει να εξηγήσουμε λίγο περισσότερο την έννοια του Transformer και γιατί ο ίδιος βρίσκεται στο επίκεντρο της BERT υλοποίησης.

```
!pip install pytorch-transformers
!pip install transformers
```

Εικόνα 42: Φόρτωμα Transformers στον κώδικα

Ο Μετασχηματιστής – Transformer, επιδιώκει (η BERT τεχνική ουσιαστικά) να κατασκευάσει ένα μοντέλο το οποίο να είναι σε θέση να κατανοήσει την σύνδεση μεταξύ των λέξεων σε μία πρόταση, για να είναι σε θέση, στην περίπτωσή μας, να αναγνωρίσει ψευδείς/αληθείς ειδήσεις. Τα δύο μεγάλα κομμάτια ενός Transformer, είναι ο κωδικοποιητής του (**encoder**) ο οποίος είναι υπεύθυνος για την ανάγνωση του κειμένου (σε άλλες προσεγγίσεις όπως είπαμε η ανάγνωση γίνεται προς μία φορά, ενώ εδώ όλη η αλληλουχία των λέξεων διαβάζεται με την μία) (48).

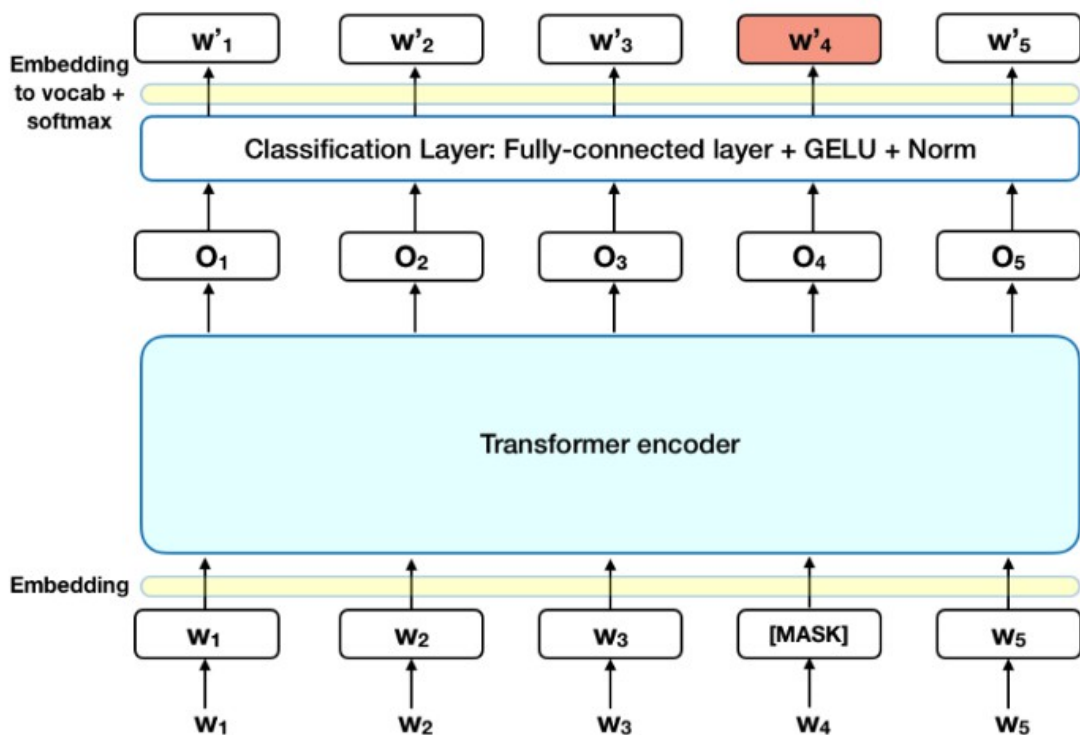


Εικόνα 43: Encoders – Decoders

Αντίθετα, ο αποκωδικοποιητής (**decoder**) είναι εκείνος που είναι υπεύθυνος για την παραγωγή του γλωσσικού μοντέλου, σε περίπτωση όπως στην εικόνα που αναφερόμαστε σε εφαρμογές μετάφρασης. Περισσότερο αναλυτικά, θα τον δούμε στη συνέχεια, αφού ολοκληρώσουμε την ανάλυση του encoder.

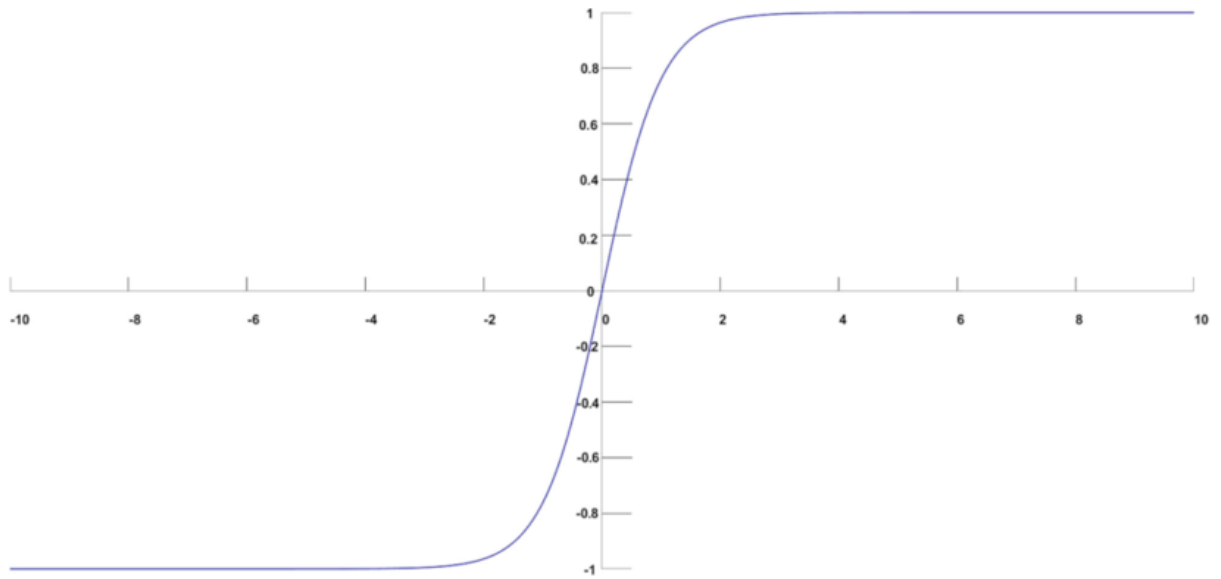
Συνεχίζοντας, λοιπόν, στην είσοδο, ο encoder, δέχεται μία αλληλουχία από token (βλ. παραπάνω) τα οποία οργανώνονται στα αντίστοιχα διανύσματα εισόδου για το νευρωνικό δίκτυο το οποίο είναι υπεύθυνο για την επεξεργασία του. Το αποτέλεσμα της παραπάνω διαδικασίας, παράγει μία αλληλουχία διανυσμάτων συγκεκριμένου μεγέθους, τα οποία αντιστοιχούν στο εκάστοτε token μέσω μίας index τιμής.

Στο ίδιο paper που αναφέραμε στην αρχή (47), οι ερευνητές παρείχαν λεπτομέρειες και για μία τεχνική που θα επέτρεπε αυτή την εκπαίδευση δύο κατευθύνσεων, σε αντίθεση με όσες υπήρχαν μέχρι τότε που ευνοούσαν συγκεκριμένη φορά. Η τεχνική αυτή ονομάστηκε **Masked LM (MLM)** και εντοπίζεται στην αντικατάσταση του 15% των λέξεων σε κάθε ακολουθία που θα τοποθετηθεί ως είσοδος στο νευρωνικό δίκτυο με ένα MASK token. Συγκεκριμένα από αυτό το 15%, το 80% είναι πράγματι MASK, ενώ από 10% αντιστοιχεί σε τυχαίες λέξεις και τις πραγματικές – εκείνες δηλαδή που υπήρχαν όντως στην ίδια θέση στην ακολουθία εισόδου.



Εικόνα 44: Δομή του Encoder

Για να υλοποιηθεί το συγκεκριμένο, τοποθετείται, στην έξοδο του κωδικοποιητή ένα επίπεδο (layer) που σχετίζεται με την κατηγοριοποίηση, ώστε στη συνέχεια να πολλαπλασιάσουμε τα διανύσματα εξόδου με έναν αντίστοιχο πίνακα που θα τα μετατρέψει το μέγεθός τους σε κατάλληλο για λεξικό (ουσιαστικά το βήμα **Embedding to vocab +**) και τέλος θα προβλέψει την πιθανότητα της κάθε λέξης σε αυτό το λεξικό με την χρήση της softmax συναρτήσεως (**softmax** που γράφει και στο σχήμα) η οποία έχει την παρακάτω μορφή:



Εικόνα 45: Softmax συνάρτηση

Έχοντας τα παραπάνω ως δεδομένο, η συνάρτηση κόστους (βλ. παραπάνω) του BERT αλγορίθμου, λαμβάνει υπόψιν μόνο την πρόβλεψη για την MASK τιμή και όχι εκείνες για τα token – λέξεις που δεν έχει χρησιμοποιηθεί, γεγονός που προκαλεί μεγάλες καθυστερήσεις στην εκπαίδευση του μοντέλου. Το γεγονός βέβαια ότι η ίδια η εφαρμογή αξιοποιείται και ως εργαλείο ερωταπαντήσεων μπορεί να δικαιολογηθεί από την ύπαρξη του επόμενου βήματος – διαδικασίας, της πρόβλεψης επόμενης πρότασης (**Next Sentence Prediction - NSP**). Η βασική παραδοχή της, μπορεί να εντοπιστεί στην υπόθεση ότι κατασκευάζουμε δύο προτάσεις (οι οποίες έρχονται ανά ζεύγη στην είσοδο) από το μοντέλο και το ίδιο επιδιώκει να μάθει να προβλέπει αν η δεύτερη είναι αυτή που πράγματι ακολουθεί την πρώτη στο πραγματικό κείμενο.



```

# Create a mask of 1s for each token followed by 0s for padding
for seq in input_test_ids:
    seq_mask = [float(i>0) for i in seq]
    test_attention_masks.append(seq_mask)
test_masks = np.array(test_attention_masks)

train_inputs = torch.tensor(input_train_ids)
validation_inputs = torch.tensor(input_val_ids)
test_inputs = torch.tensor(input_test_ids)
train_labels = torch.tensor(train_labels)
validation_labels = torch.tensor(val_labels)
train_masks = torch.tensor(train_masks)
validation_masks = torch.tensor(validation_masks)
test_masks = torch.tensor(test_masks)

train_data = TensorDataset(torch.tensor(train_token_ids), train_inputs, train_masks, train_labels)
train_sampler = RandomSampler(train_data)
train_dataloader = DataLoader(train_data, sampler=train_sampler, batch_size=batch_size)

validation_data = TensorDataset(torch.tensor(val_token_ids), validation_inputs, validation_masks, validation_labels)
validation_sampler = SequentialSampler(validation_data)
validation_dataloader = DataLoader(validation_data, sampler=validation_sampler, batch_size=batch_size)

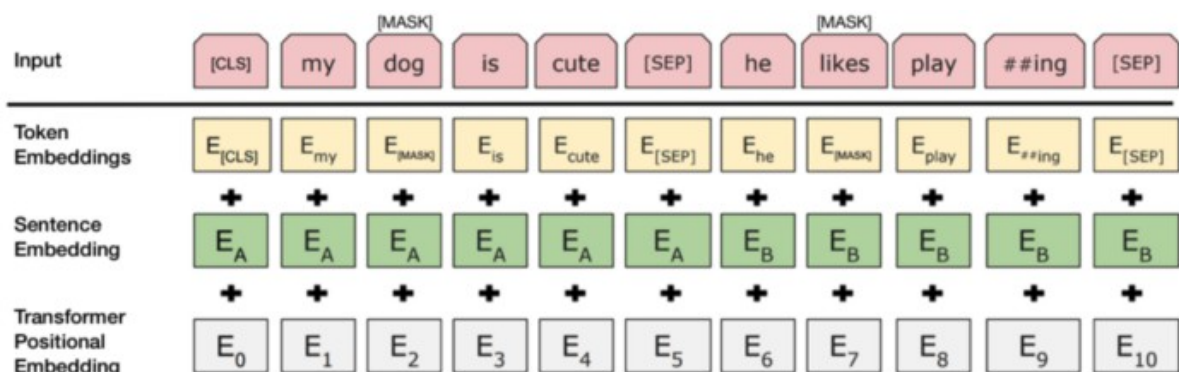
test_data = TensorDataset(torch.tensor(test_token_ids), test_inputs, test_masks)
test_sampler = SequentialSampler(test_data)
test_dataloader = DataLoader(test_data, sampler=test_sampler, batch_size=batch_size)

return train_dataloader, validation_dataloader, test_dataloader

```

Εικόνα 46: Δημιουργία μασκών

Για να υλοποιηθεί αυτό, οι εισοδοί χωρίζονται στη μέση (50-50) όπου οι μισές πράγματι η μία πρόταση ακολουθεί την άλλη, ενώ στο υπόλοιπο όχι. Πριν την είσοδο στο μοντέλο, ένα CLS token τοποθετείται στην αρχή της πρώτης πρότασης και ένα αντίστοιχο SEP στο τέλος. Ταυτόχρονα δύο διαφορετικές μεθόδους προσθήκης αξιοποιούνται και προστίθενται σε κάθε token, η μία που αφορά την θέση του token στην ακολουθία και η άλλη αν ανήκει στο πρώτο γκρουπ ή στο δεύτερο. Τέλος, τα παραπάνω αξιοποιούνται με την εισαγωγή όλης της εισόδου ακολουθίας από τον Transformer, την δημιουργία στην έξοδο του CLS token, ενός 2X1 διανύσματος, αξιοποιώντας ένα layer κατηγοριοποίησης και πάλι τον υπολογισμό για το αν η πρόταση είναι η επόμενη μέσω της Softmax.

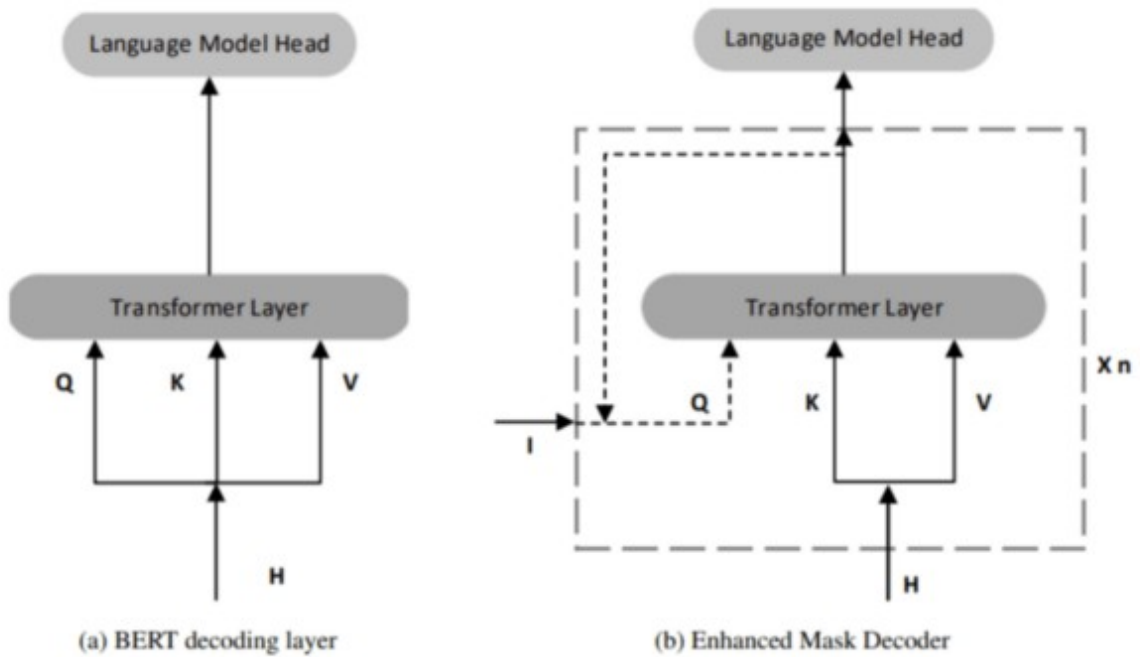


Εικόνα 47: Next Sentence Prediction

Όπως είδαμε στο paper των ερευνητών, όμως, οι ίδιοι αξιοποιούν και παραλλαγές του βασικού μοντέλου όπως τον **ROBERTa** (Robustly Optimized BERT Pretraining Approach).

Εδώ, οι ερευνητές κατάφεραν να πετύχουν μικρότερο χρόνο εκπαίδευσης (αναφερόμαστε πάντα σε pre trained μοντέλα), προχωρώντας, αρχικά στην τροποποίηση της NSP διαδικασίας, όπου αφού δοκίμασαν διαφορετικές εκδόσεις του σφάλματος της πρόβλεψης, κατέληξαν ότι η αφαίρεσή του βοηθούσε σημαντικά στην μείωση του απαραίτητου χρόνου. Επιπλέον, αυξήθηκε σημαντικά η είσοδος για την εκπαίδευση της συγκεκριμένης αρχιτεκτονικής αφού πλέον και τα batches αλλά και οι ακολουθίες, γενικότερα, ήταν σημαντικά μεγαλύτερες, γεγονός που βελτίωσε τόσο την απόδοση όσο και την μοντελοποίηση των ζητούμενων γλωσσικών μοντέλων. Τέλος, η χρήση των MASK tokens που αναφέραμε πιο πάνω, αντί να πραγματοποιηθεί μία φορά κατά την προεπεξεργασία, γεγονός που θα οδηγούσε στην κατασκευή μίας στατικής μάσκας, τα δεδομένα αντιγράφηκαν και η διαδικασία έτρεξε 10 φορές σε 40 διαφορετικές εποχές εκπαίδευσης, γεγονός που σημαίνει ότι μόνο 4 εποχές είχαν κάθε φορά την ίδια μάσκα. Από τα παραπάνω και από την ευρεία βελτίωση της αποδοτικότητας του συγκεκριμένου αλγορίθμου, προέκυψε και το όνομα **Robustly Optimized**

Σχετικά με τις υπόλοιπες υλοποιήσεις του BERT, ο **DeBERTa** (Decoding enhanced **BERT** with disentangled attention), ήταν αυτός που σαν υλοποίηση, πρότεινε τον Enhanced Mask Decoder (EMD) ο οποίος, περιλάμβανε τις απόλυτες θέσεις του εκάστοτε token μέσα στην πρόταση, σε αντίθεση με τις σχετικές που υπήρχαν στον στάνταρ μηχανισμό που έχουμε μέχρι τώρα μελετήσει.



Εικόνα 48: Λειτουργία του EMD

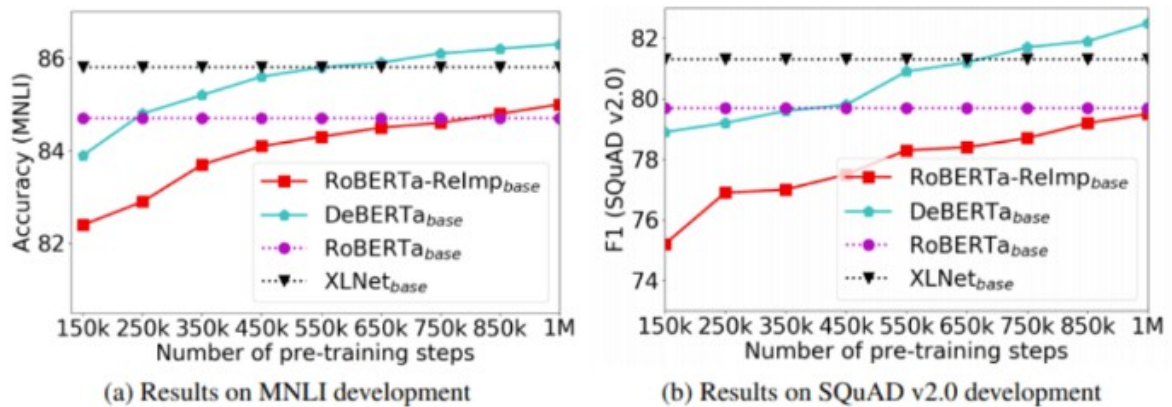


Figure 1: Pre-training performance curve between DeBERTa and its counterparts on the MNLi and SQuAD v2.0 development set.

Εικόνα 49: Συγκρίσεις μεταξύ των διαφορετικών μοντέλων

Τα μοντέλα αυτά, αξιοποιήθηκαν από τους ερευνητές, σε συνδυασμό, αφού, μετά την εκάστοτε, ατομική εκπαίδευσή τους, το σύνολο των προβλέψεών τους ήταν αυτό που έδινε την τελική πρόβλεψη για το αν ένα tweet ήταν αληθές ή όχι (Soft/Hard Training όπως αναφέρεται και μέσα στη βασική μελέτη).

```

def get_model_preds_labels(model, dataloader, mode='train'):
    data_vectors = []
    labels = []
    ids = []

    with torch.no_grad():
        correct = 0
        total = 0
        for i, batch in enumerate(dataloader):
            batch = tuple(t.to(device) for t in batch)
            if mode=='test':
                token_ids, b_input_ids, b_input_mask = batch
                outputs = model(b_input_ids, token_type_ids=None, attention_mask=b_input_mask)
                for point, tk_id in zip(F.softmax(outputs.logits).tolist(), token_ids.tolist()):
                    data_vectors.append(point)
                    ids.append(tk_id)
            else:
                # Unpack the inputs from our dataloader
                token_ids, b_input_ids, b_input_mask, b_labels = batch
                # Forward pass
                outputs = model(b_input_ids, token_type_ids=None, attention_mask=b_input_mask)
                for point, lab, tk_id in zip(F.softmax(outputs.logits).tolist(), b_labels.tolist(), token_ids.tolist()):
                    data_vectors.append(point)
                    labels.append(lab)
                    ids.append(tk_id)

    if mode=='test':
        return data_vectors, ids
    else:
        return data_vectors, labels, ids

```

Εικόνα 50: Παραγωγή προβλέψεων

## BIBΛΙΟΓΡΑΦΙΑ

1. **Rajan, Amol.** Fake news: Too important to ignore. *BBC*. [Ηλεκτρονικό] 16 January 2017. <https://www.bbc.com/news/entertainment-arts-38636042>.
2. **Dhir, Shalini Talwar & Amandeep.** Sharing of fake news on social media. [Ηλεκτρονικό] November 2020. [www.sciencedirect.com](http://www.sciencedirect.com).
3. **Benjamin D. Horne, Sibel Adali.** This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News. [Ηλεκτρονικό] 26 March 2017.
4. **Samah M. Alzanina, Aqil M. Azmia.** Detecting rumors in social media: A survey . Dubai, United Arab Emirates : s.n., 2018.
5. **Γεωργακόπουλος, Θεόδωρος.** Ο ορισμός των fake news και οι τρόποι καταπολέμησής τους. [Ηλεκτρονικό] 06 Μαρτίου 2019. <https://www.euractiv.gr/section/ekloges/opinion/o-orismos-ton-fake-news-kai-oi-tropoi-katapolemisis-toys-ti-kanei-to-fainomeno-pio-epikairo-apo-pote/>.
6. **Dawes, Sarah T. Malamut & Molly.** Characteristics of Rumors and Rumor Victims in Early Adolescence: Rumor Content and Social Impact. [Ηλεκτρονικό] March 2018.
7. **Vosoughi, Soroush.** Automatic Detection and Verification of Rumors on. [Ηλεκτρονικό] June 2015.
8. **Σωτηρόπουλος, Βασίλης.** Ορισμός της σάτιρας από ελληνικό δικαστήριο. *elawyer.com*. [Ηλεκτρονικό] 22 Ιουνίου 2006. [http://elawyer.blogspot.com/2006/06/blog-post\\_22.html](http://elawyer.blogspot.com/2006/06/blog-post_22.html).
9. **ΤΣΙΑΜΗΤΡΟΣ, ΙΩΑΝΝΗΣ.** ΠΡΟΠΑΓΑΝΔΑ ορισμός, ιστορική αναδρομή, σύγχρονη προπαγάνδα. [Ηλεκτρονικό] 12 Απριλίου 2017.
10. <https://www.kaggle.com/general>.
11. **Matthew L Newman's , James W. Pennebaker.** Lying Words: Predicting Deception from Linguistic Styles. [Ηλεκτρονικό] June 2003.
12. **James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn.** The Development and Psychometric Properties of LIWC. *University of Texas at Austin*. [Ηλεκτρονικό] 2015.

13. YOUNGOV. [www.yougov.gr](http://www.yougov.gr). <https://docs.cdn.yougov.com/1gat0ur42b/YouGov%20-%20Social%20media%20and%20charity%20Results.pdf>. [Ηλεκτρονικό]
14. Lukeprog. A History of Bayes' Theorem. [Ηλεκτρονικό] 29 August 2011.
15. Χαράλαμπος, Γούλας. "Μηχανική Μάθηση," στο Σχεδιασμός και Ανάπτυξη Αλγορίθμου Συσταδοποίησης Μεγάλης Κλίμακας Δεδομένων. . [Ηλεκτρονικό] 2015.
16. D. Anderson, G. McNeill. *Artificial Neural Networks Technology*. 1992.
17. Bishop, Christopher M. *Neural Networks for Pattern Recognition*. 1995.
18. Εφαρμογές, Νευρωνικά Δίκτυα και. *Νευρωνικά Δίκτυα και Εφαρμογές*. s.l. : Ελληνικό Ανοικτό Πανεπιστήμιο, 2001.
19. Iqbal, Muhammad. SUPERVISED MACHINE LEARNING APPROACHES: A SURVEY. [Ηλεκτρονικό] April 2015.
20. Geoffrey E. Hinton, Terrence Joseph Sejnowski. *Unsupervised Learning: Foundations of Neural Computation*. 1999.
21. Haykin, Simon. *Neural Networks and Learning Machines*. 2009.
22. Dwivedi, Rohit. What Is Naive Bayes Algorithm In Machine Learning? <https://www.analyticssteps.com/blogs/what-naive-bayes-algorithm-machine-learning>. [Ηλεκτρονικό] 29 April 2020.
23. Gandhi, Rohith. Support Vector Machine — Introduction to Machine Learning Algorithms. [Ηλεκτρονικό] 7 June 2018. <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>.
24. Stecanella, Bruno. <https://monkeylearn.com/blog/introduction-to-support-vector-machines-svm/>. [Ηλεκτρονικό] 22 June 2017.
25. Smallcombe, Mark. What Is Data Mining? [Ηλεκτρονικό] 01 July 2020. [www.xplenty.com](http://www.xplenty.com).
26. CHEN, JAMES. <https://www.investopedia.com/terms/n/neuralnetwork.asp>. *What is a Neural Network?* [Ηλεκτρονικό] 24 September 2020.
27. Rosenber, T. J. Sejnowski and C. R. «*NETtalk: A Parallel network that learns*. s.l. : University Electrical Engineering and Computer, 1986.

28. Ruchansky, N., Sungyong S. & Liu, Y. CSI: A Hybrid Deep Model for Fake News Detection. [Ηλεκτρονικό] 2017.
29. Δημητρακόπουλος, Διογένης. [www.dw.com/el](http://www.dw.com/el). [Ηλεκτρονικό] 03 Σεπτεμβρίου 2018. <https://www.dw.com/el/%CE%B1%CE%BD%CF%84%CE%B9%CE%BC%CE%B5%CF%84%CF%89%CF%80%CE%AF%CE%B6%CE%BF%CE%BD%CF%84%CE%B1%CF%82-%CF%84%CE%B1-fake-news/a-45329714>.
30. TWIN, ALEXANDRA. Data Mining. [Ηλεκτρονικό] 20 September 2020. [www.investopedia.com](http://www.investopedia.com).
31. Rouse, Margaret. DEFINITION DATA MINING. [Ηλεκτρονικό] September 2020. [www.searchsqlserver.com](http://www.searchsqlserver.com).
32. Knowledge Base. Πόσο σημαντικό είναι το data mining. [Ηλεκτρονικό] 24 Σεπτέμβριος 2019.
33. Narkhede, Sarang. Understanding Confusion Matrix. [Ηλεκτρονικό] 09 May 2018. <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>.
34. Shung, Koo Ping. Accuracy, Precision, Recall or F1? [Ηλεκτρονικό] 15 March 2018. <https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9>.
35. Shrivastav, Namratesh. Confusion Matric(TPR,FPR,FNR,TNR), Precision, Recall, F1-Score. [Ηλεκτρονικό] Januar 2018.
36. *Logistic Regression Analysis*. Hoffman, Julien I.E. 2015, Biostatistics for Medical and Biomedical Practitioners.
37. *Overfitting in prediction models – Is it a problem only in high dimensions?* Simon, Richard. 2013, Contemporary Clinical Trials.
38. *Simple Linear Regression*. 2017, Biostatistics and Computer-based Analysis of Health Data using SAS.
39. Walczak, Steven και Cerpa, Narciso. Artificial Neural Networks. *Encyclopedia of Physical Science and Technology (Third Edition)*. 2003.
40. *The Importance of the Nearest Neighbor Classifier*. Davies, E.R. 2005, Statistical Pattern Recognition.
41. *Naive Bayes Classifier*. 2018, Digital Communications and Networks.

42. *Decision Trees*. Dev, Vikrant A. και Eden, Mario R. 2019, Proceedings of the 9th International Conference on Foundations of Computer-Aided Process Design.
43. *Support Vector Machine*. Pisner, Derek A. και Schnyer, David M. 2020, Machine Learning.
44. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. Karen, Simonyan και Andrew, Zisserman. 2015, Computer Vision and Pattern Recognition.
45. *Going deeper with convolutions*. Christian, Szegedy, και συν. 2015.
46. *ImageNet Classification with Deep Convolutional Neural Networks*. Alex, Krizhevsky, Ilya, Sutskever και Geofrey, Hinton. 2012.
47. *BERT: Pre-training of Deep Bidirectional Transformers for Language Processing*. Toutanova, Jacob Devlin Ming-Wei Chang Kenton Lee Kristina. 2019.
48. *Attention Is All You Need*. Ashish Vaswani, Noam Shazeer, Niki Parmar. 2017.
49. <https://el.wikipedia.org/wiki/Προπαγάνδα>. [Ηλεκτρονικό]
50. Filippo Menczer & Thomas Hills. The Attention Economy. [Ηλεκτρονικό] December 2020. <https://www.scientificamerican.com/article/information-overload-helps-fake-news-spread-and-social-media-knows-it/>.