



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Ταξινόμηση και Διαδικτυακή Εφαρμογή  
Αναζήτησης Ηλεκτρονικών Βιβλίων με  
Τεχνικές Βαθιάς Μάθησης

Εφαρμογή στη συλλογή ebooks του Springer

---

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΞΗΡΟΥ Γ. ΒΑΣΙΛΕΙΟΥ

Επιβλέπων

Μήτρου Νικόλαος  
Καθηγητής Ε.Μ.Π.

Αθήνα, Απρίλιος 2021





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Ταξινόμηση και Διαδικτυακή Εφαρμογή  
Αναζήτησης Ηλεκτρονικών Βιβλίων με  
Τεχνικές Βαθιάς Μάθησης

Εφαρμογή στη συλλογή ebooks του Springer

---

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

**ΞΗΡΟΥ Γ. ΒΑΣΙΛΕΙΟΥ**

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 2<sup>η</sup> Απριλίου 2021

.....  
Μήτρου Νικόλαος  
Καθηγητής Ε.Μ.Π.

.....  
Παπαβασιλείου Συμεών  
Καθηγητής Ε.Μ.Π.

.....  
Συκάς Ευστάθιος  
Καθηγητής Ε.Μ.Π.

Αθήνα, Απρίλιος 2021

Επιμέλεια κειμένου

Ελισάβετ Γκρίτση  
Φιλολόγος



.....

Ξηρός Γ. Βασίλειος

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

Copyright © Ξηρός Γ. Βασίλειος, 2021

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

# Περίληψη

Οι εφαρμογές της Τεχνητής Νοημοσύνης και πιο συγκεκριμένα της Μηχανικής Μάθησης, αυξάνονται εκθετικά χρόνο με τον χρόνο. Αυξάνεται διαρκώς η ανάγκη για συστηματικότερη μελέτη και έρευνα σε αυτά τα επιστημονικά πεδία. Σε αυτήν την διπλωματική εργασία, θα εφαρμοστούν τεχνικές Μηχανικής Μάθησης για κατανόηση και επεξεργασία φυσικής γλώσσας, σε ένα σύνολο δεδομένων που αποτελείται από βιβλία.

Πιο συγκεκριμένα, θα χρησιμοποιηθεί ο πίνακας περιεχομένων του κάθε βιβλίου. Αρχικά, αφού γίνει η απαραίτητη ανάλυση και επεξεργασία τους, θα κατασκευαστούν μοντέλα Επιβλεπόμενης Μάθησης, τα οποία, με τη χρήση Νευρωνικών Δικτύων, θα είναι ικανά να ταξινομήσουν ένα βιβλίο με βάση τον πίνακα περιεχομένων του. Στη συνέχεια, με τη βοήθεια των νέων χαρακτηριστικών που θα αποκτήσει το κάθε βιβλίο, θα κατασκευαστούν μοντέλα Μη Επιβλεπόμενης Μάθησης, που σκοπό θα έχουν τη δημιουργία νέων κατηγοριών (clusters) στις οποίες θα ανήκει το κάθε βιβλίο. Επιπλέον, με τη βοήθεια των χαρακτηριστικών αυτών, θα υλοποιηθεί ένα σύστημα πρότασης συναφών συγγραμμάτων. Με σκοπό την παρουσίαση και διάθεση όλων των ανωτέρω μοντέλων, αλλά και των διαθέσιμων πληροφοριών για κάθε βιβλίο, θα υλοποιηθεί μια web εφαρμογή, η οποία θα υποστηρίζει όλα τα παραπάνω και θα είναι ικανή να εξυπηρετήσει πλήθος αιτημάτων χρήστη σε διάφορα σενάρια χρήσης.

**Λέξεις-Κλειδιά:** επιστήμη υπολογιστών, τεχνητή νοημοσύνη, μηχανική μάθηση, βαθιά μάθηση, επεξεργασία φυσικής γλώσσας, επιβλεπόμενη μάθηση, μη επιβλεπόμενη μάθηση, μοντέλα, νευρωνικά δίκτυα, ανάλυση δεδομένων, οπτικοποίηση δεδομένων, ταξινόμηση, ομαδοποίηση, web εφαρμογή, βάση δεδομένων, frontend server, backend server, σενάριο χρήσης

# Abstract

Applications of artificial intelligence and more specifically of machine learning are increasing exponentially each year. The need for more study and research is getting bigger. In this diploma thesis, machine learning techniques will be applied for natural language processing tasks in a dataset that consists of books.

More specifically, the table of contents of each book will be used and, after the necessary data analysis and preprocessing is done, supervised learning models will be built using neural networks in order to be able to classify a book based on its table of contents. Subsequently, using the new features that each book gets, unsupervised learning models will be built in order to create new clusters, in which each book will belong to. Furthermore, using these features, a system that detects related books for a book, will be implemented. In order to make available all these models and also all the available information of each book, a web application will be implemented that supports all the above features and is capable of serving various user requests for various user scenarios.

**Keywords:** computer science, artificial intelligence, machine learning, deep learning, natural language processing, supervised learning, unsupervised learning, models, neural networks, data analysis, data visualizations, classification, clustering, web application, database, frontend server, backend server, user scenario

# Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κ. Νικόλαο Μήτρου για την εμπιστοσύνη που μου έδειξε αναθέτοντάς μου τη διπλωματική αυτή εργασία μέσω της οποίας μου έδωσε την ευκαιρία να ασχοληθώ με ένα πολύ ενδιαφέρον θέμα, όπως αυτό της Τεχνητής Νοημοσύνης και της φυσικής επεξεργασίας γλώσσας, καθώς και για τη στήριξη και την καθοδήγησή του καθ' όλη τη διάρκεια της εργασίας.

Αυτή η διπλωματική εργασία είναι το επιστέγασμα των προπτυχιακών σπουδών μου στο τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Ολοκληρώνοντας αυτήν την πορεία, θα ήθελα να ευχαριστήσω τους γονείς μου Γιώργο και Ιωάννα και την αδερφή μου Μαρία για όλη την υποστήριξη και την πίστη τους σε εμένα καθ' όλη τη διάρκεια των σπουδών μου. Επιπλέον, θα ήθελα να ευχαριστήσω τον θείο μου Γιάννη Ψυχάρη και τον θείο μου Γιώργο Ψυχάρη, οι οποίοι μέσω της ακαδημαϊκής εμπειρίας τους και των γνώσεών τους με βοήθησαν να κάνω σωστές επιλογές και με συμβούλευσαν σε πολλά ζητήματα και προβληματισμούς μου. Τέλος, θα ήθελα να απευθύνω ιδιαίτερες ευχαριστίες στους συμφοιτητές μου Δημήτρη Ξενία και Μιχαήλ Φιλίππου τους οποίους γνώρισα στη σχολή και περάσαμε μαζί όλα τα φοιτητικά μας χρόνια, στηρίξαμε ο ένας τον άλλον σε δύσκολες στιγμές, απαιτητικές εργασίες και εξεταστικές περιόδους, και σίγουρα τα χρόνια αυτά θα ήταν διαφορετικά χωρίς αυτούς.

# Οργάνωση του τόμου

Η παρούσα διπλωματική εργασία διαιρείται σε δύο βασικές ενότητες. Στην πρώτη ενότητα (Μέρος I) αναλύεται το επιστημονικό πεδίο της Μηχανικής Μάθησης και το δεύτερο είναι η δημιουργία και ανάπτυξη της web εφαρμογής.

Το Κεφάλαιο 1 του Μέρους I αποτελεί εισαγωγή στο αντικείμενο της Τεχνητής Νοημοσύνης και της Μηχανικής Μάθησης. Ορίζονται κάποιες βασικές έννοιες και αναλύονται βασικές κατηγορίες και τεχνικές των αντικειμένων αυτών. Στη συνέχεια, στο Κεφάλαιο 2 καλύπτονται βασικές θεωρητικές έννοιες, τύποι και θεωρία τα οποία είναι οι απαραίτητες γνώσεις που πρέπει να διαθέτει ο αναγνώστης για να προσπελάσει και να κατανοήσει τα επόμενα κεφάλαια. Το Κεφάλαιο 3 περιέχει την παρουσίαση των δεδομένων, όλες τις αναλύσεις και την επεξεργασία που έγινε, τα οποία στο σύνολό του αποτελούν μια βασική διαδικασία προτού δημιουργηθούν και εκπαιδευτούν τα μοντέλα βαθιάς μάθησης. Το Κεφάλαιο 4, το πιο βασικό κεφάλαιο της διπλωματικής, αναλύει όλες τις διαδικασίες που πραγματοποιήθηκαν προκειμένου να δημιουργηθούν και να βελτιστοποιηθούν τα μοντέλα επιβλεπόμενης και μη επιβλεπόμενης μάθησης. Πιο αναλυτικά, περιλαμβάνει τα μοντέλα ταξινόμησης, τα μοντέλα δημιουργίας των clusters και το σύστημα εύρεσης συναφών συγγραμμάτων.

Στο Μέρος II, το Κεφάλαιο 5 παρουσιάζει το θεωρητικό υπόβαθρο που πρέπει να διαθέτει ο χρήστης για να κατανοήσει όλες τις διαδικασίες και τις τεχνικές υλοποίησης της web εφαρμογής. Αυτές περιγράφονται αναλυτικά στο τελευταίο κεφάλαιο (Κεφάλαιο 6) της διπλωματικής εργασίας. Αναλύονται οι απαιτήσεις του συστήματος, η υποδομή του frontend και του backend server της εφαρμογής και τα χαρακτηριστικά τους. Ακολούθως, παρουσιάζεται η εφαρμογή, οι οθόνες και τα διάφορα σενάρια χρήσης από υπολογιστή ή από κινητό.

# Περιεχόμενα

Περίληψη	1
Abstract	2
Ευχαριστίες	3
Οργάνωση του τόμου	4
Κατάλογος Σχημάτων	8
Κατάλογος Πινάκων	11
<b>I Σχεδιασμός και Επιλογή Μοντέλων Βαθιάς Μάθησης</b>	<b>13</b>
<b>1 Εισαγωγή</b>	<b>14</b>
1.1 Τεχνητή Νοημοσύνη . . . . .	15
1.1.1 Ορισμός . . . . .	15
1.1.2 Μέθοδοι και Τεχνικές . . . . .	16
1.2 Μηχανική Μάθηση . . . . .	17
1.2.1 Ορισμός . . . . .	18
1.2.2 Μοντέλο Μηχανικής Μάθησης . . . . .	19
1.2.3 Είδη Μηχανικής Μάθησης . . . . .	20
1.3 Βαθιά Μηχανική Μάθηση . . . . .	21
1.4 Νευρωνικά Δίκτυα . . . . .	23
1.4.1 Ο Ανθρώπινος Εγκέφαλος . . . . .	23
1.4.2 Τεχνητά Νευρωνικά Δίκτυα . . . . .	25

1.5	Εισαγωγή στην Επεξεργασία Φυσικής Γλώσσας . . . . .	27
1.5.1	Δυσκολίες στην Κατανόηση Φυσικής Γλώσσας . . . . .	28
1.5.2	Εφαρμογές Επεξεργασίας Φυσικής Γλώσσας . . . . .	29
<b>2</b>	<b>Θεωρητικό Υπόβαθρο</b>	<b>33</b>
2.1	Διαδικασίες Μάθησης . . . . .	34
2.1.1	Επιβλεπόμενη Μάθηση . . . . .	34
2.1.2	Μη Επιβλεπόμενη Μάθηση . . . . .	36
2.2	Νευρωνικά Δίκτυα . . . . .	37
2.2.1	Μοντέλο Νευρώνα . . . . .	37
2.2.2	Πολυεπίπεδα Νευρωνικά Δίκτυα . . . . .	46
2.2.3	Εκπαίδευση . . . . .	48
2.2.4	Συνελικτικά Νευρωνικά Δίκτυα . . . . .	59
2.2.5	Αναδρομικά Νευρωνικά Δίκτυα . . . . .	64
<b>3</b>	<b>Παρουσίαση, Επεξεργασία και Ανάλυση Δεδομένων</b>	<b>67</b>
3.1	Βάση δεδομένων Springer . . . . .	68
3.2	Κατηγοριοποίηση Δεδομένων . . . . .	68
3.3	Ανάλυση και Προεπεξεργασία Δεδομένων . . . . .	72
3.3.1	Χαρακτηριστικά Δεδομένων . . . . .	72
3.3.2	Προεπεξεργασία Δεδομένων . . . . .	73
3.3.3	Word Embeddings . . . . .	77
3.3.4	Οπτικοποιήσεις Δεδομένων . . . . .	79
<b>4</b>	<b>Εφαρμογές Βαθιάς Μάθησης και Κατασκευή Μοντέλων</b>	<b>82</b>
4.1	Σχεδιασμός Απλών Μοντέλων . . . . .	83
4.1.1	Προετοιμασία Δεδομένων . . . . .	83
4.1.2	Δοκιμή Απλών Μοντέλων . . . . .	84
4.2	Δοκιμές και Πειραματικά Αποτελέσματα Σύνθετων Μοντέλων . . . . .	86
4.3	Επιλογή Τελικού Μοντέλου . . . . .	95
4.4	Οπτικοποιήσεις Δεδομένων . . . . .	101
4.5	Δημιουργία Clusters . . . . .	103
4.5.1	Επίπεδα και Αλγόριθμοι Clustering . . . . .	103
4.5.2	Δοκιμές Μοντέλων και Αποτελέσματα . . . . .	106
4.5.3	Σχεδιασμός Τελικού Μοντέλου και Παρουσίαση των Clusters . . . . .	114

---

4.6	Δημιουργία Συστήματος Εύρεσης Συναφών Συγγραμμάτων . . . . .	117
<b>II</b>	<b>Υλοποίηση Web Application</b>	<b>122</b>
<b>5</b>	<b>Θεωρητικό Υπόβαθρο</b>	<b>123</b>
5.1	Αλληλεπίδραση Ανθρώπου-Υπολογιστή . . . . .	124
5.2	React Framework . . . . .	124
5.3	Βάση Δεδομένων MongoDB . . . . .	125
<b>6</b>	<b>Σχεδιασμός και Υλοποίηση Συστήματος</b>	<b>126</b>
6.1	Απαιτήσεις Συστήματος . . . . .	127
6.1.1	Λειτουργικές Απαιτήσεις Συστήματος . . . . .	127
6.1.2	Μη Λειτουργικές Απαιτήσεις Συστήματος . . . . .	128
6.2	Αρχιτεκτονική Εφαρμογής . . . . .	128
6.2.1	Υποδομή Client (Frontend) . . . . .	129
6.2.2	Υποδομή Server και Βάσης Δεδομένων (Backend) . . . . .	130
6.3	Χαρακτηριστικά και Λειτουργικότητα Οθονών Εφαρμογής . . . . .	132
6.3.1	Αρχική Σελίδα . . . . .	132
6.3.2	Προβολή πολλαπλών βιβλίων . . . . .	134
6.3.3	Αναλυτική προβολή βιβλίου . . . . .	136
6.3.4	Ταξινόμηση νέου βιβλίου . . . . .	140
	<b>Συμπεράσματα</b>	<b>145</b>
	<b>Βιβλιογραφία</b>	<b>147</b>
	<b>Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια</b>	<b>150</b>
	<b>Απόδοση ξενόγλωσσων όρων</b>	<b>151</b>



# Κατάλογος Σχημάτων

1.1	Δέντρο αποφάσεων για αλγόριθμο επιλογής γεύματος . . . . .	16
1.2	Γενικός τρόπος λειτουργίας των ML αλγορίθμων . . . . .	20
1.4	Αναπαράσταση βιολογικού νευρώνα [1] . . . . .	23
1.8	Παράδειγμα autocomplete σε αναζήτηση στο Google . . . . .	30
1.9	Παράδειγμα του chatbot της Vodafone . . . . .	31
2.1	Δείγματα εικόνων με τις ετικέτες τους . . . . .	34
2.3	Σχηματικό διάγραμμα Μη Επιβλεπόμενης Μάθησης . . . . .	36
2.4	Μοντέλο νευρώνα . . . . .	38
2.5	Αφινικός μετασχηματισμός λόγω ύπαρξης πόλωσης . . . . .	40
2.6	Γραμμική συνάρτηση . . . . .	41
2.7	Συνάρτηση κατωφλίου . . . . .	42
2.8	Συνάρτηση προσήμου . . . . .	43
2.9	Σιγμοειδής συνάρτηση . . . . .	44
2.10	Συνάρτηση relu . . . . .	45
2.11	Παράδειγμα εφαρμογής της συνάρτησης softmax . . . . .	46
2.15	Σχηματική απεικόνιση συνόλων των εννοιών True Positive, True Negative, False Positive και False Negative. Με πράσινο χρώμα είναι οι σωστές προβλέψεις και με κόκκινο οι λανθασμένες. . . . .	56
2.16	Νευρωνικό Δίκτυο πριν και μετά την εφαρμογή dropout. Αριστερά: Ένα κλασικό μοντέλο Νευρωνικού Δικτύου. Δεξιά: Το μοντέλο που προκύπτει εφαρμόζοντας τη μέθοδο της απόσυρσης στο αριστερό Νευρωνικό Δίκτυο. [2] . . . . .	58
2.17	Δημιουργία νέων εικόνων από την αρχική με τη μέθοδο data augmentation [3] . . . . .	59
2.18	Παράδειγμα χειρόγραφων ψηφίων του MNIST dataset . . . . .	59

2.19	Παράδειγμα τοπικών χαρακτηριστικών στην εικόνα της ενότητας 1.3. Τα πράσινα εικονοστοιχεία είναι τα κοντινά, τα οποία βλέπουμε ότι αφορούν όλα το ίδιο στοιχείο στην εικόνα, το αυτοκίνητο. Αντίθετα, τα κίτρινα δεν έχουν καμία εννοιολογική σχέση μεταξύ τους και αφορούν διαφορετικά στοιχεία της εικόνας. [4] . . . . .	61
2.20	Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5] . . . . .	61
2.21	Λειτουργία του συνελικτικού επιπέδου. Επάνω φαίνεται ένα φίλτρο το οποίο προχωράει στην εικόνα για να παραγάγει μια τιμή της εξόδου. Κάτω φαίνεται αναλυτικά ένα αριθμητικό παράδειγμα ενός στοιχείου-προς-στοιχείο πολλαπλασιασμού μεταξύ της αρχικής εικόνας και των τιμών του συνελικτικού φίλτρου. [6] . . . . .	62
2.24	Παράδειγμα συνελικτικού δικτύου μιας διάστασης. [7] . . . . .	64
2.25	Απεικόνιση ενός κόμβου από ένα RNN. Από κάτω προς τα πάνω: επίπεδο εισόδου - κρυφό επίπεδο - επίπεδο εξόδου και $\{U, V, W\}$ είναι τα βάρη του δικτύου. Αριστερά φαίνεται ο κόμβος και δεξιά το ανάπτυγμά του. [8] . . . . .	65
3.1	Αρχική σελίδα του ηλεκτρονικού αποθετηρίου Springer. [9] . . . . .	68
3.2	Συχνότητες εμφάνισης δειγμάτων για τις 26 κλάσεις. . . . .	71
3.3	Αριθμός λέξεων στον πίνακα περιεχομένων του κάθε βιβλίου. . . . .	73
3.4	Παράδειγμα αναπαράστασης διανυσμάτων λέξεων σε δισδιάστατο χώρο. . . . .	77
3.5	Συχνότητες εμφάνισης δειγμάτων για τις 26 κλάσεις. . . . .	79
3.6	Οπτικοποίηση όλων των βιβλίων στον δισδιάστατο χώρο. . . . .	80
4.1	Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 3 με Αλγόριθμο Βελτιστοποίησης Adam . . . . .	87
4.2	Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 3 με Αλγόριθμο Βελτιστοποίησης RMSProp . . . . .	88
4.3	Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 4 . . . . .	89
4.4	Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 5 . . . . .	91
4.5	Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 6 . . . . .	93
4.6	Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 7 . . . . .	94
4.7	Διαγράμματα μετρικών εκπαίδευσης Τελικού Μοντέλου 26 κλάσεων . . . . .	97
4.9	Οπτικοποίηση δεδομένων μετά την έξοδο τους από το επίπεδο dense 1. . . . .	101

4.12	Ομαδοποίηση δεδομένων με τους 3 clustering αλγορίθμους και οπτικοποίηση των clusters [10] . . . . .	105
4.13	Αποτελέσματα της μετρικής silhouette για κάθε αριθμό clusters . . . . .	106
4.14	Αποτελέσματα της μετρικής silhouette για κάθε αριθμό clusters . . . . .	107
4.15	Μέση απόδοση silhouette για κάθε αριθμό pca components . . . . .	108
4.16	Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5] . . . . .	109
4.17	Μέση απόδοση silhouette για κάθε αριθμό pca components . . . . .	109
4.18	Μέση απόδοση silhouette για κάθε αριθμό threshold . . . . .	110
4.19	Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5] . . . . .	111
4.20	Μέση απόδοση silhouette για κάθε αριθμό pca components . . . . .	111
4.21	Μέση απόδοση silhouette για κάθε αριθμό threshold . . . . .	112
4.24	Μέση απόδοση silhouette για κάθε αριθμό threshold . . . . .	114
4.25	Οπτικοποίηση των 5 clusters . . . . .	117
6.1	Αρχική σελίδα εφαρμογής από Desktop. . . . .	133
6.2	Αρχική σελίδα εφαρμογής από Tablet. . . . .	134
6.3	Αρχική σελίδα εφαρμογής από Smartphone. . . . .	134
6.5	Σελίδα προβολής πολλαπλών βιβλίων από Smartphone. . . . .	136
6.6	Αναλυτική προβολή βιβλίου (βασικά στοιχεία) από Desktop. . . . .	137
6.7	Αναλυτική προβολή βιβλίου (βασικά στοιχεία) από Smartphone. . . . .	137
6.8	Αναλυτική προβολή βιβλίου (πίνακα περιεχομένων) από Desktop. . . . .	138
6.9	Αναλυτική προβολή βιβλίου (πίνακα περιεχομένων) από Smartphone. . . . .	138
6.10	Αναλυτική προβολή βιβλίου (περίληψη και παρόμοια βιβλία) από Desktop. . . . .	139
6.11	Αναλυτική προβολή βιβλίου (περίληψη και παρόμοια βιβλία) από Smartphone. . . . .	140
6.12	Σελίδα ταξινόμησης νέου βιβλίου από Desktop και Smartphone. . . . .	141
6.13	Εισαγωγή πίνακα περιεχομένων νέου βιβλίου από Desktop και Smartphone. . . . .	142
6.14	Αποτέλεσμα ταξινόμησης νέου βιβλίου από Desktop και Smartphone. . . . .	143

# Κατάλογος Πινάκων

3.1	Αριθμός δειγμάτων ανά κατηγορία. . . . .	70
3.2	Αριθμός δειγμάτων ανά κατηγορία για τις 7 επιλεγμένες κατηγορίες. . .	72
3.3	Αποτελέσματα όμοιων λέξεων της λέξης “car”. . . . .	78
4.1	Classification Μοντέλο 1 . . . . .	84
4.2	Αποτελέσματα Μοντέλου 1 . . . . .	84
4.3	Classification Μοντέλο 2 . . . . .	85
4.4	Αποτελέσματα Μοντέλου 2 . . . . .	85
4.5	Classification Μοντέλο 3 . . . . .	86
4.6	Αποτελέσματα Μοντέλου 3 . . . . .	87
4.7	Classification Μοντέλο 4 . . . . .	88
4.8	Αποτελέσματα Μοντέλου 4 . . . . .	89
4.9	Classification Μοντέλο 5 . . . . .	90
4.10	Αποτελέσματα Μοντέλου 5 . . . . .	90
4.11	Classification Μοντέλο 6 . . . . .	92
4.12	Αποτελέσματα Μοντέλου 6 . . . . .	92
4.13	Classification Μοντέλο 7 . . . . .	93
4.14	Αποτελέσματα Μοντέλου 7 . . . . .	95
4.15	Τελικό Classification Μοντέλο 26 κλάσεων . . . . .	95
4.16	Αποτελέσματα Τελικού Μοντέλου 26 κλάσεων . . . . .	96
4.17	Τελικό Classification Μοντέλο 7 κλάσεων . . . . .	99
4.18	Αποτελέσματα Τελικού Μοντέλου 6 κλάσεων . . . . .	99
4.19	Clustering Μοντέλο 1 . . . . .	106
4.20	Clustering Μοντέλο 2 . . . . .	107
4.21	Clustering Μοντέλο 3 . . . . .	108
4.22	Clustering Μοντέλο 4 . . . . .	110

---

4.23 Clustering Μοντέλο 5 . . . . .	112
4.24 Περιεχόμενα του Cluster 1 . . . . .	115
4.25 Περιεχόμενα του Cluster 2 . . . . .	115
4.26 Περιεχόμενα του Cluster 3 . . . . .	116
4.27 Περιεχόμενα του Cluster 4 . . . . .	116
4.28 Περιεχόμενα του Cluster 5 . . . . .	116
4.29 Παράδειγμα 1 εύρεσης όμοιων βιβλίων . . . . .	118
4.30 Παράδειγμα 2 εύρεσης όμοιων βιβλίων . . . . .	119

# Μέρος Ι

## Σχεδιασμός και Επιλογή Μοντέλων Βαθιάς Μάθησης

# Κεφάλαιο 1

## Εισαγωγή

Σε αυτό το κεφάλαιο θα γίνει μια εισαγωγή στις διάφορες έννοιες και στο αντικείμενο της διπλωματικής αυτής εργασίας. Θα δοθούν ορισμοί και περιγραφές για την Τεχνητή Νοημοσύνη, την Μηχανική Μάθηση, τη Βαθιά Μάθηση και την επεξεργασία φυσικής γλώσσας με τα οποία ασχολείται κυρίως το Μέρος I της διπλωματικής. Θα επεξηγηθούν βασικές λειτουργίες του εγκεφάλου οι οποίες σχετίζονται με τη λογική των νευρωνικών δικτύων, τα οποία χρησιμοποιούνται για την κατασκευή των μοντέλων στα επόμενα κεφάλαια.

## 1.1 Τεχνητή Νοημοσύνη

Ο άνθρωπος συνεχώς εξελίσσεται, σκέφτεται, μαθαίνει, ερευνά τον κόσμο γύρω του και έχει την τάση να δημιουργεί και να ανακαλύπτει. Τι είναι, όμως, αυτό που διαθέτει και του δίνει αυτές τις δυνατότητες; Μα, φυσικά η νοημοσύνη του. Η νοημοσύνη είναι το σύνολο των γνωστικών ικανοτήτων του ανθρώπου, δηλαδή η αντίληψη, η μνήμη, ο συνειρμός, η φαντασία, η προσοχή, η διανόηση και, ειδικότερα, η ικανότητα προσαρμογής σε νέες καταστάσεις και αντίληψης ομοιοτήτων, διαφορών και σχέσεων [11]. Η επιστήμη μελετά και προσπαθεί να κατανοήσει αυτούς τους μηχανισμούς που διαθέτει ο άνθρωπος και να απαντήσει στο ερώτημα για το πώς μπορεί να εκτελεί τέτοιες πολύπλοκες διαδικασίες τόσο απλά αλλά και τόσο στοχευμένα.

### 1.1.1 Ορισμός

Το πεδίο της Τεχνητής Νοημοσύνης (Artificial Intelligence), ή για συντομία TN, ουσιαστικά δεν προσπαθεί απλά να καταλάβει τους μηχανισμούς αυτούς, αλλά επιχειρεί να σχεδιάσει και να κατασκευάσει ευφυή μοντέλα και συστήματα, τα οποία συμπεριφέρονται σαν να έχουν την δική τους νοημοσύνη. Έχουν δοθεί πολλοί ορισμοί για την TN, ένας γενικός που περιλαμβάνει όσο το δυνατόν περισσότερα στοιχεία είναι ο εξής:

*“TN είναι ο τομέας της Επιστήμης των Υπολογιστών που ασχολείται με τη σχεδίαση και την υλοποίηση προγραμμάτων τα οποία είναι ικανά να μιμηθούν τις ανθρώπινες γνωστικές ικανότητες, εμφανίζοντας έτσι χαρακτηριστικά που αποδίδουμε συνήθως σε ανθρώπινη συμπεριφορά, όπως για παράδειγμα η επίλυση προβλημάτων, η αντίληψη μέσω της όρασης, η μάθηση, η εξαγωγή συμπερασμάτων, η κατανόηση φυσικής γλώσσας, κτλ.”*  
[1]

Τα τελευταία χρόνια διεξάγεται εντατικότερη έρευνα στο πεδίο της TN, καθώς έχει προσελκύσει πολλούς επιστήμονες και μηχανικούς του κλάδου της επιστήμης των υπολογιστών (computer science). Η τεχνολογία έχει αλλάξει κατά πολύ τις ζωές μας και θα συνεχίσει να τις αλλάζει όσο εξελίσσεται και αναπτύσσεται. Η TN είναι ένας σημαντικός κλάδος της τεχνολογίας που θα συμβάλει σε αυτό και σταδιακά εφαρμόζεται σε όλο και περισσότερες συσκευές, εφαρμογές, πλατφόρμες και υπηρεσίες. Μεταμορφώνει την καθημερινότητά μας, εξυπηρετεί ανάγκες και μας βοηθάει να παίρνουμε αποφάσεις·



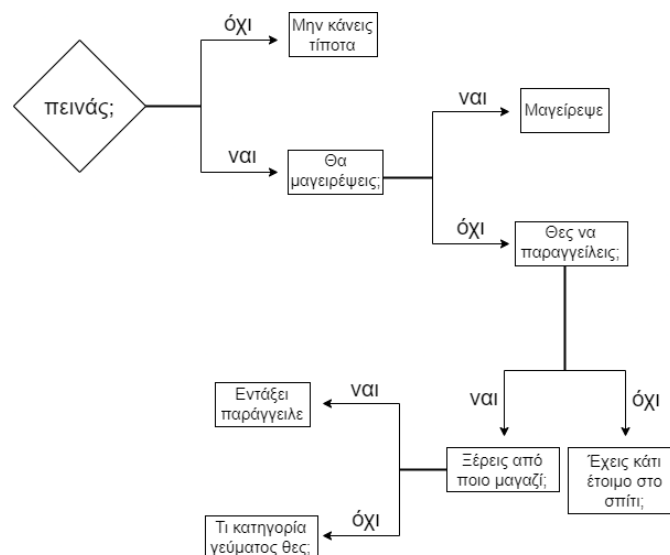
κάποιες φορές αντικαθιστά και τη σκέψη μας. Είναι δύσκολο να φανταστεί κανείς τι θα μπορούμε να κάνουμε στο μέλλον με την τεχνητή νοημοσύνη και τις εφαρμογές της.

### 1.1.2 Μέθοδοι και Τεχνικές

Η Τεχνητή Νοημοσύνη έχει αναπτυχθεί αρκετά και έχει δημιουργήσει νέα ερευνητικά πεδία και θέματα. Υπάρχουν πολλές μέθοδοι και τεχνικές για την εφαρμογή της. Κάποιες από αυτές αναλύονται παρακάτω:

- **Δέντρα αποφάσεων**

Οι πρώτοι ερευνητές προσπαθώντας να προσομοιώσουν την ανθρώπινη συλλογιστική, κατασκεύασαν κάποιους αλγορίθμους. Η λογική των αλγορίθμων αυτών είναι ο σχεδιασμός ενός δέντρου αποφάσεων με όλες τις πιθανές εκδοχές και καταστάσεις που είναι δυνατό να προκύψουν.



Σχήμα 1.1: Δέντρο αποφάσεων για αλγόριθμο επιλογής γεύματος

- **Κατασκευή μιας βάσης γνώσης**

Αυτή η μέθοδος περιλαμβάνει ορισμένα συστήματα τα οποία προσπαθούν να συγκεντρώσουν γνώση για ένα πεδίο χρησιμοποιώντας μια ή περισσότερες πηγές (π.χ. Η Αθήνα είναι η πρωτεύουσα της Ελλάδας, ο δήμος Περιστερίου βρίσκεται

στην Αθήνα). Όλες αυτές οι πληροφορίες αποθηκεύονται σε μια βάση γνώσης στην οποία ύστερα υπάρχει η δυνατότητα επεξεργασίας των δεδομένων που περιέχει και η κατασκευή ερωτημάτων (queries) τα οποία θα φέρουν χρήσιμα αποτελέσματα. Ουσιαστικά πρόκειται για μια αποθήκη γνώσεων στην οποία μπορούν να ανατρέχουν τα συστήματα και να αντλούν πληροφορίες προτού προβούν στην λήψη κάποιας απόφασης.

- **Μηχανική Μάθηση**

Η Μηχανική Μάθηση είναι ένα από τα πιο ενδιαφέροντα πεδία της ΤΝ με πολλές εφαρμογές. Ουσιαστικά είναι η δυνατότητα των μηχανών να μαθαίνουν από τα δεδομένα και εκτελώντας αντίστοιχους αλγορίθμους Μηχανικής Μάθησης, το σύστημα βελτιώνεται και γίνεται πιο αποτελεσματικό όσον αφορά τις προβλέψεις και τους υπολογισμούς του με κάθε εκτέλεση των αλγορίθμων γύρω από κάποια ομάδα δεδομένων. Η δυνατότητα αυτή των μηχανών να μπορούν να μαθαίνουν από τα δεδομένα έχει βοηθήσει σε μεγάλο βαθμό την ΤΝ να προχωρήσει και να εξελιχθεί. Περισσότερα για την μηχανική μάθηση θα αναλυθούν στην επόμενη ενότητα.

## 1.2 Μηχανική Μάθηση

Η Μηχανική Μάθηση (Machine Learning ή για συντομία ML) προσελκύει όλο και περισσότερους επιστήμονες να ασχοληθούν με αυτή, να την εφαρμόσουν σε έργα και εφαρμογές τους και να την αναπτύξουν. Γιατί όμως υπάρχει τόσο μεγάλη δραστηριότητα γύρω από αυτή; Το 2006, ο Clive Humby, μαθηματικός και επιχειρηματίας στον τομέα της επιστήμης των δεδομένων (data science), είπε την εξής φράση:

*“Data is the new oil.”* [12]

Πρόκειται για μια πολύ διαδεδομένη και γνωστή έκφραση που ουσιαστικά παρομοιάζει το πετρέλαιο, το ορυκτό αυτό καύσιμο το οποίο αποτελεί μια από τις σπουδαιότερες πηγές ενέργειας, με τα δεδομένα, λέγοντας ουσιαστικά ότι τα δεδομένα είναι το “νέο” πετρέλαιο. Ο Humby θέλει να δείξει ότι τα δεδομένα, όπως και το πετρέλαιο, είναι πολύτιμα μόνο αν αξιοποιηθούν σωστά, αν φιλτραριστούν και αν αναλυθούν με τον κατάλληλο τρόπο για να δώσουν χρήσιμα και ουσιώδη αποτελέσματα τα οποία θα αποτελέσουν βασικό κριτήριο στη λήψη αποφάσεων και στον σχεδιασμό στρατηγικής.

Με λίγα λόγια, η πληροφορία μπορεί να εξαχθεί από τα δεδομένα, όπως και η ενέργεια από το πετρέλαιο.

Αυτή η έκφραση, λοιπόν, ακούγεται συχνά και κάθε χρόνο βλέπουμε ότι τα δεδομένα γίνονται όλο και πιο πολύτιμα. Εταιρίες προσπαθούν να αποκτήσουν όσο το δυνατό περισσότερο δεδομένα που αφορούν τους πελάτες τους και επενδύουν όλο και περισσότερα στα τμήματα ανάλυσης δεδομένων των επιχειρήσεών τους. Επιστήμονες του data science μελετούν και ερευνούν όλο και πιο βαθιά τα δεδομένα και προσπαθούν να τα αξιοποιήσουν στο 100% χωρίς να χαθεί καμία πληροφορία. Νέες κατευθύνσεις, ερευνητικά προγράμματα αλλά και θέσεις εργασίας δημιουργούνται που έχουν σχέση με δεδομένα και όπως φαίνεται στο μέλλον θα δημιουργηθούν ακόμη περισσότερα. Είναι φανερό λοιπόν η χρησιμότητα της Μηχανικής Μάθησης στην επιστήμη αλλά και ο λόγος που υπάρχει μια αυξητική τάση από τον κόσμο να ασχοληθούν με αυτή.

### 1.2.1 Ορισμός

Η Μάθηση (Learning) είναι μια πολύ σημαντική διαδικασία που μπορεί να εκτελέσει ο άνθρωπος. Μέσω παρατηρήσεων από το περιβάλλον του έχει την ικανότητα να μάθει και να κατανοήσει έννοιες, να διακρίνει χαρακτηριστικά τους και να οργανώσει τη σκέψη του. Για παράδειγμα κανένας δεν μας έχει δώσει τον ορισμό του σκύλου αλλά μπορούμε να τον διακρίνουμε και να τον ξεχωρίσουμε από άλλα ζώα. Αυτό συμβαίνει γιατί έχουμε δει τόσες εικόνες από σκύλους στην καθημερινότητά μας έτσι ώστε το μυαλό μας έχει σχηματίσει ένα εμπειρικό μοντέλο για τον σκύλο και τα χαρακτηριστικά του. Η ΤΝ μπορεί να δημιουργήσει υπολογιστικά συστήματα με αυτή τη δυνατότητα, ικανά δηλαδή να μάθουν από δεδομένα· αυτή είναι η Μηχανική Μάθηση.

Υπάρχουν πολλοί ορισμοί για την Μηχανική Μάθηση. Κάποιοι από αυτούς είναι οι εξής:

- *“Το φαινόμενο κατά το οποίο ένα σύστημα βελτιώνει την απόδοσή του κατά την εκτέλεση μιας συγκεκριμένης εργασίας, χωρίς να υπάρχει ανάγκη να προγραμματιστεί εκ νέου.”* [13]
- Mitchell (1997)

*“Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία  $E$  ως προς κάποια κλάση εργασιών  $T$  και μέτρο απόδοσης  $P$ , αν η απόδοσή*

του σε εργασίες από το  $T$ , όπως μετρίεται από το  $P$ , βελτιώνεται μέσω της εμπειρίας  $E$ .”

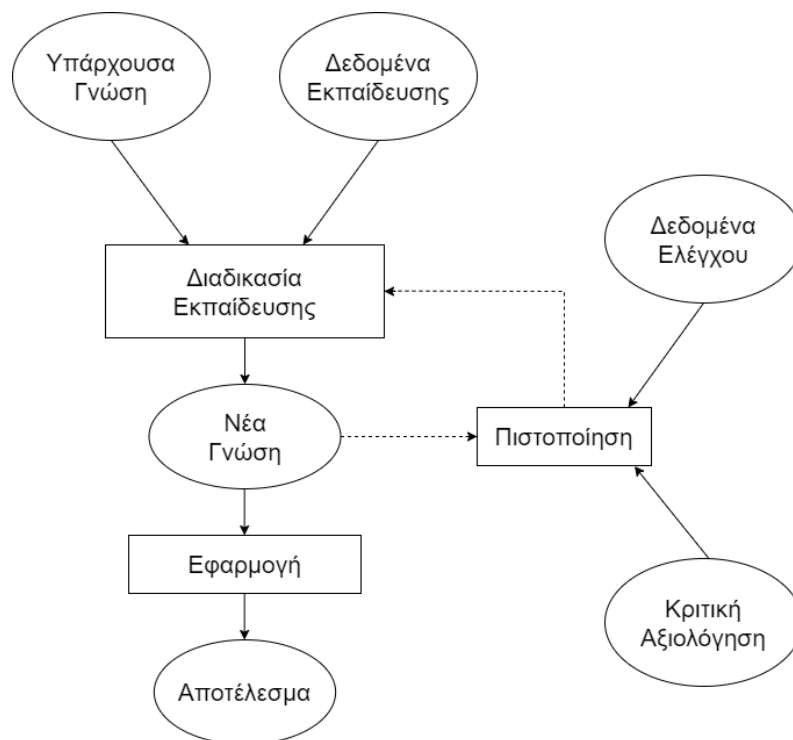
Ο δεύτερος ορισμός του Mitchell χρησιμοποιεί τέσσερις συνιστώσες για να περιγράψει τη Μηχανική Μάθηση.

- Το πρόγραμμα, το οποίο μαθαίνει και είναι ο αλγόριθμος ή η συνάρτηση που εκπαιδεύεται
- Η εμπειρία  $E$ , η οποία είναι το σύνολο των δεδομένων εκπαίδευσης
- Το έργο  $T$ , το οποίο είναι το πρόβλημα
- Το μέτρο απόδοσης  $P$ , το οποίο είναι μια συνάρτηση εκτίμησης της ακρίβειας και της επιτυχίας του προγράμματός μας στο συγκεκριμένο πρόβλημα

Θέλοντας να συνδυάσουμε όλα τα παραπάνω, ένας γενικός ορισμός της Μηχανικής Μάθησης είναι όλες οι τεχνικές και οι διαδικασίες με τις οποίες τα υπολογιστικά συστήματα, με την βοήθεια ειδικών αλγορίθμων, μπορούν να μαθαίνουν από τα δεδομένα και να βελτιώνονται, να κάνουν προβλέψεις σχετικές με αυτά, να εξάγουν συμπεράσματα και να αποκτούν εμπειρία.

### 1.2.2 Μοντέλο Μηχανικής Μάθησης

Για κάθε πρόβλημα στον χώρο της Μηχανικής Μάθησης υπάρχει και ένας κατάλληλος τρόπος και μέθοδος εκμάθησης. Η διαδικασία αυτή χρειάζεται έναν αλγόριθμο, τον αλγόριθμο εκμάθησης. Αυτός θα εκτελείται επαναλαμβανόμενα με είσοδο τα δεδομένα του προβλήματος λαμβάνοντας επιπλέον υπόψη και άλλες μετρικές απόδοσης, παρατηρήσεις του συστήματος ή την προϋπάρχουσα γνώση που διαθέτει. Ο γενικός τρόπος λειτουργίας τέτοιων αλγορίθμων φαίνεται στο παρακάτω σχήμα:



Σχήμα 1.2: Γενικός τρόπος λειτουργίας των ML αλγορίθμων

Στο Σχήμα 1.2 φαίνεται η γενική διαδικασία εκμάθησης όπου ο αλγόριθμος λαμβάνει ως είσοδο την υπάρχουσα γνώση, την οποία διαθέτει από προηγούμενες εκτελέσεις του, και τα δεδομένα εκπαίδευσης (training dataset) προκειμένου να δημιουργήσει νέα γνώση. Με την βοήθεια κάποιων δεδομένων ελέγχου (test dataset) και με μια διαδικασία ανάκλησης (recall), ελέγχει αν υπήρξε κάποια βελτίωση στο σύστημα και αν “έμαθε” από τα δεδομένα εκπαίδευσης. Η παραπάνω μεθοδολογία επαναλαμβάνεται μέχρι η πιστοποίηση να δείξει τις επιθυμητές μετρικές και έτσι το σύστημα να είναι έτοιμο να εφαρμοστεί σε κάποια εφαρμογή ή υπηρεσία.

### 1.2.3 Είδη Μηχανικής Μάθησης

Οι μέθοδοι και οι τεχνικές της Μηχανικής Μάθησης είναι ποικίλες. Κάποιες είναι απλές στην εφαρμογή τους ενώ άλλες είναι περίπλοκες και απαιτούν γνώσεις για την κατανόηση και εφαρμογή τους. Ανάλογα όμως με την φύση του προβλήματος και των δεδομένων, υπάρχουν τρία είδη της Μηχανικής Μάθησης, τα οποία αναφέρονται παρακάτω:

- **Επιβλεπόμενη Μάθηση (Supervised Learning)**

Στην Επιβλεπόμενη Μάθηση το σύστημα καλείται να “μάθει” μια έννοια ή ένα μοντέλο από κάποια δεδομένα τα οποία το περιγράφουν. Ονομάζεται έτσι διότι υπάρχει κάποιος “επιβλέπων” ο οποίος παρέχει την επιθυμητή τιμή εξόδου της συνάρτησης. Δύο βασικές κατηγορίες προβλημάτων με τα οποία ασχολείται η Επιβλεπόμενη Μάθηση είναι προβλήματα:

- Ταξινόμησης (Classification)
- Πρόβλεψης (Prediction)

- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**

Εδώ το μοντέλο προσπαθεί να “ανακαλύψει” μόνο του πόσα και ποια είναι τα μοντέλα και οι ομάδες που ανήκουν τα δεδομένα καθώς και να βρει τα κοινά τους χαρακτηριστικά, χωρίς να γνωρίζει εξαρχής τη σωστή τιμή εξόδου. Δύο βασικές κατηγορίες προβλημάτων με τα οποία ασχολείται η Μη Επιβλεπόμενη Μάθηση είναι προβλήματα:

- Ομαδοποίησης (Clustering)
- Ανάλυσης Συσχετισμών (Association Analysis)

- **Ενισχυτική Μάθηση (Reinforcement Learning)**

Στην Ενισχυτική Μάθηση, το σύστημα εκπαιδεύεται αλληλεπιδρώντας άμεσα με το περιβάλλον του, προσδοκώντας να έχει βέλτιστη συμπεριφορά. Εφαρμόζεται κυρίως σε:

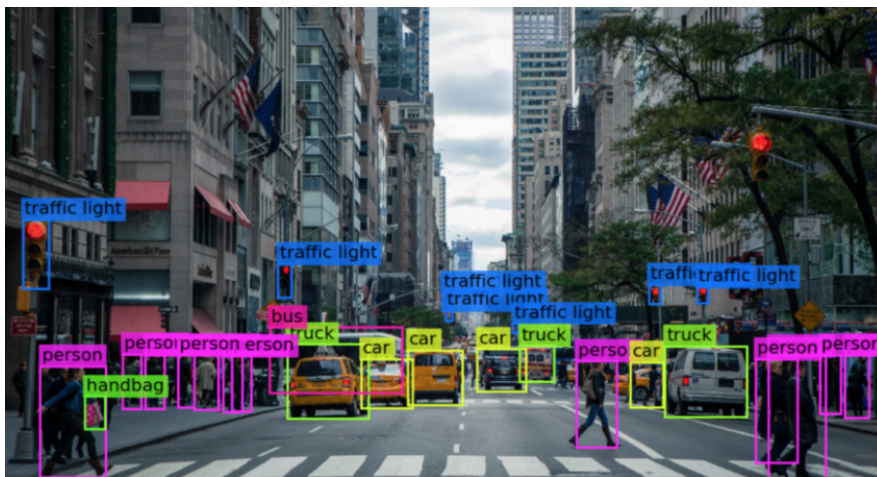
- Έλεγχο κίνησης ρομπότ
- Εκμάθηση παιχνιδιών (π.χ. σκάκι)
- Βελτιστοποίηση αυτοματοποιημένων μηχανημάτων στα εργοστάσια

### 1.3 Βαθιά Μηχανική Μάθηση

Η Βαθιά Μάθηση αποτελεί εξέλιξη και υποκατηγορία του κλάδου της Μηχανικής Μάθησης. Με την χρήση πολλαπλών επιπέδων στις αρχιτεκτονικές των μοντέλων

και πιο πολύπλοκων αλγορίθμων, οι τεχνικές της Βαθιάς Μάθησης είναι πολύ αποδοτικές και μπορούν να εξάγουν περισσότερα και υψηλότερου επιπέδου χαρακτηριστικά από τα δεδομένα εισόδου.

Λόγω της δυνατότητας αυτής, οι αλγόριθμοι Βαθιάς Μάθησης μπορούν να μαθαίνουν από σύνθετα δεδομένα με πολλές συνιστώσες. Έτσι, είναι δυνατό, αν χρησιμοποιηθούν ως δεδομένα εικόνες, βίντεο ή κείμενα κάποιας γλώσσας, τα μοντέλα αυτά να είναι ικανά, με τον κατάλληλο σχεδιασμό, να τα επεξεργαστούν και δημιουργήσουν μια γνώση για αυτά. Αυτό έχει φέρει τεράστια πρόοδο στην Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing), βλ. ενότητα 1.5, στα αυτόνομα αυτοκίνητα (self-driving cars), στην Όραση Υπολογιστών (Computer Vision) και σε πολλούς άλλους τομείς.



Σχήμα 1.3: Κάμερα αυτόνομου αυτοκινήτου [4]

Βασικό πλεονέκτημα, λοιπόν, της Βαθιάς Μάθησης είναι η ικανότητά της να εξάγει χαρακτηριστικά από περίπλοκα δεδομένα τα οποία είναι δύσκολο να κάνουμε με κάποιον άλλο τρόπο. Επιπλέον, όσο μεγαλύτερος είναι ο όγκος δεδομένων που διαθέτουμε, υπό κάποιες προϋποθέσεις, τόσο καλύτερη θα είναι η εκπαίδευση του μοντέλου Βαθιάς Μάθησης και τόσο καλύτερα αποτελέσματα θα υπολογίζει. Έτσι, για άλλη μια φορά, επιβεβαιώνεται η σημαντικότητα των δεδομένων αλλά και πόσο πολύτιμα είναι στην ΤΝ και τις εφαρμογές της.

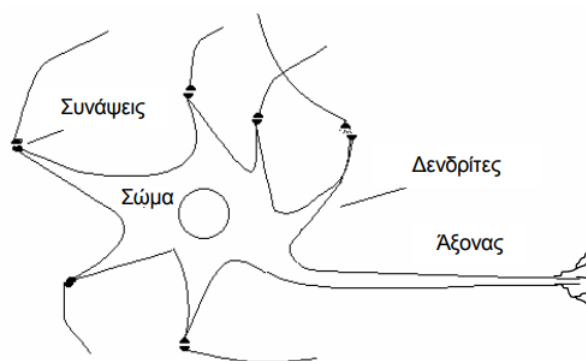
## 1.4 Νευρωνικά Δίκτυα

Η βασική ιδέα της Τεχνητής Νοημοσύνης είναι τα υπολογιστικά συστήματα να “έχουν” τη δική τους νοημοσύνη, να δρουν και να παίρνουν αποφάσεις όπως οι άνθρωποι. Πολλά συστήματα, λοιπόν, έχουν σχεδιαστεί και αναπτυχθεί με βάση λειτουργίες και χαρακτηριστικά των ανθρώπων. Ιδιαίτερη σημασία έχει δοθεί στον ανθρώπινο εγκέφαλο και στις πολύπλοκες λειτουργίες τις οποίες εκτελεί. Εμπνευσμένα λοιπόν από τον ανθρώπινο εγκέφαλο, σχεδιάστηκαν τα Νευρωνικά Δίκτυα (Neural Networks) ή πιο απλά NN. Τι είναι όμως ένα NN και πώς η δομή του εγκεφάλου έχει βοηθήσει στη δημιουργία του;

### 1.4.1 Ο Ανθρώπινος Εγκέφαλος

Το έργο στο επιστημονικό πεδίο των τεχνητών νευρωνικών δικτύων βασίστηκε στο γεγονός ότι ο ανθρώπινος εγκέφαλος εκτελεί τις λειτουργίες και τους υπολογισμούς του με διαφορετικό τρόπο από ένα τυπικό υπολογιστικό σύστημα.

*Ο εγκέφαλος είναι ένας εξαιρετικά πολύπλοκος, μη γραμμικός, παράλληλος υπολογιστής, ο οποίος έχει τη δυνατότητα να οργανώνει τα δομικά του στοιχεία, γνωστά ως νευρώνες (neurons), με τρόπο ώστε να εκτελούν συγκεκριμένους υπολογισμούς (π.χ. αναγνώριση προτύπων, αντίληψη, έλεγχο της κίνησης) με πολλαπλάσια ταχύτητα από αυτή του πιο γρήγορου υπολογιστή σήμερα. [14]*

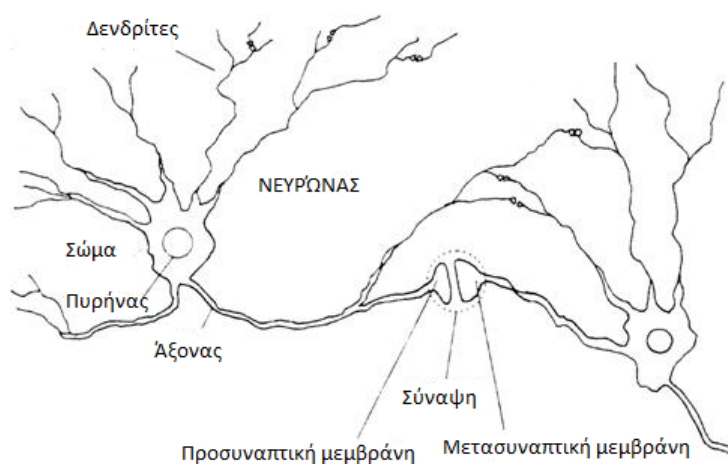


Σχήμα 1.4: Αναπαράσταση βιολογικού νευρώνα [1]



Τα νευρικά κύτταρα λειτουργούν με ηλεκτρικά σήματα και ταυτόχρονα ελκύουν βιοχημικές ουσίες για να επικοινωνούν με άλλα κύτταρα [15]. Ένα τυπικό νευρικό κύτταρο, ή νευρώνας, (Σχήμα 1.4) αποτελείται από το σώμα (body), το οποίο αποτελεί τον πυρήνα του, τους δενδρίτες (dendrites), οι οποίοι είναι υπεύθυνοι για την λήψη σημάτων από γειτονικούς νευρώνες, τον άξονα (axon), που είναι η έξοδος του νευρώνα και το μέσο διασύνδεσής του με άλλους νευρώνες. Η εξειδικευμένη επαφή μεταξύ δύο νευρώνων, όπου ο ένας επηρεάζει την λειτουργία του άλλου, ονομάζεται σύναψη (synapse).

Ο ανθρώπινος εγκέφαλος είναι αποτέλεσμα βιολογικής εξέλιξης εκατομμυρίων ετών και, δίχως αμφιβολία, αποτελεί ένα από τα μεγαλύτερα δημιουργήματα της φύσης. Αποτελείται από περίπου  $10^{10}$  νευρώνες, οι οποίοι συνδέονται μεταξύ τους με συνολικά  $10^{14}$  έως  $10^{15}$  σημεία. Αυτή η μικρή μάζα λοιπόν που ζυγίζει μόλις 1,5 κιλό, μέσω αυτού του εξαιρετικά πολυσύνθετου δικτύου κυττάρων, είναι υπεύθυνη για τη δημιουργία της σκέψης, τη μνήμη, τα συναισθήματα, τον έλεγχο των κινήσεων του σώματος και για άλλες λειτουργίες. Παρόλη την πολυπλοκότητα αυτών των διαδικασιών, ο ανθρώπινος εγκέφαλος φαίνεται να τις εκτελεί σε λίγα χιλιοστά του δευτερολέπτου. Αυτό συμβαίνει, διότι, επεξεργάζεται την πληροφορία κατανεμημένα και παράλληλα μέσα στο βιολογικό νευρωνικό σύστημά του.



Σχήμα 1.5: Επικοινωνία μεταξύ νευρώνων [15]

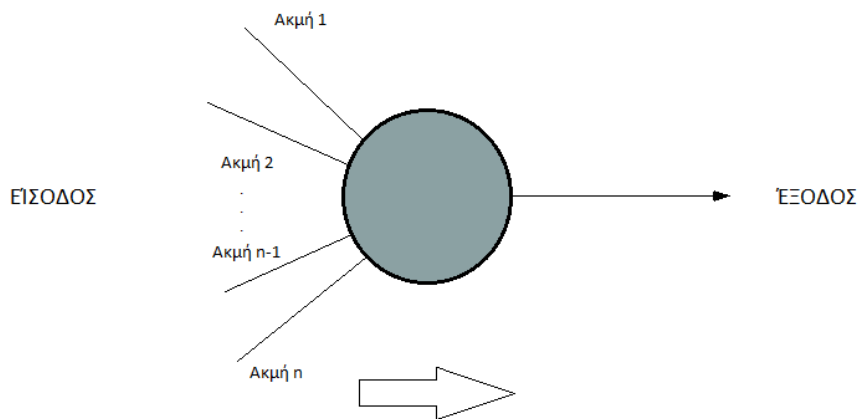
Οι νευρώνες επικοινωνούν μεταξύ τους και ανταλλάσσουν σήματα μέσω των συ-

νάψεων. Οι άξονες του κάθε νευρώνα καταλήγουν στην προσυναπτική μεμβράνη, η οποία συμμετέχει στον σχηματισμό της σύναψης και καταλήγει στην μετασυναπτική μεμβράνη του άλλου νευρώνα η οποία συλλέγει τα σήματα που εκπέμπονται σε αυτόν και τα μεταφέρει στο σώμα του.

### 1.4.2 Τεχνητά Νευρωνικά Δίκτυα

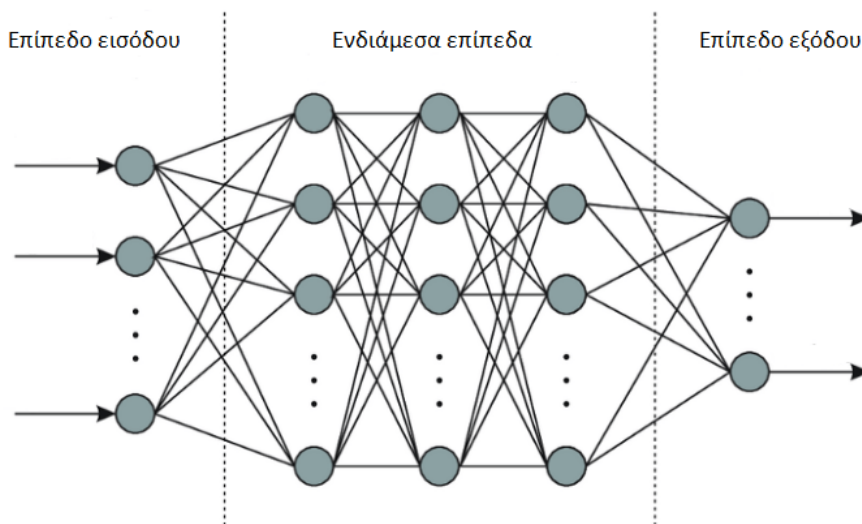
Η αρχιτεκτονική και ο βασικός τρόπος λειτουργίας λοιπόν του εγκεφάλου, ενέπνευσαν τους επιστήμονες να δημιουργήσουν μοντέλα τα οποία λειτουργούν και επεξεργάζονται την πληροφορία με παρόμοιο τρόπο και τεχνικές. Τα μοντέλα αυτά ονομάζονται Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks), για συντομία στη συνέχεια θα αναφερόμαστε σε αυτά ως Νευρωνικά Δίκτυα, και είναι βασικό εργαλείο αυτής της διπλωματικής εργασίας.

Ο τεχνητός νευρώνας (artificial neuron) είναι η βασική υπολογιστική μονάδα που περιέχει ένα Νευρωνικό Δίκτυο. Τα μέρη του αντιστοιχούν άμεσα στα μέρη ενός βιολογικού νευρώνα του εγκεφάλου. Ο νευρώνας δέχεται κάποια σήματα εισόδου τα οποία αντιστοιχούν σε συνεχείς μεταβλητές, σε αντίθεση με τους ηλεκτρικούς παλμούς τους οποίους δέχεται ως είσοδο ο βιολογικός νευρώνας. Κάθε τέτοια είσοδος συνοδεύεται από μια ακμή, ο ρόλος της οποίας αντιστοιχεί σε μια σύναψη στα βιολογικά νευρωνικά συστήματα. Η ακμή αυτή περιέχει μια τιμή βάρους (weight). Αφού συλλέξει όλες τις εισόδους και τα βάρη των ακμών, ο νευρώνας εκτελεί κάποιους υπολογισμούς και παράγει την έξοδό του, με την οποία τροφοδοτεί τους γειτονικούς του νευρώνες πάλι μέσω των αντίστοιχων ακμών που τους συνδέουν με αυτόν. Με αυτή τη διαδικασία λοιπόν, οι νευρώνες ανταλλάσσουν πληροφορίες μεταξύ τους.



Σχήμα 1.6: Μοντέλο τεχνητού νευρώνα

Το σύνολο των νευρώνων, η δομή τους και φυσικά το σύνολο των ακμών και των βαρών τους αποτελούν ένα Νευρωνικό Δίκτυο. Συνήθως οι νευρώνες είναι οργανωμένοι σε μια σειρά από στρώματα ή επίπεδα (layers). Το πρώτο επίπεδο είναι το επίπεδο εισόδου (input layer) του νευρωνικού δικτύου και το τελευταίο το επίπεδο εξόδου (output layer). Ανάμεσα σε αυτά τα δύο επίπεδα μπορεί να ακολουθούν, προαιρετικά, ένα ή περισσότερα ενδιάμεσα ή κρυφά επίπεδα (hidden layers).



Σχήμα 1.7: Μοντέλο τεχνητού νευρωνικού δικτύου [16]

Τα δίκτυα αυτά παίρνουν γνώσεις, δηλαδή μαθαίνουν, με την εξάσκηση και την

εμπειρία, όπως ακριβώς και οι άνθρωποι, με τεχνικές και ορισμένους κανόνες διαφορετικούς από αυτούς που εκτελεί ένα συνηθισμένο υπολογιστικό σύστημα. Ο κυρίαρχος σκοπός της λειτουργίας ενός Νευρωνικού Δικτύου είναι να εκτελεί σύνθετες διεργασίες, π.χ. αναγνώριση εικόνας ή κατανόηση και επεξεργασία φυσικής γλώσσας, αφού προηγουμένως έχει εκπαιδευτεί κατάλληλα. Η διαδικασία της εκπαίδευσης γίνεται με το να παρέχουμε στο δίκτυο μια ομάδα από πρότυπα ως είσοδο, για τα οποία όμως γνωρίζουμε ποια θα πρέπει να είναι η αναμενόμενη έξοδος για αυτά. Ουσιαστικά είναι σαν να δίνουμε στο δίκτυο μια ερώτηση και την απάντησή της. Το Νευρωνικό Δίκτυο εκτελώντας τον αλγόριθμο εκμάθησης, τροποποιεί την εσωτερική δομή του, ώστε να μπορέσει με τις εισόδους που του δώσαμε να παραγάγει τις αντίστοιχες εξόδους. Έτσι, το δίκτυο αποκτά γνώση και ύστερα μπορεί να εκτιμήσει την έξοδο από εισόδους για τις οποίες δεν τη γνωρίζει. Διεξοδικότερη ανάλυση στην θεωρία των Νευρωνικών Δικτύων θα πραγματοποιηθεί στο επόμενο κεφάλαιο.

## 1.5 Εισαγωγή στην Επεξεργασία Φυσικής Γλώσσας

Μια από τις πιο δημοφιλείς εφαρμογές της ΤΝ είναι η Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing - NLP), ή για συντομία ΕΦΓ. Η βασική ιδέα είναι να μπορεί ο άνθρωπος να επικοινωνεί με τον υπολογιστή, να αλληλεπιδρά μαζί του και να έχει τη δυνατότητα να τον ελέγχει και να του δίνει εντολές με την φυσική του γλώσσα και όχι με τεχνητές γλώσσες προγραμματισμού, γλώσσες μηχανής κλπ. Πλήθος τομέων της επιστήμης μπορούν να επωφεληθούν από την ΕΦΓ και την Κατανόηση Φυσικής Γλώσσας (Natural Language Understanding - NLU) και τις εφαρμογές της. Ο κυριότερος είναι αυτός της επικοινωνίας ανθρώπου-μηχανής (human-computer interaction), ο οποίος έχει αναπτυχθεί αρκετά τα τελευταία χρόνια και έχει ενθαρρύνει πολλούς επιστήμονες της ΤΝ να ασχοληθούν με τον κλάδο αυτό. Ένας ορισμός που έχει δοθεί είναι ο παρακάτω:

*Η Επεξεργασία Φυσικής Γλώσσας είναι ένας διεπιστημονικός κλάδος της επιστήμης της πληροφορικής, της τεχνητής νοημοσύνης και της υπολογιστικής γλωσσολογίας και ασχολείται με τις αλληλεπιδράσεις μεταξύ των υπολογιστών και των ανθρώπινων (φυσικών) γλωσσών. Κα-*

*τά συνέπεια, η ΕΦΓ συνδέεται στενά με την αλληλεπίδραση ανθρώπου-υπολογιστή. [17]*

### 1.5.1 Δυσκολίες στην Κατανόηση Φυσικής Γλώσσας

Για να υλοποιηθούν όμως συστήματα που έχουν την δυνατότητα κατανόησης και επεξεργασίας της φυσικής γλώσσας, υπάρχουν πολλές δυσκολίες που πρέπει να εξομαλυνθούν και να ξεπεραστούν. Κάποιες από τις πιο συνήθεις δυσκολίες είναι οι εξής:

- **ασάφεια σε επίπεδο σύνταξης**

Κάποιες συντακτικά ορθές προτάσεις επιδέχονται παραπάνω από μια ερμηνεία, ανάλογα με τη συντακτική προσέγγιση. Για παράδειγμα, στην πρόταση “Χτύπησα τον κλέφτη με το τσεκούρι”, το τσεκούρι είναι του κλέφτη ή εγώ χτύπησα τον κλέφτη με ένα τσεκούρι;

- **ασάφεια σε επίπεδο λεξιλογικό**

Όταν το νόημα μιας λέξης είναι διφορούμενο, δηλαδή έχει παραπάνω από μια έννοιες. Για παράδειγμα, η λέξη “γράμμα” έχει δύο έννοιες, της επιστολής και του γράμματος του αλφαβήτου.

- **ασάφεια σε αναφορικό επίπεδο**

Εδώ η δυσκολία παρατηρείται στο γεγονός ότι σε κάποιες περιπτώσεις δεν είναι εφικτό και ευκρινές σε ποιον, πού ή σε τί αναφέρεται μια πρόταση. Για παράδειγμα στην πρόταση “Ο Γιώργος δεν βλέπει την Μαρία συχνά επειδή δουλεύει πολύ.” Ποιος δουλεύει πολύ, ο Γιώργος ή η Μαρία;

- **ασάφεια σε σημασιολογικό επίπεδο**

Όταν, με διατήρηση της ίδιας συντακτικής ανάλυσης, η πρόταση επιδέχεται τουλάχιστον δύο διαφορετικές ερμηνείες. Για παράδειγμα στην πρόταση “Τον άφησε στα χρύα του λουτρού” υπάρχει η κυριολεκτική και η μεταφορική ερμηνεία.

- **ασάφεια σε πραγματολογικό επίπεδο**

Όταν κατά την διερμηνεία μιας πρότασης λαμβάνουμε υπόψη το πλαίσιο του κειμένου που την περιέχει. Στην πρόταση “Οι δεινόσαυροι έχουν εξαφανιστεί πολλά χρόνια πριν” δεν είναι εύκολο να προσδιοριστούν πόσα χρόνια είναι τα “πολλά χρόνια”. [18]

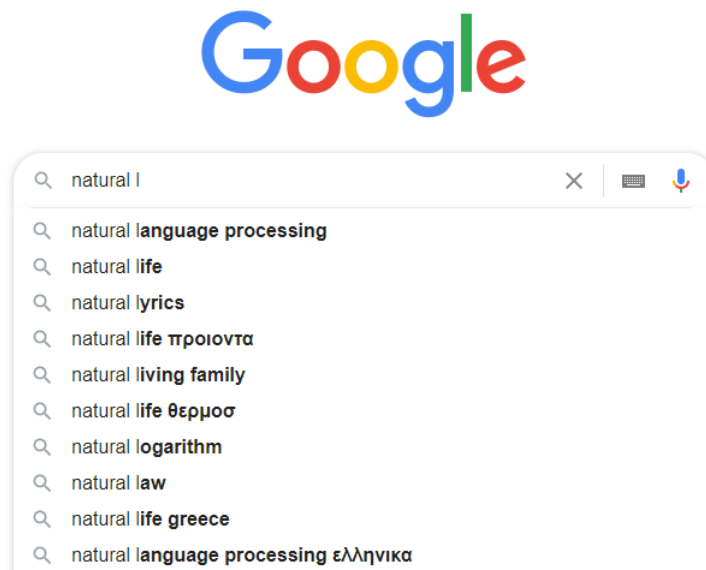
Όπως φαίνεται, υπάρχουν αρκετές προκλήσεις για να επιτευχθεί η ορθή κατανόηση φυσικής γλώσσας από έναν υπολογιστή. Ο τομέας της Βαθιάς Μάθησης έρχεται εδώ για να βοηθήσει και να δώσει λύσεις σε πολλά προβλήματα. Ειδικότερες τεχνικές, διαδικασίες αλλά και μοντέλα Επεξεργασίας Φυσικής Γλώσσας θα παρουσιαστούν λεπτομερώς στο επόμενο κεφάλαιο.

### 1.5.2 Εφαρμογές Επεξεργασίας Φυσικής Γλώσσας

Η χρησιμότητα της ΕΦΓ είναι τεράστια στην πληροφορική και οι εφαρμογές της είναι πολύ περισσότερες από αυτές που μπορούμε να φανταστούμε με μια πρώτη σκέψη. Πολλά από τα εργαλεία και τα προγράμματα που χρησιμοποιούμε σήμερα δεν θα υπήρχαν ή δεν θα λειτουργούσαν το ίδιο καλά και αποτελεσματικά, χωρίς τη συμβολή της ΕΦΓ. Κάποιες από τις πιο σημαντικές και ευρέως γνωστές εφαρμογές της αναλύονται παρακάτω:

- **Autocomplete σε μηχανές αναζήτησης**

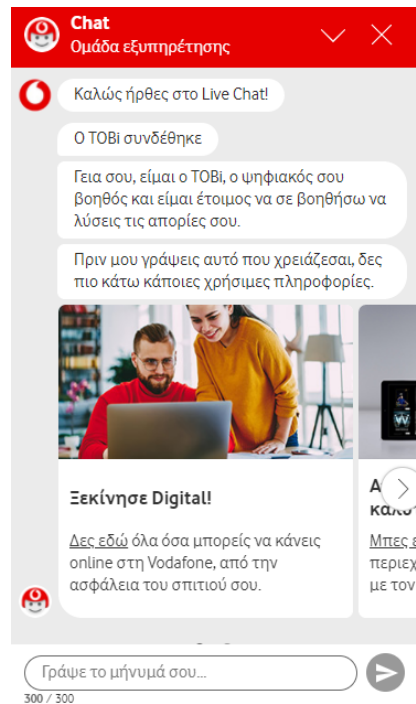
Καθημερινά, το μεγαλύτερο ποσοστό των χρηστών του διαδικτύου, αν όχι όλο, χρησιμοποιεί Μηχανές Αναζήτησης (Search Engines) για να αναζητήσει κάποια πληροφορία σε μια βάση δεδομένων. Η Google είναι μια από τις πιο δημοφιλείς μηχανές αναζήτησης με εξαιρετικά πολύπλοκους αλγόριθμους να τρέχουν στους εξυπηρετητές (servers) της. Ιδιαίτερο ενδιαφέρον έχει όταν, κατά την αναζήτηση, οι χρήστες πληκτρολογούν μόλις 2-3 γράμματα και αμέσως εμφανίζονται υποψήφια προτάσεις συμπληρωμένες αυτόματα, οι οποίες τις περισσότερες φορές είναι αυτές που όντως ο χρήστης επιθυμεί να αναζητήσει. Αυτή η διαδικασία είναι αποτέλεσμα ΕΦΓ και διευκολύνει καθημερινά πάρα πολλούς χρήστες.



Σχήμα 1.8: Παράδειγμα autocomplete σε αναζήτηση στο Google

- **Ταξινόμηση κειμένου** Η ταξινόμηση κειμένου (text classification) είναι μια διαδικασία της ΕΦΓ η οποία περιέχει την αυτόματη ανάλυση, κατανόηση και κατηγοριοποίηση ενός κειμένου σε μια συγκεκριμένη κατηγορία (π.χ. μια κριτική αν είναι θετική ή αρνητική, ένα άρθρο αν αφορά πολιτικά, αθλητικά, επιστημονικά ζητήματα κλπ). Η αυτόματη ταξινόμηση κειμένων μπορεί να φανεί πολύ χρήσιμη όταν υπάρχει μεγάλο πλήθος κειμένων τα οποία θα απαιτούσαν αρκετό χρόνο από κάποιον να τα διαβάσει και να τα κατηγοριοποιήσει. Επιπλέον, μπορεί να βοηθήσει κάποιον να καταλάβει τη θεματολογία και το περιεχόμενο ενός κειμένου χωρίς να χρειαστεί να το αναγνώσει ολόκληρο.
- **Chatbots**  
Άλλη μια σημαντική εφαρμογή της ΕΦΓ εντοπίζεται στην δημιουργία chatbots ή αλλιώς εικονικών βοηθών. Πρόκειται για συστήματα με τα οποία ο χρήστης μπορεί να έχει μια κανονική συνομιλία μαζί τους μέσω μηνυμάτων με σκοπό την εξυπηρέτησή του. Το chatbot έχει την ικανότητα να αναγνωρίσει τί ζητάει ο χρήστης και το πρόβλημα που έχει και να αναλάβει δράση είτε επιλύοντάς το απευθείας αν είναι απλό είτε προωθώντας το σε κάποιον υπεύθυνο για την περαιτέρω επεξεργασία του. Τα chatbots μπορούν να βοηθήσουν στην εξυπηρέτηση

πελατών, δίνοντας την δυνατότητα να εξυπηρετούνται περισσότεροι την ίδια στιγμή αλλά και πιο γρήγορα, κατηγοριοποιώντας το πρόβλημα που έχει ο καθένας αλλά και επιλύοντας απλά ζητήματα τα οποία αποτελούν σπατάλη χρόνου για κάποιον υπάλληλο. Επίσης, μειώνουν τον χρόνο αναμονής των πελατών και είναι μια σημαντική αναβάθμιση για μια επιχείρηση.

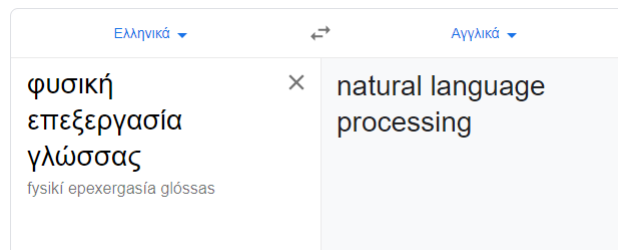


Σχήμα 1.9: Παράδειγμα του chatbot της Vodafone

- **Μετάφραση**

Η μετάφραση ενός κειμένου σε μια άλλη γλώσσα από αυτή που είναι γραμμένο είναι μια διαδικασία η οποία προϋποθέτει γνώση και των δύο γλωσσών και ίσως την βοήθεια λεξικού. Με την συμβολή της ΕΦΓ, όμως, είναι εφικτό να κατασκευαστούν συστήματα τα οποία μεταφράζουν αυτόματα κείμενα από μια γλώσσα σε μια άλλη και μάλιστα με τεράστια επιτυχία. Η Google, που αναφέρθηκε προηγουμένως, έχει υλοποιήσει το Google Translate, ένα εργαλείο που πραγματοποιεί μεταφράσεις κειμένων πολύ γρήγορα αλλά και με μεγάλη ακρίβεια, σαν να τις έκανε μεταφραστής.





Σχήμα 1.10: Παράδειγμα Google Translate για μετάφραση γλώσσας

Αυτές είναι κάποιες από τις πολλές εφαρμογές της ΕΦΓ, οι οποίες συνεχώς βελτιώνονται και αυξάνονται. Επιχειρηματίες αναζητούν μηχανικούς και επιστήμονες για να αξιοποιήσουν τα δεδομένα που έχουν στη διάθεσή τους και να αναπτύξουν τις επιχειρήσεις τους με τέτοιες τεχνικές. Όλο και περισσότεροι στρέφουν το ενδιαφέρον τους προς αυτόν τον κλάδο και όλο και περισσότερες προκλήσεις δημιουργούνται. Μέρα με την μέρα τα συστήματα μαθαίνουν και αποκτούν περισσότερη γνώση, γίνονται πιο ευφυή, επεξεργάζονται και κατανοούν την φυσική γλώσσα, κάποιες φορές, καλύτερα και από τους ανθρώπους. Βοηθούν σε αμέτρητα προβλήματα, εκτελούν εργασίες διευκολύνοντας τους ανθρώπους, εξοικονομώντας τους χρόνο και σκέψη.

## Κεφάλαιο 2

### Θεωρητικό Υπόβαθρο

Στο κεφάλαιο αυτό, γίνεται μια πιο βαθιά και επιστημονική ανάλυση στις έννοιες που αναφέρθηκαν στο 1<sup>ο</sup> Κεφάλαιο. Αρχικά, γίνεται μια εισαγωγή στα είδη μάθησης και στη συνέχεια, επεξηγούνται οι τρόποι με τους οποίους πραγματοποιείται η διαδικασία της μάθησης. Επιπλέον, αναλύονται η αρχιτεκτονική και η τρόποι λειτουργίας των Νευρωνικών Δικτύων καθώς και οι διάφορες δομές και κύριες κατηγορίες τους.

## 2.1 Διαδικασίες Μάθησης

Δύο από τις πιο βασικές διαδικασίες μάθησης μέσω των οποίων τα συστήματα μαθαίνουν και αποκτούν γνώση είναι, όπως αναφέραμε και στην ενότητα 1.2.3, η Επιβλεπόμενη και η Μη Επιβλεπόμενη Μάθηση. Με αυτές τις δύο θα ασχοληθούμε ως επί το πλείστον σε αυτή τη διπλωματική εργασία. Για αυτό λοιπόν είναι αναγκαία η πλήρης κατανόησή τους και η ανάλυση των μεθόδων και τεχνικών τους, προτού προχωρήσουμε στην εφαρμογή τους σε προβλήματα.

### 2.1.1 Επιβλεπόμενη Μάθηση

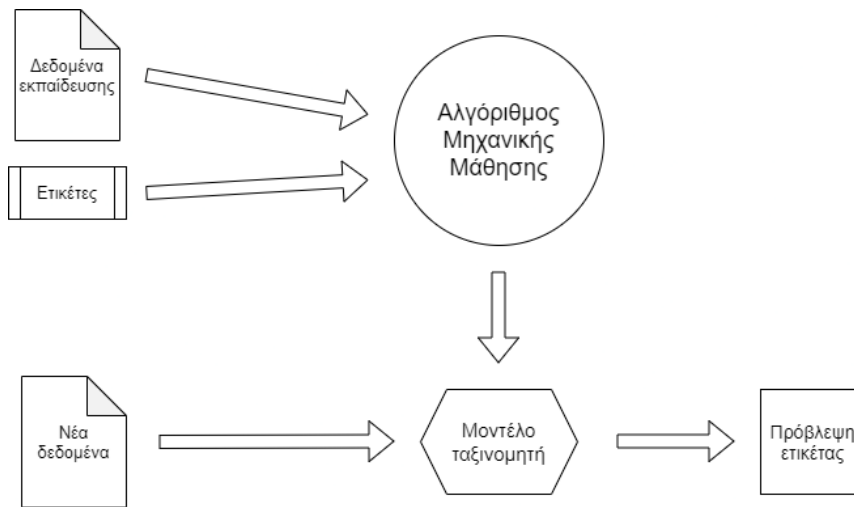
Η πιο συνηθισμένη και απλή μορφή Μηχανικής Μάθησης είναι αυτή της Επιβλεπόμενης Μάθησης. Έστω ότι έχουμε ένα σύνολο  $N$  δειγμάτων  $X = \{x_1, x_2, \dots, x_N\}$ , το κάθε ένα από τα οποία ανήκει σε μια κατηγορία  $y_i$ , δηλαδή  $Y = \{y_1, y_2, \dots, y_N\}$ , από ένα σύνολο κατηγοριών. Το κάθε δείγμα  $x_i$  αποτελείται από  $M$  χαρακτηριστικά τα οποία το περιγράφουν,  $x_i = \{x_i^1, x_i^2, \dots, x_i^M\}$ . Την κατηγορία  $y_i$  θα την αποκαλούμε και ετικέτα (label). Το σύνολο των δειγμάτων μπορεί να είναι από εικόνες και κείμενα μέχρι ιατρικές εξετάσεις και καιρικές συνθήκες: γενικά, οτιδήποτε έχει ετικέτα και διαθέτει κάποια χαρακτηριστικά μπορεί να αποτελέσει το δείγμα. Ένα παράδειγμα εικόνων με μέσα μεταφοράς μαζί με τις ετικέτες τους φαίνεται παρακάτω:



Σχήμα 2.1: Δείγματα εικόνων με τις ετικέτες τους

Σε αυτή την διπλωματική εργασία θα ασχοληθούμε κυρίως με δείγματα βιβλίων και πιο συγκεκριμένα με τον πίνακα περιεχομένων τους. Ως ετικέτα το κάθε δείγμα θα έχει τη θεματική κατηγορία στην οποία ανήκει το βιβλίο (π.χ. επιστήμη υπολογιστών, ιατρική, μαθηματικά κλπ).

Στην επιβλεπόμενη μάθηση, ο σκοπός είναι η εκπαίδευση ενός μοντέλου, έστω  $g$ , το οποίο θα είναι ικανό να βρίσκει τη σωστή ετικέτα  $y_i$  για ένα δείγμα  $x_i$ . Αυτό το πρόβλημα είναι γνωστό ως πρόβλημα κατηγοριοποίησης ή ταξινόμησης (classification). Ένα τέτοιο μοντέλο μπορεί να ονομαστεί και ταξινομητής (classifier). Για να μπορέσει το μοντέλο να ταξινομεί σωστά τα δείγματα, θα πρέπει να διακρίνει και να μάθει τις συσχετίσεις που υπάρχουν ανάμεσα στα δείγματα και τις ετικέτες τους, πώς δηλαδή οι τιμές των χαρακτηριστικών του κάθε  $x_i$  επηρεάζουν την κατηγορία  $y_i$ .



Σχήμα 2.2: Σχηματικό διάγραμμα Επιβλεπόμενης Μάθησης για μοντέλο ταξινόμητη

Ακολουθώντας τον ορισμό του Mitchell στην ενότητα 1.2.1, η εμπειρία  $E$  αποτελείται από το σύνολο των δειγμάτων με τις ετικέτες τους και το έργο  $T$  είναι το πρόβλημα ταξινόμησης. Αυτό που μένει, είναι να ορίσουμε ένα μέτρο επίδοσης  $P$ , το οποίο θα υποδεικνύει πόσο καλά αποδίδει ο αλγόριθμος. Ως ένα τέτοιο μέτρο μπορούμε να ορίσουμε το κατά πόσο συγκλίνουν οι προβλέψεις του μοντέλου  $\hat{y}$  με τις πραγματικές ετικέτες  $y$ . Ρόλος λοιπόν της εκπαίδευσης είναι να αυξήσει το μέτρο  $P$  φέρνοντας τις προβλέψεις του μοντέλου όσο πιο κοντά γίνεται στις αναμενόμενες [6]. Ως μέτρο επίδοσης  $P$  θα μπορούσαμε να χρησιμοποιήσουμε το μέσο τετραγωνικό σφάλμα ή το άθροισμα των τετραγώνων των σφαλμάτων επί του δείγματος εκπαίδευσης.

Τα δεδομένα εκπαίδευσης μαζί με τις αντίστοιχες ετικέτες τους εισάγονται ως είσοδος στον αλγόριθμο Μηχανικής Μάθησης. Οι ετικέτες δείχνουν την επιθυμητή απόκριση, δηλαδή τη βέλτιστη ενέργεια, που πρέπει να εκτελείται από το μοντέλο. Οι παράμετροι του συστήματος ενημερώνονται συνέχεια επηρεασμένες από το διάνυσμα

εκπαίδευσης, το οποίο προκύπτει από τα δεδομένα εκπαίδευσης και το σήμα σφάλματος. Ως “σήμα σφάλματος” ορίζεται η διαφορά μεταξύ επιθυμητής και πραγματικής απόκρισης του μοντέλου. Αυτή η προσαρμογή εκτελείται επαναληπτικά μέσω του αλγορίθμου μάθησης. Με αυτόν τον τρόπο η γνώση του περιβάλλοντος που είναι διαθέσιμη, μεταφέρεται στο μοντέλο μέσω της μάθησης και αποθηκεύεται στις παραμέτρους του. Όταν ολοκληρωθεί η παραπάνω διαδικασία, το μοντέλο ταξινομητή, διαθέτοντας πλέον γνώση, είναι έτοιμο να κάνει προβλέψεις σε νέα δεδομένα, όπως φαίνεται και στο Σχήμα 2.2 και να αντιμετωπίσει το περιβάλλον εντελώς μόνο του. Αυτή η μορφή της επιβλεπόμενης μάθησης που περιγράφηκε αποτελεί τη βάση της μάθησης μέσω διόρθωσης σφάλματος.

### 2.1.2 Μη Επιβλεπόμενη Μάθηση

Στη Μη Επιβλεπόμενη Μάθηση ή αλλιώς αυτο-οργανούμενη, δεν υπάρχει εξωτερικός εκπαιδευτής που επιβλέπει τη μάθηση ή πιο απλά, τα δεδομένα εκπαίδευσης είναι ανοργάνωτα και μη κατηγοριοποιημένα με ετικέτες. Ο αλγόριθμος εδώ καλείται μέσω της μάθησης να προσεγγίσει μια συνάρτηση η οποία να περιγράφει την κρυφή δομή στα μη επισημασμένα δεδομένα, να τα ομαδοποιήσει και να διακρίνει κοινά χαρακτηριστικά τους, συσχετίσεις και διαφορές μεταξύ τους. Στόχος, δηλαδή, είναι να βρεθεί μια κατηγοριοποίηση ή ταξινόμηση η οποία δεν περιλαμβάνεται αρχικά στα δεδομένα.



Σχήμα 2.3: Σχηματικό διάγραμμα Μη Επιβλεπόμενης Μάθησης

Αφού, λοιπόν, δεν έχουμε ετικέτες για τα δεδομένα, δεν θα μπορέσουμε να ορίσουμε το μέτρο επίδοσης  $P$ , όπως πριν. Πρέπει να ορίσουμε όμως ένα τέτοιο μέτρο μιας και είναι αναγκαίο για τη διαδικασία της μάθησης. Υπάρχει εδώ ένα, ανεξάρτητο από την εργασία, μέτρο της ποιότητας της αναπαράστασης που καλείται να μάθει το μοντέλο και οι ελεύθερες παράμετροί του βελτιστοποιούνται σε σχέση με αυτό το μέτρο.

*“Για ένα συγκεκριμένο ανεξάρτητο από την εργασία μέτρο, αφού το*

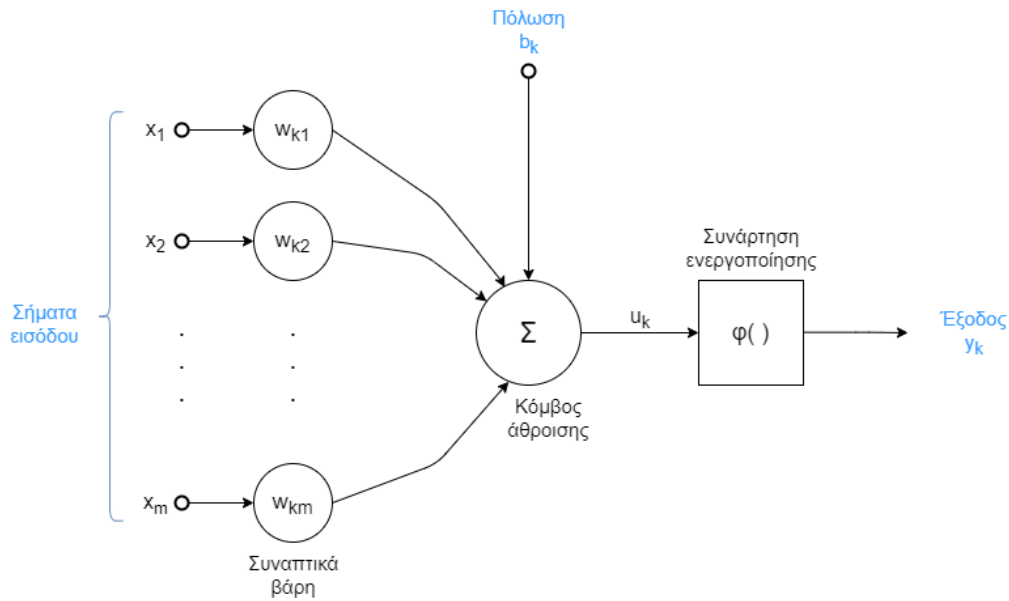
*δίκτυο συντονιστεί με τις στατιστικές κανονικότητες των δεδομένων εισόδου, αναπτύσσει τη δυνατότητα να σχηματίζει εσωτερικές αναπαραστάσεις για την κωδικοποίηση χαρακτηριστικών της εισόδου και, μέσω αυτών, να δημιουργεί νέες κλάσεις αυτόματα.” (Becker, 1991)*

## 2.2 Νευρωνικά Δίκτυα

Το μοντέλο g, το οποίο εκπαιδεύεται με διαδικασίες μάθησης που περιγράψαμε παραπάνω, μπορεί να είναι και ένα Νευρωνικό Δίκτυο. Όπως αναφέραμε και στην ενότητα 1.4, τα νευρωνικά δίκτυα δημιουργήθηκαν βασισμένα στον ανθρώπινο εγκέφαλο και τα βιολογικά του Νευρωνικά Δίκτυα. Βασική υπολογιστική τους μονάδα είναι ο νευρώνας.

### 2.2.1 Μοντέλο Νευρώνα

Ένας νευρώνας είναι μια μονάδα επεξεργασίας πληροφορίας, τα μέρη του οποίου μπορεί να αντιστοιχιστούν άμεσα σε αυτά του βιολογικού νευρώνα. Η λειτουργία του είναι θεμελιώδης για τη λειτουργία ενός Νευρωνικού Δικτύου. Στο Σχήμα 2.4 παρουσιάζεται το μοντέλο ενός νευρώνα που αποτελεί τη βάση για τη σχεδίαση μιας μεγάλης οικογένειας Νευρωνικών Δικτύων βασισμένα σε αυτόν με διαφορετικές αρχιτεκτονικές και λειτουργίες.



Σχήμα 2.4: Μοντέλο νευρώνα

Τα τρία βασικά στοιχεία του μοντέλου αυτού είναι τα εξής:

1. Ένα **σύνολο συνάψεων** όπου κάθε μια χαρακτηρίζεται από το δικό της βάρος, το συναπτικό βάρος, η τιμή του οποίου μπορεί να είναι θετική ή αρνητική, σε αντιστοιχία με την επιταχυντική ή επιβραδυντική λειτουργία της σύναψης. Συγκεκριμένα ένα σήμα  $x_j$  στην είσοδο της σύναψης  $j$ , η οποία συνδέεται με τον νευρώνα  $k$ , πολλαπλασιάζεται με το συναπτικό βάρος  $w_{kj}$ . Ο συμβολισμός του βάρους  $w_{kj}$  προκύπτει ως εξής:
  - Ο πρώτος δείκτης  $k$  αναφέρεται στον εν λόγω νευρώνα που βρίσκεται η σύναψη.
  - Ο δεύτερος δείκτης  $j$  αναφέρεται στο άκρο εισόδου της σύναψης στην οποία αναφέρεται το βάρος.
2. Έναν **αθροιστή** (adder) ο οποίος αθροίζει τα σήματα εισόδου, σταθμισμένα με τα αντίστοιχα συναπτικά βάρη του νευρώνα, παράγοντας την ποσότητα  $u_k$ .
3. Μια **συνάρτηση ενεργοποίησης** (activation function) η οποία λειτουργεί σαν φίλτρο, διαμορφώνοντας την τελική τιμή σήματος του νευρώνα  $y_k$ . Η συνάρτηση ενεργοποίησης καλείται και συνάρτηση περιορισμού (squashing function),

επειδή περιορίζει το επιτρεπτό εύρος πλάτους του σήματος εξόδου σε κάποια πεπερασμένη τιμή. Διευκρινίζεται ότι η τιμή εξόδου του κάθε νευρώνα είναι μοναδική, ανεξάρτητα από το πλήθος των εξόδων που έχει ο νευρώνας· αν έχει παραπάνω από μια, τότε όλες θα έχουν την ίδια τιμή.

Το μοντέλο νευρώνα που παρουσιάζεται στο Σχήμα 2.4 περιέχει και την πόλωση  $b_k$ . Η εξωτερικά εφαρμοζόμενη αυτή πόλωση έχει ως αποτέλεσμα την αύξηση ή μείωση της δικτυακής διέγερσης της συνάρτησης ενεργοποίησης του νευρώνα, ανάλογα με το αν είναι θετική ή αρνητική.

Μπορούμε να περιγράψουμε το μοντέλο νευρώνα  $k$  με τους παρακάτω μαθηματικούς όρους γράφοντας το παρακάτω ζεύγος εξισώσεων:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (2.1)$$

και

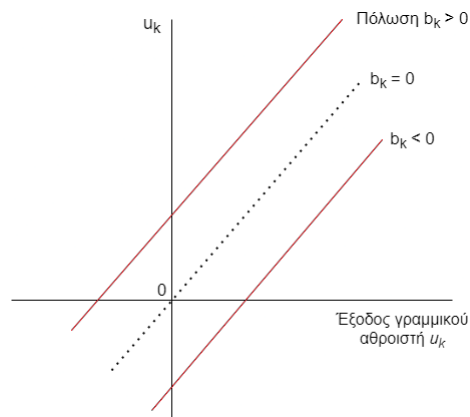
$$y_k = \phi(u_k + b_k) \quad (2.2)$$

όπου  $x_1, x_2, \dots, x_m$  είναι τα σήματα εισόδου και  $w_{k1}, w_{k2}, \dots, w_{km}$  είναι τα αντίστοιχα συναπτικά βάρη του νευρώνα  $k$ . Ακολούθως,  $b_k$  είναι η πόλωση,  $u_k$  είναι η έξοδος του αθροιστή των σημάτων εισόδου,  $\phi()$  είναι συνάρτηση ενεργοποίησης και  $y_k$  είναι το σήμα εξόδου του νευρώνα. Η χρήση της πόλωσης  $b_k$  έχει ως αποτέλεσμα έναν αφινικό μετασχηματισμό στην έξοδο  $u_k$  του γραμμικού αθροιστή, όπως φαίνεται και από την σχέση:

$$u_k = u_k + b_k \quad (2.3)$$

Ένας νευρώνας με  $m$  εισόδους ορίζει ένα υπερεπίπεδο στον  $m$ -διάστατο χώρο, το οποίο τον διαχωρίζει σε δύο περιοχές. Οι τιμές των συναπτικών βαρών του νευρώνα καθορίζουν τη θέση του υπερεπιπέδου αυτού. Χωρίς την εξωτερική παρέμβαση της πόλωσης  $b_k$ , δηλαδή  $b_k = 0$  στην σχέση (2.3), το υπερεπίπεδο θα ήταν αναγκαίο να διέρχεται από την αρχή των αξόνων, κάτι το οποίο δεν είναι επιθυμητό σε ορισμένες περιπτώσεις.





Σχήμα 2.5: Αφινικός μετασχηματισμός λόγω ύπαρξης πόλωσης

### 2.2.1.1 Τύποι της Συνάρτησης Ενεργοποίησης

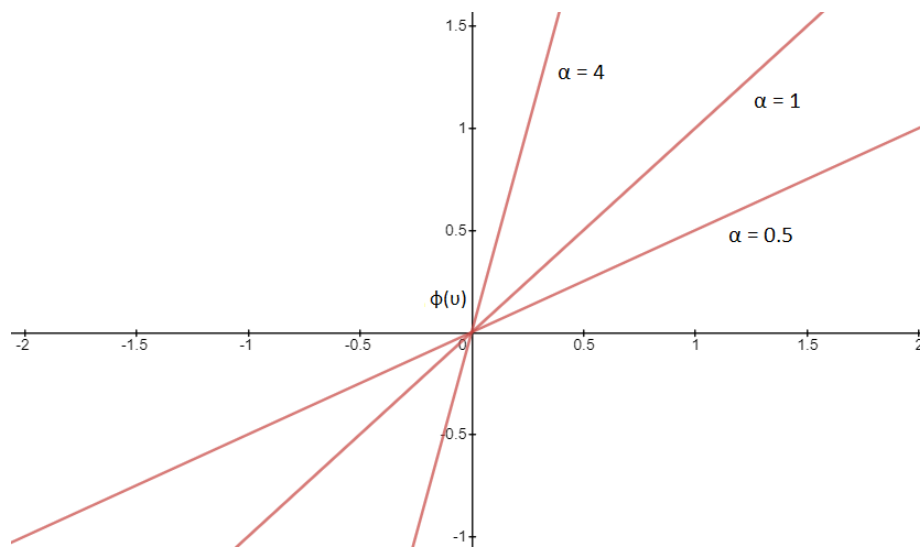
Η συνάρτηση ενεργοποίησης του νευρώνα, η οποία συμβολίζεται ως  $\phi(u)$ , ορίζεται ως η έξοδος του νευρώνα αυτού βάσει του τοπικού πεδίου  $u$ . Παρακάτω θα παρουσιαστούν κάποιοι από τους πιο βασικούς τύπους συναρτήσεων ενεργοποίησης, οι οποίοι θα εφαρμοστούν στη συνέχεια πάνω στην κατασκευή μοντέλων και νευρωνικών δικτύων. Αρχικά, οι συναρτήσεις ενεργοποίησης χωρίζονται σε δύο κατηγορίες οι οποίες είναι οι:

1. **Γραμμικές Συναρτήσεις (Linear Functions)**
2. **Μη Γραμμικές Συναρτήσεις (Non-Linear Functions)**

Στην κατηγορία των γραμμικών συναρτήσεων όλες οι συναρτήσεις είναι της μορφής

$$\phi = \alpha u \quad (2.4)$$

όπου  $\alpha \in \mathbb{R}$ , και η αντίστοιχη γραφική τους παράσταση φαίνεται παρακάτω:



Σχήμα 2.6: Γραμμική συνάρτηση

Οι γραμμικές συναρτήσεις ενεργοποίησης δεν χρησιμοποιούνται ιδιαίτερα στην Μηχανική Μάθηση και σε Νευρωνικά Δίκτυα αφού δε βοηθούν σε προβλήματα με μεγάλη πολυπλοκότητα και ποικιλία μεταβλητών.

Οι Μη Γραμμικές Συναρτήσεις ενεργοποίησης είναι οι πιο διαδεδομένες, με πολλές δυνατότητες και χαρακτηριστικά η κάθε μια από αυτές. Παρακάτω παρουσιάζονται μερικές από τις πιο βασικές οι οποίες θα χρησιμοποιηθούν και στη διπλωματική αυτή εργασία:

- **Συνάρτηση Κατωφλίου (Threshold Function)**

Ο τύπος αυτής της συνάρτησης είναι ο εξής:

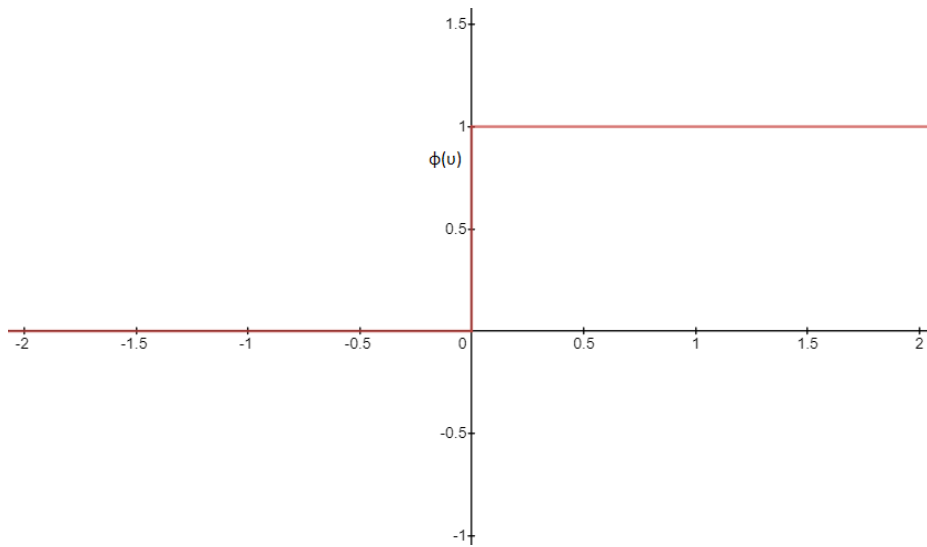
$$\phi(v) = \begin{cases} 1 & \text{εάν } v \geq 0 \\ 0 & \text{εάν } v < 0 \end{cases} \quad (2.5)$$

Στους κλάδους της μηχανικής, αυτή η μορφή συνάρτησης κατωφλίου είναι γνωστή και ως συνάρτηση Heaviside. Έτσι, η έξοδος του νευρώνα  $k$  είναι:

$$y_k = \begin{cases} 1 & \text{εάν } v_k \geq 0 \\ 0 & \text{εάν } v_k < 0 \end{cases} \quad (2.6)$$

όπου  $v_k$  είναι το τοπικό πεδίο του νευρώνα, δηλαδή:

$$v_k = \sum_{j=1}^m w_{kj}x_j + b_k \quad (2.7)$$



Σχήμα 2.7: Συνάρτηση κατωφλίου

Το παραπάνω μοντέλο νευρώνα αναφέρεται ως μοντέλο McCulloch-Pitts, εις ένδειξιν αναγνώρισης του πρωτοποριακού έργου των McCulloch και Pitts το 1943 [14]. Σε αυτό το μοντέλο ουσιαστικά η έξοδος του νευρώνα λαμβάνει την τιμή 1, αν το τοπικό πεδίο είναι θετικό ή μηδέν και την τιμή 0, αν είναι αρνητικό.

- **Συνάρτηση Προσήμου (Sign Function)**

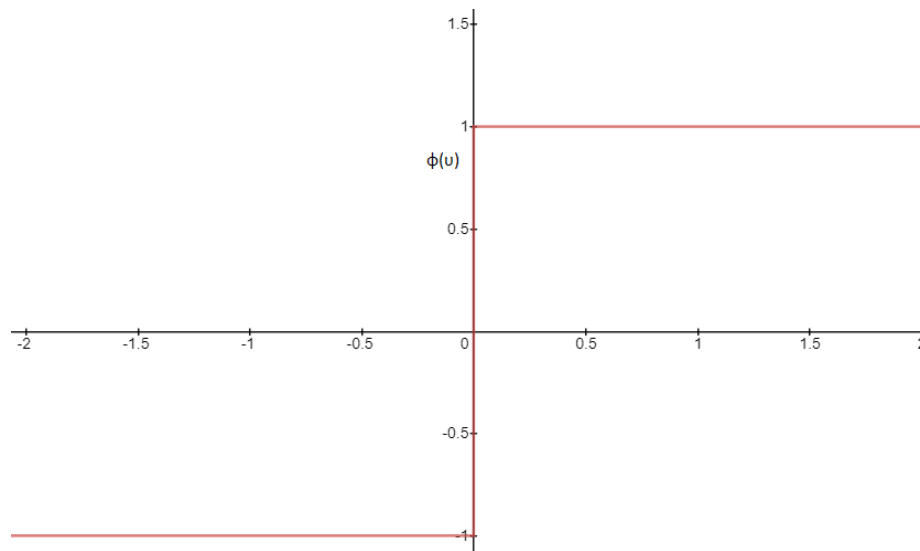
Ο τύπος αυτής της συνάρτησης είναι ο εξής:

$$\phi(v) = \begin{cases} 1 & \text{εάν } v > 0 \\ 0 & \text{εάν } v = 0 \\ -1 & \text{εάν } v < 0 \end{cases} \quad (2.8)$$

Και αντίστοιχα η έξοδος του νευρώνα  $k$  είναι:

$$y_k = \begin{cases} 1 & \text{εάν } v_k > 0 \\ 0 & \text{εάν } v_k = 0 \\ -1 & \text{εάν } v_k < 0 \end{cases} \quad (2.9)$$

Η γραφική παράσταση της συνάρτησης προσήμου είναι η εξής:



Σχήμα 2.8: Συνάρτηση προσήμου

Ουσιαστικά πρόκειται για μια συνάρτηση η οποία μοιάζει με την συνάρτηση κατωφλίου με τη μόνη διαφορά ότι για αρνητικές τιμές δίνει έξοδο -1 και όχι 0.

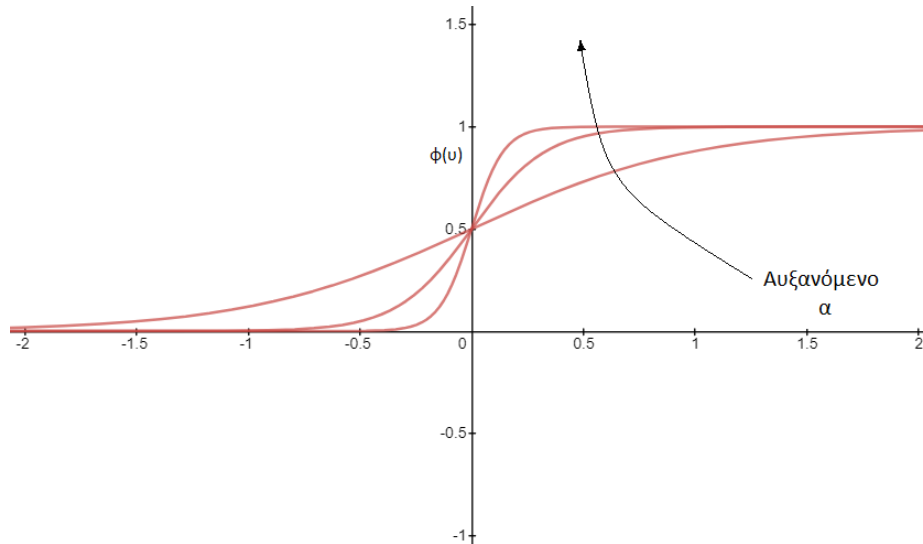
- **Σιγμοειδής Συνάρτηση (Sigmoid Function)**

Η σιγμοειδής ή αλλιώς λογιστική συνάρτηση (logistic function), είναι η πλέον κοινή μορφή συνάρτησης ενεργοποίησης στην κατασκευή Νευρωνικών Δικτύων. Ορίζεται ως αυστηρά αύξουσα συνάρτηση που επιδεικνύει κομψή ισορροπία γραμμικής και μη γραμμικής συμπεριφοράς. Ο τύπος της σιγμοειδούς συνάρτησης είναι ο εξής:

$$\phi(v) = \frac{1}{1 + e^{-av}} \quad (2.10)$$

όπου  $a$  είναι η παράμετρος κλίσης της. Μεταβάλλοντας την παράμετρο αυτή, παίρνουμε διαφορετικές κλίσεις, όπως φαίνεται και στο παρακάτω σχήμα που

δείχνει τη γραφική παράσταση της σιγμοειδούς συνάρτησης, η οποία έχει σχήμα «S»:



Σχήμα 2.9: Σιγμοειδής συνάρτηση

Να σημειωθεί ότι καθώς η παράμετρος κλίσης πλησιάζει το άπειρο, τότε η σιγμοειδής συνάρτηση τείνει να γίνει η συνάρτηση κατωφλίου.

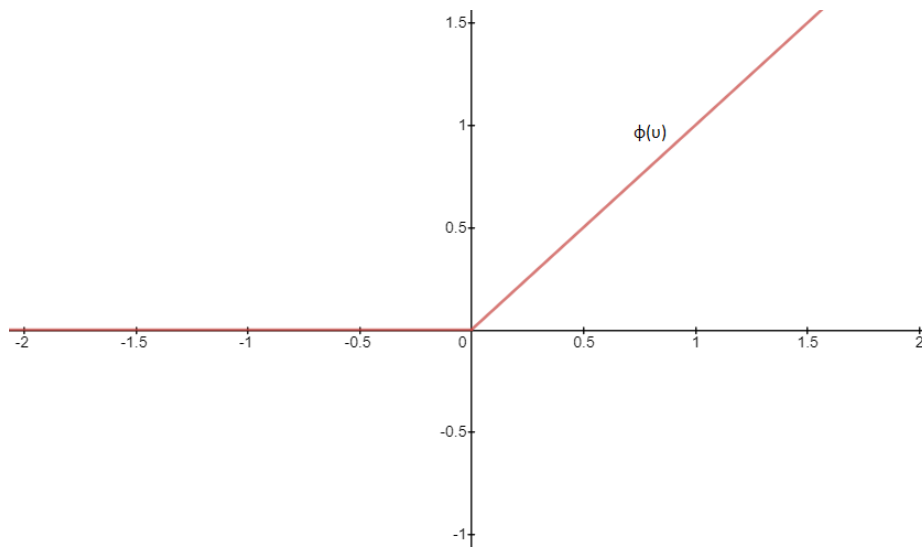
- **Συνάρτηση ReLU (Rectified Linear Unit)** Μια πολύ ιδιαίτερη και ισχυρή συνάρτηση ενεργοποίησης, είναι η συνάρτηση Rectified Linear Unit, εν συντομία ReLU. Πρόκειται για μια συνάρτηση η οποία, σε αντίθεση με τη σιγμοειδή, μοιάζει και “συμπεριφέρεται” σαν γραμμική αλλά ουσιαστικά είναι μη γραμμική βοηθώντας στη διαδικασία της μάθησης πολύπλοκων σχέσεων και συσχετίσεων. Ο τύπος της είναι ο παρακάτω:

$$\phi(v) = \begin{cases} v & \text{εάν } v \geq 0 \\ 0 & \text{εάν } v < 0 \end{cases} \quad (2.11)$$

ή πιο απλά

$$\phi(v) = \max(0, v) \quad (2.12)$$

Η γραφική παράσταση της συνάρτησης relu, για διάφορες τιμές της παραμέτρου  $\alpha$ , παρουσιάζεται παρακάτω:



Σχήμα 2.10: Συνάρτηση relu

Η συνάρτηση relu έχει δώσει λύσεις σε αρκετά προβλήματα μάθησης που έχουν παρουσιαστεί. Οι ιδιότητές της βοηθούν το δίκτυο να μάθει και να διαχειριστεί σωστά την πληροφορία των κόμβων του αλλά επιπλέον απαιτεί λιγότερη υπολογιστική ισχύ σε σχέση με άλλες μη γραμμικές συναρτήσεις.

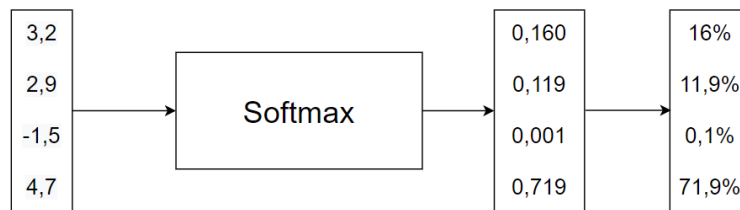
*“Άλλη μια σημαντική αλγοριθμική αλλαγή που έχει βελτιώσει σημαντικά την απόδοση των δικτύων πρόσθιας τροφοδότησης ήταν η αντικατάσταση των σιγμοειδών κρυφών μονάδων με γραμμικές κρυφές μονάδες, όπως μονάδες τύπου *rectified linear unit*” [19]*

#### • Συνάρτηση Softmax

Η συνάρτηση softmax, γνωστή και ως softargmax, είναι μια πιο γενικευμένη μορφή της σιγμοειδούς συνάρτησης με περισσότερες διαστάσεις. Η συνάρτηση αυτή δέχεται ως είσοδο έναν πίνακα  $z$  που περιέχει  $K$  πραγματικούς αριθμούς και υπολογίζει την κατανομή πιθανότητας (probability distribution) που αποτελείται από  $K$  πιθανότητες ανάλογες με τα εκθετικά των αριθμών του πίνακα  $z$ . Η έξοδος της συνάρτησης softmax εκφράζει πιθανότητα. Ο τύπος της είναι ο παρακάτω:

$$\phi(v) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (2.13)$$

όπου  $i = 1, 2, \dots, K$  και  $z = (z_1, z_2, \dots, z_K) \in \mathbb{R}^K$ . Η softmax χρησιμοποιείται κυρίως στους νευρώνες του τελευταίου επιπέδου σε ένα Νευρωνικό Δίκτυο το οποίο ταξινομεί δείγματα σε παραπάνω από μια κλάσεις. Κι αυτό γιατί έχει τη δυνατότητα να κανονικοποιεί την έξοδο κάθε νευρώνα, ο οποίος αντιστοιχεί σε κάποια κλάση, στο διάστημα  $[0, 1]$  και να το διαιρεί με το συνολικό άθροισμα όλων των κλάσεων βρίσκοντας έτσι τις πιθανότητες του δείγματος να ανήκει σε κάθε μια από τις κλάσεις του προβλήματος. Ένα παράδειγμα εφαρμογής της συνάρτησης softmax στον πίνακα  $[3.2 \ 2.9 \ -1.5 \ 4.7]$  φαίνεται παρακάτω:



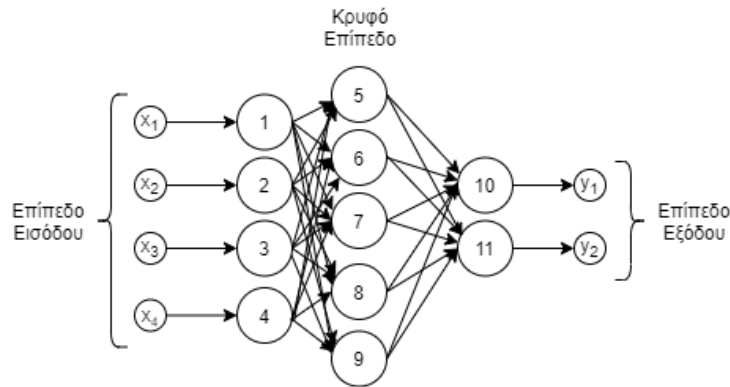
Σχήμα 2.11: Παράδειγμα εφαρμογής της συνάρτησης softmax

### 2.2.2 Πολυεπίπεδα Νευρωνικά Δίκτυα

Τα Νευρωνικά Δίκτυα είναι συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος νευρώνων οργανωμένων κυρίως σε επίπεδα. Τα επίπεδα διατάσσονται έτσι ώστε η έξοδος του ενός να είναι η είσοδος στο επόμενο. Με αυτόν τον τρόπο, ένα Νευρωνικό Δίκτυο έχει τη δυνατότητα να προσεγγίζει και μη γραμμικές συναρτήσεις. Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, κάθε Νευρωνικό Δίκτυο περιέχει:

1. το επίπεδο εισόδου (input layer)
2. το επίπεδο εξόδου (output layer)
3. τα ενδιάμεσα, ή αλλιώς κρυφά, επίπεδα (hidden layers).

Τα ενδιάμεσα επίπεδα ονομάζονται και κρυφά διότι η έξοδος τους δεν είναι παρατηρήσιμη, αλλά υπάρχει με σκοπό να τροφοδοτεί τα επόμενα επίπεδα. Όσο περισσότερα κρυφά επίπεδα έχει ένα Νευρωνικό Δίκτυο, τόσο πιο περίπλοκη είναι η δομή του αλλά και τόσο πιο βαθιά εξετάζει τα δεδομένα και ψάχνει συσχετίσεις μεταξύ τους. Ουσιαστικά Νευρωνικό Δίκτυο βαθιάς μάθησης θεωρείται ένα δίκτυο με παραπάνω από ένα κρυφά επίπεδα.



Σχήμα 2.12: Παράδειγμα Νευρωνικού Δικτύου με πλήρη διασύνδεση μεταξύ διαδοχικών επιπέδων

Συνολικά, ένα Νευρωνικό Δίκτυο με  $M$  επίπεδα έχει μια συνάρτηση εξόδου της μορφής:

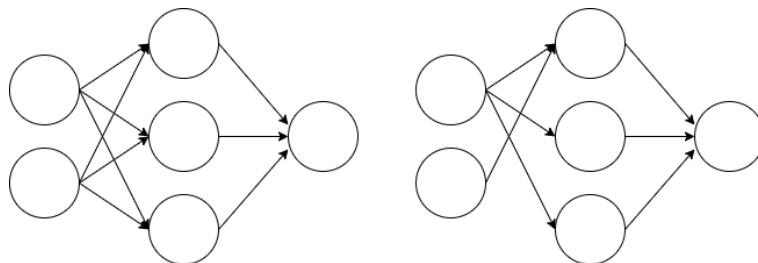
$$\hat{y} = z^M(z^{M-1}(\dots z^1(\hat{x}))) \quad (2.14)$$

όπου  $z^m, m \in \{1, \dots, M\}$  είναι η έξοδος του  $m$ -οστού επιπέδου του δικτύου,  $\hat{x}$  είναι το διάνυσμα εισόδου και  $\hat{y}$  το διάνυσμα εξόδου. Στην περίπτωση πολλαπλών επιπέδων η έξοδος ενός επιπέδου είναι:

$$z^m = f(W \cdot z^{m-1} + b) \quad (2.15)$$

όπου  $W$  είναι το διάνυσμα των βαρών του επιπέδου και  $b$  η αντίστοιχη πόλωση.

Οι νευρώνες σε ένα Νευρωνικό Δίκτυο μπορεί να είναι πλήρως ή μερικώς συνδεδεμένοι. Πλήρως συνδεδεμένοι (fully connected) είναι οι νευρώνες οι οποίοι συνδέονται με όλους τους υπόλοιπους και μερικώς συνδεδεμένοι (partially connected) είναι σε κάθε άλλη περίπτωση.



Σχήμα 2.13: Παράδειγμα πλήρως συνδεδεμένων και μερικώς συνδεδεμένων νευρώνων



Όταν δεν υπάρχουν συνδέσεις μεταξύ νευρώνων ενός επιπέδου και νευρώνων ενός προηγούμενου, δηλαδή όταν η ροή πληροφορίας είναι πρόσθια κατεύθυνσης, τότε τα Νευρωνικά Δίκτυα αυτά χαρακτηρίζονται ως δίκτυα με πρόσθια τροφοδότηση (feedforward). Σε αντίθετη περίπτωση και σε περίπτωση όπου νευρώνες του ίδιου επιπέδου είναι συνδεδεμένοι μεταξύ τους, τότε το Νευρωνικό Δίκτυο χαρακτηρίζεται ως δίκτυο με ανατροφοδότηση (feedback).

### 2.2.3 Εκπαίδευση

Αρχικά, αφού είδαμε τη βασική δομή των Νευρωνικών Δικτύων και τα χαρακτηριστικά των νευρώνων, τώρα θα αναλυθεί η διαδικασία της εκπαίδευσης. Σε αυτή τη φάση το Νευρωνικό Δίκτυο χρησιμοποιεί τα δεδομένα εκπαίδευσης προκειμένου να μάθει από αυτά και να είναι ικανό να εκτελεί την εργασία που του έχει ανατεθεί με όσο μεγαλύτερη ακρίβεια μπορεί. Ουσιαστικά, κατά τη διαδικασία αυτή το Νευρωνικό Δίκτυο ανανεώνει τα βάρη των ακμών του έτσι ώστε το αποτέλεσμα που παράγουν να προσεγγίζει, κατά το δυνατόν, το επιθυμητό. Επομένως, όταν ένα Νευρωνικό Δίκτυο λέμε ότι είναι εκπαιδευμένο σε ένα πρόβλημα, τότε σημαίνει ότι έχει ενημερωμένα τα βάρη του. Έτσι, η βάση γνώσης που περιέχει ένα δίκτυο μπορεί να αναπαρασταθεί από τα βάρη των ακμών του.

Τα Νευρωνικά Δίκτυα πραγματοποιούν δύο βασικές λειτουργίες: τη μάθηση και την ανάκληση (recall). Μάθηση, όπως αναφέρθηκε, είναι η διαδικασία της τροποποίησης της τιμής των βαρών του δικτύου, προκειμένου για ένα διάνυσμα εισόδου να δίνει το επιθυμητό διάνυσμα εξόδου. Ανάκληση είναι η διαδικασία υπολογισμού του διανύσματος εξόδου δεδομένου του διανύσματος εισόδου και των τιμών των βαρών του δικτύου. Στην Επιβλεπόμενη Μάθηση δίνονται στο δίκτυο ζευγάρια εισόδων - επιθυμητών εξόδων και αυτό παράγει, με την τρέχουσα κατάσταση των βαρών, μια έξοδο η οποία ενδεχομένως να διαφέρει από την επιθυμητή. Αυτή η διαφορά ονομάζεται σφάλμα (error) και βάσει αυτής και της χρήσης ειδικών αλγορίθμων γίνεται η αναπροσαρμογή των βαρών του δικτύου. Στην Μη Επιβλεπόμενη Μάθηση, το δίκτυο αυτο-οργανώνεται με βάση τα δεδομένα εισόδου, μιας και δεν υπάρχουν δεδομένα εξόδου σε αυτή την κατηγορία μάθησης.

Στα Νευρωνικά Δίκτυα, κατά τη διαδικασία της εκπαίδευσης, υπάρχει πιθανότητα να εμφανιστούν φαινόμενα υποπροσαρμογής (underfitting) και υπερπροσαρμογής (overfitting). Ένα δίκτυο το οποίο δεν είναι αρκετά περίπλοκο σε σχέση με τη φύση των

δεδομένων εκπαίδευσης, μπορεί να αποτύχει να τα μοντελοποιήσει επιτυχώς. Αντίθετα, αν ένα δίκτυο είναι υπερβολικά περίπλοκο, τότε ενδέχεται να μοντελοποιήσει υπερβολικά τα δεδομένα εκπαίδευσης καθώς και το θόρυβο που ίσως υπάρχει σε αυτά. Σε αυτή την περίπτωση, το δίκτυο μπορεί να προβλέψει σωστά τα δεδομένα εκπαίδευσης αλλά λανθασμένα τα δεδομένα στα οποία δεν έχει εκπαιδευτεί.

Έτσι, είναι βασικό να λαμβάνουμε υπόψη το παραπάνω φαινόμενο σχεδιάζοντας το δίκτυό μας και τα επίπεδά του. Φαινόμενα υπερπροσαρμογής ενδέχεται να εμφανιστούν σε Νευρωνικά Δίκτυα πολλών επιπέδων, ακόμη κι αν τα δεδομένα εκπαίδευσης δεν περιέχουν θόρυβο [1]. Ένας άλλος όμως σημαντικός παράγοντας είναι το μέγεθος των δεδομένων εκπαίδευσης που έχουμε διαθέσιμα. Όσο περισσότερα δεδομένα διαθέτουμε, τόσο λιγότερο πιθανό είναι να έχουμε υποπροσαρμογή και υπερπροσαρμογή στο δίκτυό μας.

Ο πιο συνηθής τρόπος χρήσης των δεδομένων εκπαίδευσης είναι σε κύκλους εκπαίδευσης. Σε κάθε κύκλο, τα δεδομένα χωρίζονται σε δέσμες (batches). Για κάθε δέσμη δεδομένων, το δίκτυο δέχεται ένα-ένα όλα τα διανύσματα εισόδου, υπολογίζει την έξοδο και στη συνέχεια βρίσκει την απόκλισή της από την επιθυμητή έξοδο. Όταν εξετάσει όλα τα δεδομένα της δέσμης, και έχοντας αθροίσει όλες τις αποκλίσεις μεταξύ επιθυμητής και πραγματικής εξόδου, υπολογίζονται τα νέα αναπροσαρμοσμένα βάρη του δικτύου. Όταν προσπελαστούν όλα τα δεδομένα από όλες τις δέσμες, τότε έχει ολοκληρωθεί ένας κύκλος. Αυτή η μέθοδος ονομάζεται **μάθηση δέσμης** (batch learning). Εναλλακτικά, η αναπροσαρμογή των βαρών μπορεί να γίνει μετά από κάθε διάνυσμα εισόδου, αυτή η μέθοδος ονομάζεται **επαυξητική μάθηση** (incremental learning). Οι κύκλοι αυτοί ονομάζονται και **εποχές** (epochs).

Ανεξάρτητα από τον τρόπο και τη μέθοδο με την οποία υλοποιείται η διαδικασία της εκπαίδευσης, τερματίζεται συνήθως με βάση κάποια κριτήρια ελέγχου ποιότητας του δικτύου. Ως τέτοιο κριτήριο συνήθως χρησιμοποιείται το σφάλμα ή η μεταβολή του μέσου σφάλματος του συνόλου κατά την εκπαίδευση. Βέβαια, η εκπαίδευση μπορεί να διακοπεί και μετά από έναν συγκεκριμένο αριθμό εποχών, τον οποίο έχει ορίσει ο προγραμματιστής.

### 2.2.3.1 Συνάρτηση Κόστους

Όπως αναφέραμε, η διαδικασία της εκπαίδευσης περιέχει και τη φάση κατά την οποία ανατροφοδοτούνται τα βάρη του Νευρωνικού Δικτύου. Αυτά τα βάρη αλλάζουν την

τιμή τους με βάση τη διαφορά τρέχουσας εξόδου και επιθυμητής εξόδου σε ένα συγκεκριμένο διάνυσμα εισόδου. Έστω λοιπόν  $\hat{y}$  η τιμή εξόδου του δικτύου για μια είσοδο  $x$ , και  $y$  η επιθυμητή έξοδος που θα θέλαμε το δίκτυο να δίνει στην είσοδο αυτή. Θέλουμε μια συνάρτηση η οποία θα παίρνει αυτές τις δύο τιμές και θα δίνει μια τιμή που δείχνει κατά πόσο η πραγματική με την επιθυμητή έξοδο διαφέρουν. Αυτό είναι βασικό, όχι μόνο για να ξέρουμε με βάση ποια τιμή θα ανατροφοδοτήσουμε τα βάρη αλλά και για να αξιολογήσουμε την απόδοση του δικτύου μας και κατά πόσο αυτό βελτιώνεται κατά τη διάρκεια της εκπαίδευσής του σε κάθε εποχή. Για τον σκοπό αυτό, θα χρησιμοποιήσουμε μια συνάρτηση κόστους (cost function) ή αλλιώς **συνάρτηση σφάλματος** (error function)  $e$ .

Στην απλή περίπτωση της γραμμικής παλινδρόμησης, όπου τα  $\hat{y}$  και  $y$  είναι συνεχή, συνήθως χρησιμοποιούμε μια Μέση Τετραγωνική Απόκλιση (Mean Squared Error - MSE).

$$e(y, \hat{y}) = MSE(y, \hat{y}) = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2.16)$$

Συνεπώς, όταν οι προβλέψεις  $\hat{y}$  βρίσκονται κοντά στην επιθυμητή τιμή εξόδου  $y$  τότε η τιμή της συνάρτησης κόστους είναι μικρή.

Στην περίπτωση της ταξινόμησης, όπου η τιμή εξόδου είναι μια από τις  $K$  κλάσεις του προβλήματος, τότε η έξοδος  $\psi$  κωδικοποιείται σε μορφή one-hot διανυσμάτων. Αυτό σημαίνει ότι κάθε ετικέτα  $y_i$  είναι ένα διάνυσμα με  $K$  δυαδικά στοιχεία, όπου κάθε στοιχείο αναπαριστά μια από τις  $K$  κλάσεις. Το διάνυσμα αυτό περιέχει σε όλες τις θέσεις μηδενική τιμή και στη θέση που αφορά την κλάση που ανήκει το δείγμα παίρνει την τιμή 1.

**Παράδειγμα 2.1** Σε ένα πρόβλημα που έχουμε κατηγορίες ταινιών, έστω ότι οι κατηγορίες είναι οι ακόλουθες: κωμωδία, τρόμου, περιπέτειας και επιστημονικής φαντασίας. Θα αντιστοιχίσουμε τη μεταβλητή *movie\_category* στη μορφή one-hot διανύσματος. Εδώ έχουμε 4 κλάσεις άρα και το διάνυσμα θα έχει 4 τιμές.

$$\text{movie\_category} = \left( \text{κωμωδία} \quad \text{τρόμου} \quad \text{περιπέτειας} \quad \text{επιστημονικής φαντασίας} \right)$$

Έτσι, το αντίστοιχο διάνυσμα one-hot για κάθε κλάση θα είναι το εξής:

$$\begin{pmatrix} \text{κωμωδία} \\ \text{τρόμου} \\ \text{περιπέτειας} \\ \text{επιστημονικής φαντασίας} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Σε τέτοια προβλήματα ταξινόμησης χρησιμοποιούμε συνήθως **σταυροειδή εντροπία** (cross-entropy) για συνάρτηση κόστους. Η συνάρτηση αυτή, όταν έχουμε  $K$  κλάσεις, είναι η εξής:

$$e(y, \hat{y}) = H(y, \hat{y}) = \sum_{i=1}^N \sum_{j=1}^K -y_i^j \log \hat{y}_i^j \quad (2.17)$$

### 2.2.3.2 Διαδικασία Μάθησης

Όπως αναφέρθηκε και προηγουμένως, σε ένα Νευρωνικό Δίκτυο κατά τη διαδικασία της μάθησης, αλλάζουν οι τιμές των παραμέτρων  $p$  του δικτύου που είναι τα βάρη  $W$  και οι πολώσεις  $b$  αυτού. Επομένως, για ένα δίκτυο με  $N$  κόμβους, έχουμε:

$$p = \{(W_1, b_1), (W_2, b_2), \dots, (W_N, b_N)\} \quad (2.18)$$

Οι παράμετροι αυτές είναι πολύ σημαντικές διότι αλλάζοντας τις τιμές τους, επηρεάζονται και οι προβλέψεις του δικτύου. Με την εκπαίδευση του δικτύου, βρίσκουμε τις πιο κατάλληλες τιμές για τις παραμέτρους αυτές με σκοπό την πιο ακριβή πρόβλεψη. Θεωρητικά, ο στόχος είναι η εύρεση των βέλτιστων τιμών  $p$ , έστω  $p^*$ , προκειμένου να ελαχιστοποιηθεί το σφάλμα  $e$ . Επομένως, ισχύει για το  $p^*$  η εξής σχέση:

$$p^* = \arg \min_p e(y, \hat{y}) \quad (2.19)$$

όπου η συνάρτηση  $\arg \min$  ορίζεται:

$$\arg \min_x f(x) := \{x \in S | \forall y \in S : f(y) \geq f(x)\} \quad (2.20)$$

Δεδομένου ότι η διαδικασία εύρεσης των βέλτιστων τιμών  $p$  αναζητώντας όλες τις πιθανές τιμές είναι ανέφικτη, θα εκτελέσουμε μια επαναληπτική διαδικασία για την ελαχιστοποίηση του κόστους. Η διαδικασία αυτή ονομάζεται **κάθοδος μέγιστης κλίσης** (gradient descent). Αν παραγωγίσουμε την συνάρτηση σφάλματος ως προς

κάποια παράμετρο του δικτύου, τότε μπορούμε να καταλάβουμε πώς πρέπει να τροποποιήσουμε την παράμετρο αυτή για να μειωθεί το σφάλμα.

$$p^* = \arg \min_p \nabla_p e(y, \hat{y}) \quad (2.21)$$

Στην παραπάνω σχέση, προκειμένου να βρούμε μια λύση, υποθέτουμε ότι η ελαχιστοποίηση μιας παραγώγου μιας συνάρτησης οδηγεί σε κάποιο ελάχιστο. Μέσω σύνθετων αλγορίθμων προσπαθούμε να βρούμε το ολικό ελάχιστο αλλά κάποιες φορές και ένα τοπικό ελάχιστο αρκεί και είναι αποδεκτή λύση του προβλήματος.

Η επαναληπτική διαδικασία που γίνεται κατά τη διάρκεια της εκπαίδευσης ξεκινάει με την εύρεση της μερικής παραγώγου της συνάρτησης κόστους ως προς κάθε παράμετρο  $W_i$  του δικτύου.

$$\nabla_{W_i} e(y, \hat{y}) = \frac{\partial e(y, \hat{y})}{\partial W_i} \quad (2.22)$$

Στη συνέχεια η νέα τιμή της συγκεκριμένης παραμέτρου θα αλλάξει προς την κατεύθυνση της παραγώγου. Ορίζουμε ως  $\lambda$  το ρυθμό μάθησης (learning rate) και με βάση αυτόν ενημερώνεται η νέα τιμή της παραμέτρου  $W'_i$  η οποία είναι:

$$W'_i = W_i - \lambda \cdot \nabla_{W_i} e(y, \hat{y}) \quad (2.23)$$

Σε διανυσματική μορφή έχουμε:

$$p' = p - \lambda \cdot \nabla_{W_i} e(y, \hat{y}) \quad (2.24)$$

### 2.2.3.3 Αλγόριθμοι Βελτιστοποίησης

Ο αλγόριθμος που περιγράφηκε στην προηγούμενη ενότητα ανήκει σε μια κατηγορία αλγορίθμων οι οποίοι ονομάζονται αλγόριθμοι βελτιστοποίησης, συχνά στα Νευρωνικά Δίκτυα αποκαλούνται ως optimizers ή solvers. Ο παραπάνω αλγόριθμος της καθόδου παραγώγου είναι αρκετά απλός και μη αποδοτικός για τη διαδικασία της μάθησης στα Νευρωνικά Δίκτυα. Αναμφίβολα, υπάρχουν πιο σύνθετοι και αποδοτικοί αλγόριθμοι όπου ο κάθε ένας έχει τα δικά του χαρακτηριστικά, τα οποία πρέπει να λάβουμε υπόψη σχεδιάζοντας ένα Νευρωνικό Δίκτυο.

Αρχικά, ένα πρόβλημα που αντιμετωπίζει η τεχνική της καθόδου μέγιστης κλίσης είναι ότι απαιτεί τον υπολογισμό του σφάλματος ολόκληρου του συνόλου των δεδομένων, όπου για κάθε δείγμα πρέπει να υπολογιστεί η μερική παράγωγος της συνάρτησης κόστους ως προς κάθε μια παράμετρο του νευρωνικού δικτύου. Αυτό για μεγάλο όγκο δεδομένων και για δίκτυα με πολλές παραμέτρους είναι ανέφικτο καθώς απαιτείται τεράστια υπολογιστική ισχύ και χρόνος για να πραγματοποιηθεί. Για τον λόγο αυτό χρησιμοποιείται μια παραλλαγή της τεχνικής αυτής, η Στοχαστική Κάθοδος Κλίσης (Stochastic Gradient Descent - SGD).

$$p' = p - \lambda \cdot \nabla_{w_i} e(y_b, \hat{y}_b), b \subset N \quad (2.25)$$

Η βασική διαφορά από την απλή μέθοδο είναι ότι δε χρειάζεται να υπολογιστούν όλα τα σφάλματα του συνόλου των δεδομένων αλλά τα σφάλματα ενός μικρότερου συνόλου, το οποίο ορίζουμε  $\text{batch}(b)$ , και ο αριθμός των δειγμάτων του καλείται  $\text{batch size}$ . Μόλις φορτωθούν τα δεδομένα ενός  $\text{batch}$ , υπολογίζεται το σφάλμα τους και ύστερα ενημερώνονται τα νέα βάρη του δικτύου. Αυτή η διαδικασία θα λέμε ότι είναι μια επανάληψη (iteration). Όταν προσπελαστούν όλα τα δεδομένα, τότε έχουμε μια εποχή (epoch).

Το πλεονέκτημα της μεθόδου αυτής είναι ότι μπορεί να εφαρμοστεί σε μεγάλο όγκο δεδομένων. Δε χρειάζεται επίσης να φορτωθεί όλο το σύνολο δεδομένων αλλά μόνο μια δέσμη αυτών (batch).

Ένας άλλος αλγόριθμος που θα αναφέρουμε είναι ο Adam. Το όνομά του προέκυψε από την αγγλική φράση adaptive moment estimation, δηλαδή **προσαρμοστική εκτίμηση ροπής**. Ο αλγόριθμος αυτός χρησιμοποιεί δύο άλλους αλγορίθμους (επεκτάσεις του SGD), συνδυάζοντας τα πλεονεκτήματά τους. Συγκεκριμένα τους:

- **Adaptive Gradient Algorithm (AdaGrad)**

Ο αλγόριθμος αυτός διατηρεί μια επιπλέον παράμετρο learning rate η οποία βελτιώνει τη διαδικασία μάθησης

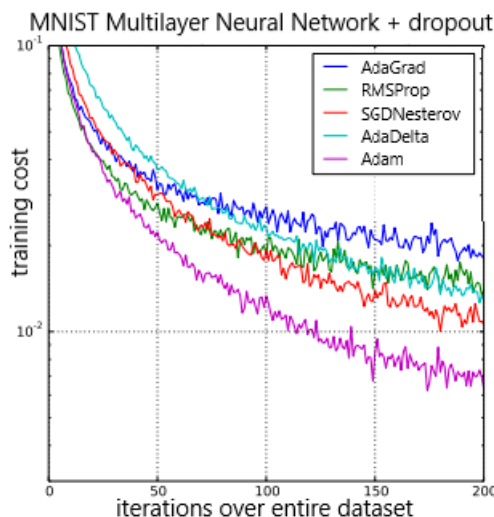
- **Root Mean Square Propagation (RMSProp)**

Εδώ, ο αλγόριθμος διατηρεί και αυτός μια επιπλέον παράμετρο learning rate η οποία είναι βασισμένη στον ρυθμό αλλαγής των βαρών κατά τη διάρκεια της εκπαίδευσης.

Αντί η παράμετρος learning rate του Adam να προσαρμόζεται μόνο στις αλλαγές των βαρών, ο αλγόριθμος ελέγχει και τις τιμές των κλίσεων και υπολογίζει τον μέσο όρο τους. Πιο αναλυτικά, ο αλγόριθμος αυτός υπολογίζει έναν κινούμενο μέσο όρο των παραγώγων και του τετραγώνου των παραγώγων τις οποίες εξομαλύνει με τη χρήση των σταθερών  $\beta_1$  και  $\beta_2$  και μετά χρησιμοποιεί αυτούς τους μέσους όρους για να αναπροσαρμόσει τα βάρη. Ο Adam είναι ένας πολύ αποδοτικός αλγόριθμος καθώς χρησιμοποιεί περισσότερες μεταβλητές και διαχειρίζεται καλύτερα τα δεδομένα εκπαίδευσης. Συγκεκριμένα, ο αλγόριθμος Adam έχει τις εξής παραμέτρους:

- **alpha** Αυτή είναι η αντίστοιχη παράμετρος learning rate. Όσο μεγαλύτερη είναι αυτή, τόσο πιο γρήγορα μεταβάλλονται τα βάρη του δικτύου ενώ όσο μικρότερη είναι, τόσο πιο αργή είναι η διαδικασία της εκπαίδευσης.
- **beta1** Σταθερά εξομάλυνσης 1.
- **beta2** Σταθερά εξομάλυνσης 2.
- **epsilon** Αυτός ο αριθμός είναι πολύ μικρός και χρησιμοποιείται για να αποφύγουμε με τυχόν διαιρέσεις με το μηδέν.

Παρακάτω φαίνεται ένα διάγραμμα το οποίο είναι αποτέλεσμα προσομοιώσεων με διάφορους αλγορίθμους πάνω σε ένα σύνολο δεδομένων.



Σχήμα 2.14: Σύγκριση του Adam με άλλους αλγορίθμους βελτιστοποίησης [20]

Στο παραπάνω διάγραμμα φαίνεται η συνάρτηση κόστους - επαναλήψεις εκπαίδευσης. Είναι προφανές ότι ο Adam μπορεί να δώσει πολύ καλά αποτελέσματα σε σύγκριση με τους υπόλοιπους αλγόριθμους του διαγράμματος. Είναι ένας αλγόριθμος ο οποίος μπορεί να χρησιμοποιηθεί σε προβλήματα βαθιάς μάθησης με σύνθετες και πολύπλοκες αρχιτεκτονικές δικτύων και να ανταπεξέλθει με επιτυχία. Στην παρούσα διπλωματική εργασία, ο συγκεκριμένος αλγόριθμος χρησιμοποιείται συχνά.

#### 2.2.3.4 Μετρικές Απόδοσης Δικτύου

Όπως αναφέρθηκε και στον ορισμό του Mitchell για την TN, πρέπει να ορίσουμε και ένα μέτρο απόδοσης  $P$  με το οποίο θα εξετάζουμε κατά πόσο το δίκτυό μας επιτυγχάνει τον σκοπό του. Σε αυτή την ενότητα θα ορίσουμε τις μετρικές αυτές.

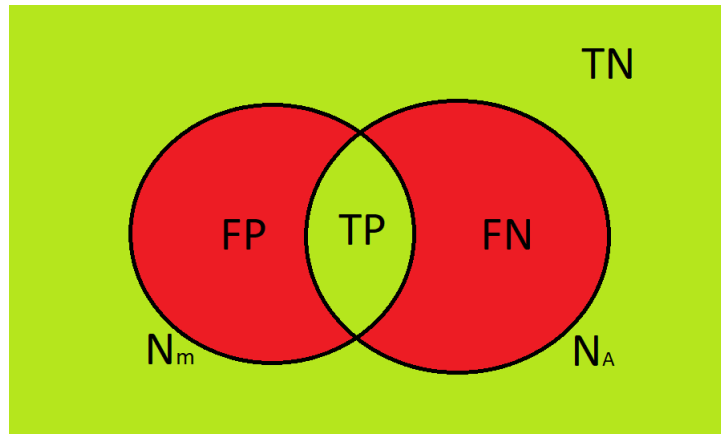
Αρχικά, όμως, πρέπει να αναφερθούν κάποιοι βασικοί όροι, προκειμένου να κατανοηθούν πλήρως οι παρακάτω μετρικές. Έστω ότι σε ένα πρόβλημα κατηγοριοποίησης θέλουμε να μελετήσουμε τις προβλέψεις ενός μοντέλου που αφορούν την κλάση  $A$ . Έχουμε:

- $N_m$  είναι το σύνολο των δειγμάτων που το μοντέλο προέβλεψε ότι ανήκουν στην κλάση  $A$
- $N_A$  είναι το σύνολο των δειγμάτων που ανήκουν στην κλάση  $A$

Ορίζουμε τις εξής έννοιες:

- **True Positive (TP)** είναι το σύνολο των δειγμάτων που ανήκουν και στο σύνολο  $N_m$  και στο  $N_A$ , δηλαδή τα δείγματα που το μοντέλο προέβλεψε σωστά ότι ανήκουν στην κλάση  $A$
- **False Positive (FP)** είναι το σύνολο των δειγμάτων που ανήκουν στο σύνολο  $N_m$  αλλά δεν ανήκουν στο  $N_A$ , δηλαδή τα δείγματα που το μοντέλο προέβλεψε λανθασμένα ότι ανήκουν στην κλάση  $A$
- **False Negative (FN)** είναι το σύνολο των δειγμάτων που δεν ανήκουν στο σύνολο  $N_m$  αλλά ανήκουν στο  $N_A$ , δηλαδή τα δείγματα που το μοντέλο προέβλεψε λανθασμένα ότι δεν ανήκουν στην κλάση  $A$
- **True Negative (TN)** είναι το σύνολο των δειγμάτων που δεν ανήκουν στο σύνολο  $N_m$  ούτε στο  $N_A$ , δηλαδή τα δείγματα που το μοντέλο προέβλεψε σωστά ότι δεν ανήκουν στην κλάση  $A$





Σχήμα 2.15: Σχηματική απεικόνιση συνόλων των εννοιών True Positive, True Negative, False Positive και False Negative. Με πράσινο χρώμα είναι οι σωστές προβλέψεις και με κόκκινο οι λανθασμένες.

Παρακάτω λοιπόν θα ορίσουμε τις μετρικές τις οποίες θα χρησιμοποιούμε για να μετρήσουμε την απόδοση των μοντέλων που θα εξετάσουμε:

- **Ακρίβεια (Accuracy)** Η ακρίβεια είναι μια βασική μετρική η οποία μας δείχνει το ποσοστό από τις σωστές προβλέψεις του μοντέλου μας προς τις συνολικές.

$$Accuracy = \frac{N_{correct}}{N} \quad (2.26)$$

όπου:

- $N_{correct}$  ο αριθμός των σωστών προβλέψεων
- $N$  ο συνολικός αριθμός των προβλέψεων

Σε προβλήματα δυαδικής κατηγοριοποίησης (binary classification), οι κλάσεις που ανήκουν τα δεδομένα είναι αυστηρά δύο, ο τύπος της ακρίβειας είναι ο εξής:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.27)$$

- **Precision** Η μετρική αυτή εκφράζει το ποσοστό των σωστών προβλέψεων του μοντέλου που ανήκουν σε μια κλάση προς τις συνολικές προβλέψεις των δειγμάτων για μια κλάση και υπολογίζεται ως εξής:

$$Precision = \frac{TP}{TP + FP} \quad (2.28)$$

- **Recall** Η μετρική αυτή εκφράζει το ποσοστό των σωστών προβλέψεων του μοντέλου για μια κλάση προς τον συνολικό αριθμό των δειγμάτων που ανήκουν στην κλάση αυτή:

$$Recall = \frac{TP}{TP + FN} \quad (2.29)$$

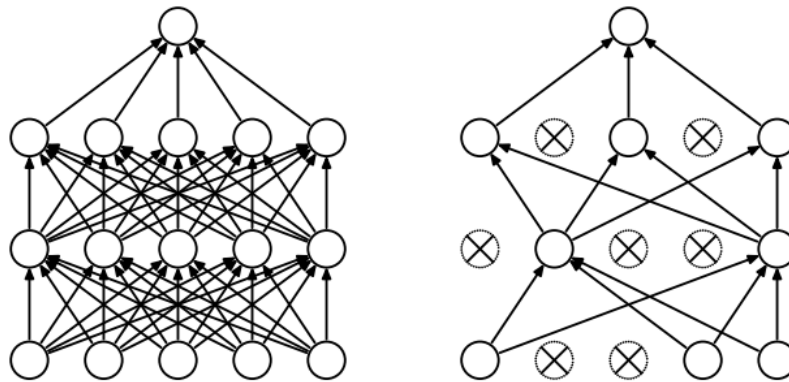
- **F1 score** Η μετρική F1 score είναι ο αρμονικός μέσος των μετρικών Precision και Recall:

$$F1 \text{ score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (2.30)$$

### 2.2.3.5 Τρόποι αποφυγής Overfitting

Όπως αναφέρθηκε παραπάνω, το overfitting είναι ένα συχνό φαινόμενο στην Μηχανική Μάθηση το οποίο μπορεί να οδηγήσει σε εσφαλμένες προβλέψεις του μοντέλου όταν εισάγονται νέα δεδομένα. Η υπερβολική εκμάθηση των δεδομένων εκπαίδευσης, η πολυεπίπεδη και πολύπλοκη αρχιτεκτονική του δικτύου και η λανθασμένη συλλογή δεδομένων εκπαίδευσης είναι κάποιοι από τους κύριους λόγους που το δίκτυό μας μπορεί να οδηγηθεί σε υπερπροσαρμογή κατά τη διάρκεια της εκμάθησης των δεδομένων. Παρακάτω, θα αναλυθούν δύο αρκετά δημοφιλείς μέθοδοι για να αποφευχθεί το overfitting.

- **Απόσυρση (Dropout)** Η τεχνική του dropout είναι μια από τις πιο γνωστές τεχνικές η οποία όχι μόνο αντιμετωπίζει τα φαινόμενα overfitting αλλά αποτελεί και μια μέθοδο να συνενωθούν δύο Νευρωνικά Δίκτυα αποδοτικά. Το όνομα της μεθόδου, απόσυρση, δόθηκε διότι αποσύρει κάποιους κόμβους. Αυτούς τους κόμβους δεν τους διαγράφει οριστικά, αλλά τους αποδεσμεύει από το δίκτυο προσωρινά και έτσι απενεργοποιεί όλες τις ακμές που εισέρχονται και εξέρχονται από τον κόμβο αυτόν, όπως φαίνεται και στο παρακάτω σχήμα. Η επιλογή των κόμβων που θα αποσυρθούν γίνεται τυχαία. Πιο συγκεκριμένα, υπάρχει μια πιθανότητα  $p$  με την οποία κάθε κόμβος συμμετέχει στη διαδικασία της εκπαίδευσης. Ως αποτέλεσμα αυτής της μεθόδου, απενεργοποιώντας και ενεργοποιώντας αντίστοιχα τυχαίους νευρώνες σε κάθε επανάληψη, εξαλείφουμε το φαινόμενο της υπερπροσαρμογής και το δίκτυο μπορεί να διακρίνει τα πραγματικά χαρακτηριστικά της κάθε κλάσης.

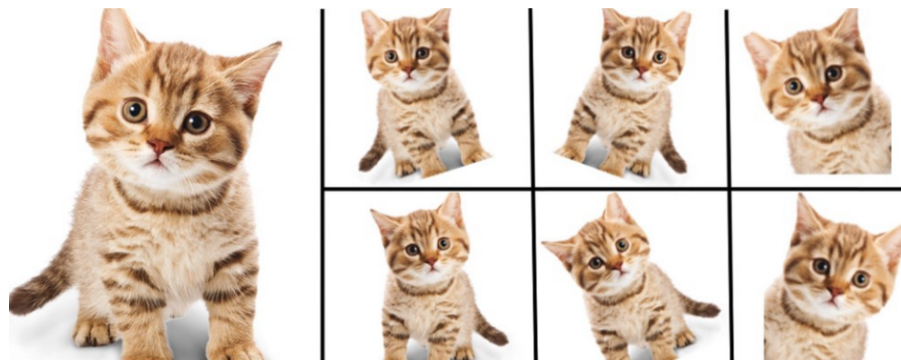


Σχήμα 2.16: Νευρωνικό Δίκτυο πριν και μετά την εφαρμογή dropout.

Αριστερά: Ένα κλασικό μοντέλο Νευρωνικού Δικτύου.

Δεξιά: Το μοντέλο που προκύπτει εφαρμόζοντας τη μέθοδο της απόσυρσης στο αριστερό Νευρωνικό Δίκτυο. [2]

- **Data Augmentation** Η έλλειψη δεδομένων στα δεδομένα εκπαίδευσης είναι μια βασική αιτία για την υπερπροσαρμογή και γενικά για τη λανθασμένη εκπαίδευση του νευρωνικού δικτύου. Επομένως, μεγαλύτερο πλήθος δεδομένων θα βοηθούσε στην αντιμετώπιση των προβλημάτων αυτών. Αυτό είναι εφικτό με την μέθοδο data augmentation. Ας πάρουμε για παράδειγμα την εικόνα μιας γάτας. Ο ανθρώπινος εγκέφαλος βλέποντας μια εικόνα, όπως αυτή της γάτας, μπορεί να τη φανταστεί και από άλλες όψεις και διαφορετικές οπτικές γωνίες και να αντιληφθεί ότι και αυτές είναι εικόνες που απεικονίζουν μια γάτα. Συνεπώς, με την μέθοδο αυτή μπορούμε να δημιουργήσουμε αυτές τις πρόσθετες εικόνες με βάση μια αρχική και να τις προσθέσουμε στα δεδομένα εκπαίδευσης με ετικέτα ίδια με αυτή της αρχικής εικόνας. Γενικά, με τη μέθοδο του data augmentation μπορούμε να δημιουργήσουμε νέα δεδομένα και να τα προσθέσουμε στο dataset βοηθώντας έτσι το μοντέλο να αποφύγει την υπερπροσαρμογή.



Σχήμα 2.17: Δημιουργία νέων εικόνων από την αρχική με τη μέθοδο data augmentation [3]

#### 2.2.4 Συνελικτικά Νευρωνικά Δίκτυα

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks - CNNs ή αλλιώς ConvNets) εισήχθησαν πρώτη φορά από τον Yann LeCun τη δεκαετία του 1990 στο paper με τίτλο “Gradient-Based Learning Applied to Document Recognition” [21]. Πιο συγκεκριμένα, ο LeCun εκπαίδευσε ένα μοντέλο πάνω στο σύνολο δεδομένων MNIST το οποίο αποτελείται από χειρόγραφα ψηφία με ετικέτες  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ .



Σχήμα 2.18: Παράδειγμα χειρόγραφων ψηφίων του MNIST dataset

Το CNN μοντέλο λοιπόν, εκπαιδεύτηκε στα δεδομένα αυτά ώστε να μπορεί να

διακρίνει ποιος αριθμός είναι γραμμένος σε μια εικόνα που του δίνεται ως είσοδος. Τα σημερινά state-of-the-art μοντέλα, τα οποία εκτελούν αυτή τη διεργασία, επιτυγχάνουν σχεδόν τέλεια ακρίβεια στις προβλέψεις τους στο MNIST dataset.

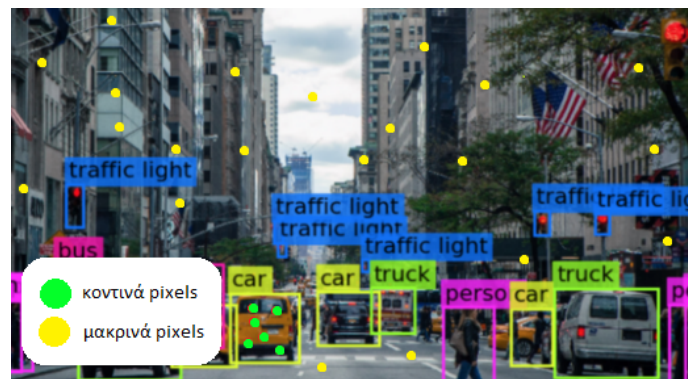
Από τότε πολλοί ερευνητές και επιστήμονες της Μηχανικής Μάθησης έχουν δοκιμάσει να εισαγάγουν τα ConvNets στα μοντέλα τους και παρατηρούν μεγάλες βελτιώσεις στην απόδοσή τους. Τα CNNs έχουν συμβάλει σημαντικά στη σχεδίαση μοντέλων πάνω σε γνωστά προβλήματα Μηχανικής Μάθησης όπως στα:

- CIFAR-10
- CIFAR-100
- VisualGenome
- CelebA

#### 2.2.4.1 Αρχιτεκτονική και ιδιότητες

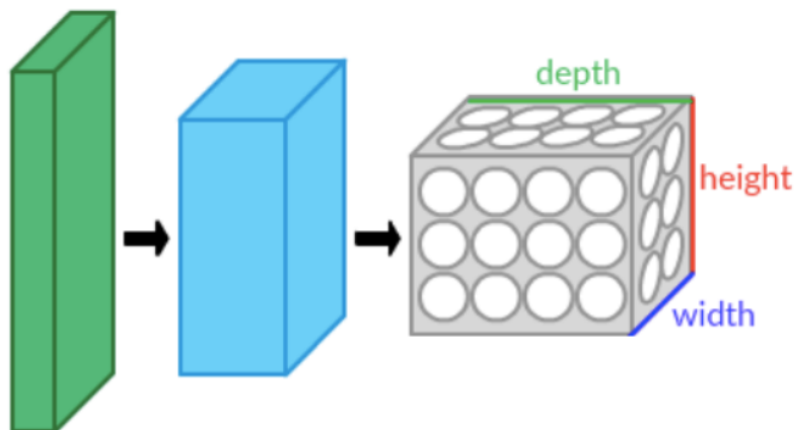
Τα CNNs είναι και αυτά Νευρωνικά Δίκτυα τα οποία αποτελούνται από νευρώνες και διαθέτουν και αυτά παραμέτρους μάθησης, δηλαδή βάρη, πολώσεις κλπ. Και εδώ οι νευρώνες διαθέτουν τη συνάρτηση ενεργοποίησης και το τελευταίο επίπεδο διαθέτει τη συνάρτηση σφάλματος. Η είσοδος που δέχονται τα δίκτυα αυτά, μπορεί να έχει περισσότερες από μια διαστάσεις καθώς έχουν τη δυνατότητα να εφαρμοστούν και σε τέτοιες εισόδους. Τα CNNs παίρνουν το όνομά τους από έναν ειδικό τύπο στρώματος, ο οποίος αποτελεί τον πυρήνα της αρχιτεκτονικής τους, τα convolutional layers. Όπως δηλώνει λοιπόν και το όνομα, τα στρώματα αυτά εκτελούν την πράξη της συνέλιξης με τη χρήση κάποιων φίλτρων πάνω στα δεδομένα και εξάγουν χρήσιμα χαρακτηριστικά από αυτά. Τα φίλτρα αυτά αποτελούν τα βάρη του δικτύου.

Στα CNNs δε συνδέονται όλοι οι νευρώνες ενός επιπέδου με όλες τους τις εισόδους. Αντίθετα, υπάρχει μια τοπική συνδεσιμότητα όπου μόνο οι κοντινοί νευρώνες έχουν ακμές που τους συνδέουν. Αυτή είναι η βασική ιδέα των συνελικτικών δικτύων, δηλαδή, τα γειτονικά δεδομένα συσχετίζονται περισσότερο μεταξύ τους. Για παράδειγμα, σε ένα κείμενο οι κοντινές λέξεις που βρίσκονται μέσα σε μια πρόταση έχουν μεγαλύτερη σχέση μεταξύ τους από ότι με λέξεις άλλων προτάσεων. Σε μια εικόνα τα γειτονικά εικονοστοιχεία (pixels) έχουν μεγαλύτερη σχέση το ένα με το άλλο σε σχέση με τα πιο μακρινά.



Σχήμα 2.19: Παράδειγμα τοπικών χαρακτηριστικών στην εικόνα της ενότητας 1.3. Τα πράσινα εικονοστοιχεία είναι τα κοντινά, τα οποία βλέπουμε ότι αφορούν όλα το ίδιο στοιχείο στην εικόνα, το αυτοκίνητο. Αντίθετα, τα κίτρινα δεν έχουν καμία εννοιολογική σχέση μεταξύ τους και αφορούν διαφορετικά στοιχεία της εικόνας. [4]

Η τυπική αρχιτεκτονική ενός CNN έχει νευρώνες διατεταγμένους σε τρεις διαστάσεις: πλάτος (width), ύψος (height) και βάθος (depth), όπου ο όρος βάθος δεν αναφέρεται στον αριθμό των επιπέδων αλλά στον αριθμό των καναλιών. Το γεγονός αυτό, ότι οι νευρώνες οργανώνονται σε τρισδιάστατους όγκους, σημαίνει ότι θα συνδεθούν μόνο με μια μικρή περιοχή του επόμενου επιπέδου και όχι με όλο το δίκτυο. Αυτό επαληθεύει την ιδιότητα της τοπικότητας των δεδομένων που αναφέραμε.



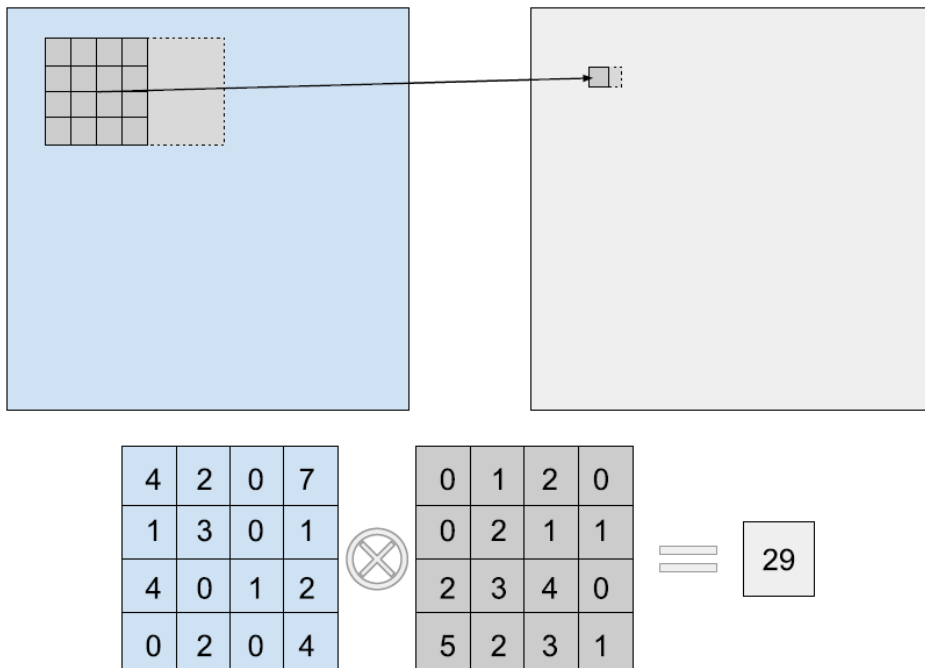
Σχήμα 2.20: Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5]

Τα συνελικτικά δίκτυα, όπως λέει και το όνομά τους, βασίζονται ουσιαστικά στη

μαθηματική πράξη της συνέλιξης. Συνέλιξη είναι μια πράξη που εφαρμόζεται σε δύο συναρτήσεις. Έστω οι διακριτές συναρτήσεις  $\phi, \gamma$ . Ορίζεται ως συνέλιξη των  $\phi, \gamma$  και συμβολίζεται με  $\phi * \gamma$  η συνάρτηση [22]:

$$(\phi * \gamma)(x) = \sum_{\kappa=-\infty}^{+\infty} \phi(\kappa) \cdot \gamma(x - \kappa) \quad (2.31)$$

Η πράξη της συνέλιξης έχει ίδιες ιδιότητες με τον πολλαπλασιασμό, δηλαδή την αντιμεταθετική, την προσεταιριστική και την επιμεριστική ιδιότητα.



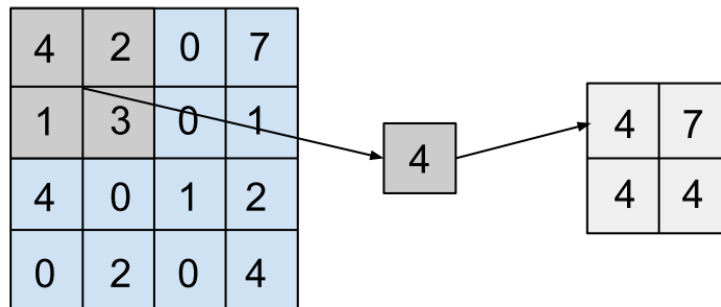
Σχήμα 2.21: Λειτουργία του συνελικτικού επιπέδου. Επάνω φαίνεται ένα φίλτρο το οποίο προχωράει στην εικόνα για να παραγάγει μια τιμή της εξόδου. Κάτω φαίνεται αναλυτικά ένα αριθμητικό παράδειγμα ενός στοιχείου-προς-στοιχείο πολλαπλασιασμού μεταξύ της αρχικής εικόνας και των τιμών του συνελικτικού φίλτρου. [6]

#### 2.2.4.2 Επίπεδα Συμψηφισμού

Τα CNNs χρησιμοποιούν και **επίπεδα ομαδοποίησης** (pooling layers). Τα επίπεδα αυτά εκτελούν υποδειγματοληψία στην είσοδό τους. Περνάνε κι αυτά ένα φίλτρο το

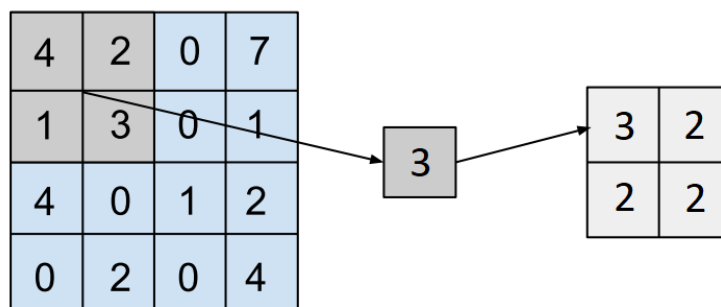
οποίο όμως δε διαθέτει παραμέτρους. Οι δύο πιο συνηθισμένες πράξεις συγκέντρωσης είναι οι εξής:

- **Max Pooling** Εδώ επιλέγεται η μέγιστη τιμή από τους κόμβους



Σχήμα 2.22: Παράδειγμα Επιπέδου Ομαδοποίησης με βάση το Μέγιστο [6]

- **Average Pooling** Εδώ επιλέγεται ο μέσος όρος των τιμών των κόμβων

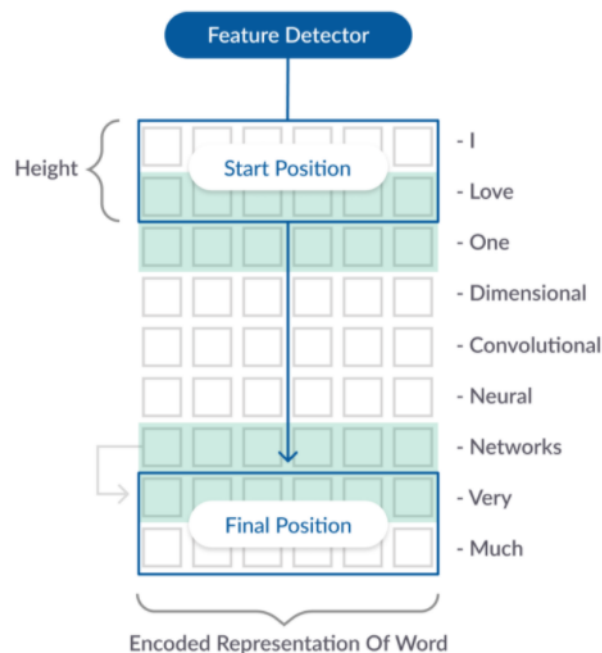


Σχήμα 2.23: Παράδειγμα CNN με βάση το Μέσο Όρο

### 2.2.4.3 Μονοδιάστατα Συνελικτικά Δίκτυα

Τα μονοδιάστατα συνελικτικά δίκτυα έχουν εφαρμογή στην επεξεργασία φυσικής γλώσσας. Ουσιαστικά το φίλτρο είναι μονοδιάστατο και μπορεί να εφαρμοστεί σε πίνακες οι οποίοι περιέχουν λέξεις, να εντοπίσει και να διακρίνει χαρακτηριστικά που το αφορούν. Τα δίκτυα αυτά είναι μονοδιάστατα μιας και τα φίλτρα κινούνται μόνο προς μια διάσταση πάνω στα δεδομένα εισόδου.





Σχήμα 2.24: Παράδειγμα συνελικτικού δικτύου μιας διάστασης. [7]

### 2.2.5 Αναδρομικά Νευρωνικά Δίκτυα

Αυτή η ενότητα ασχολείται με τα Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks ή για συντομία RNNs), τα οποία είναι πολύ σημαντικά για την Επεξεργασία Φυσικής Γλώσσας. Τα δίκτυα αυτά και η δομή τους μοιάζουν με τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου, όπου οι διαδικασίες της σκέψης και της αντίληψης είναι μη γραμμικές. Τα δίκτυα αυτά περιέχουν αναδρομές ανάμεσα σε κόμβους των ίδιων ή διαφορετικών επιπέδων. Αυτή η ιδιότητα τους προσδίδει δύο βασικά χαρακτηριστικά:

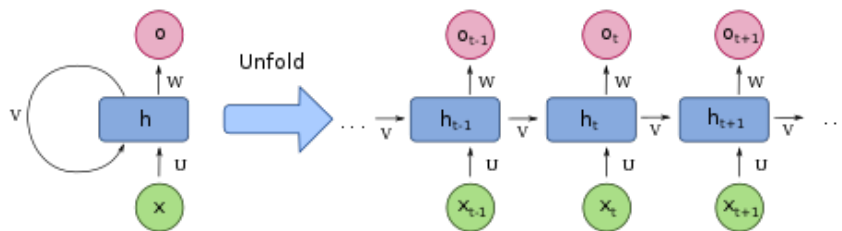
- Το πρώτο είναι η ικανότητά τους να έχουν αντίληψη του χρόνου και να μπορούν να διαχειρίζονται χρονοσειρές με μεγάλη επιτυχία
- Το δεύτερο είναι το χαρακτηριστικό της μνήμης, το οποίο μπορεί να χρησιμεύσει σε προβλήματα που υπάρχει συσχέτιση μεταξύ των δειγμάτων

#### 2.2.5.1 Αρχιτεκτονική και ιδιότητες

Τα RNNs κατηγοριοποιούνται σε πλήρως αναδρομικά, στα οποία υπάρχει πλήρης διασύνδεση των κόμβων και στα τοπικά αναδρομικά δίκτυα. Τα πλήρως αναδρομικά παρουν-

σιάζουν προβλήματα στη διαδικασία μάθησης και θέματα ευστάθειας στην εκπαίδευση. Τα τοπικά, αντιθέτως, με την πιο απλή δομή τους εκπαιδεύονται ταχύτερα και πιο αποδοτικά.

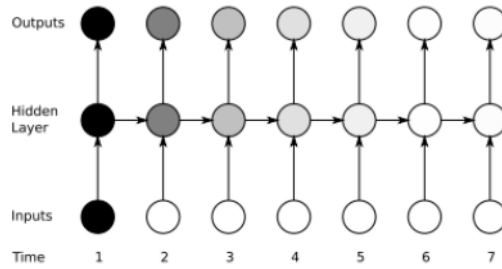
Τα βασικά RNN οργανώνονται σε κόμβους και επίπεδα. Κάθε κόμβος που ανήκει σε κάποιο επίπεδο συνδέεται με τους κόμβους του επόμενου επιπέδου. Κάθε κόμβος διαθέτει δύο ενεργοποιήσεις, τη χρονική και την πραγματική. Κάθε σύναψη μεταξύ κόμβων διαθέτει ένα πραγματικό βάρος. Στην επιβλεπόμενη μάθηση, η είσοδος στο δίκτυο έχει τη μορφή πίνακα (vector) και εισέρχεται στο δίκτυο μέσω των κόμβων εισόδου, ένας πίνακας τη φορά. Κάθε χρονική στιγμή, όλοι οι κόμβοι υπολογίζουν την ενεργοποίησή τους ως μια μη γραμμική συνάρτηση η οποία περιλαμβάνει το άθροισμα όλων των ενεργοποιήσεων των συνδεδεμένων με αυτόν κόμβων. Κάθε χρονική στιγμή  $t$ , οι κόμβοι διαθέτουν μια τιμή ενεργοποίησης, η οποία αναπαρίσταται με  $x_t$  για τους κόμβους εισόδου,  $h_t$  για τους εσωτερικούς και  $o_t$  για τους κόμβους εξόδου.



Σχήμα 2.25: Απεικόνιση ενός κόμβου από ένα RNN. Από κάτω προς τα πάνω: επίπεδο εισόδου - κρυφό επίπεδο - επίπεδο εξόδου και  $\{U, V, W\}$  είναι τα βάρη του δικτύου. Αριστερά φαίνεται ο κόμβος και δεξιά το ανάπτυγμά του. [8]

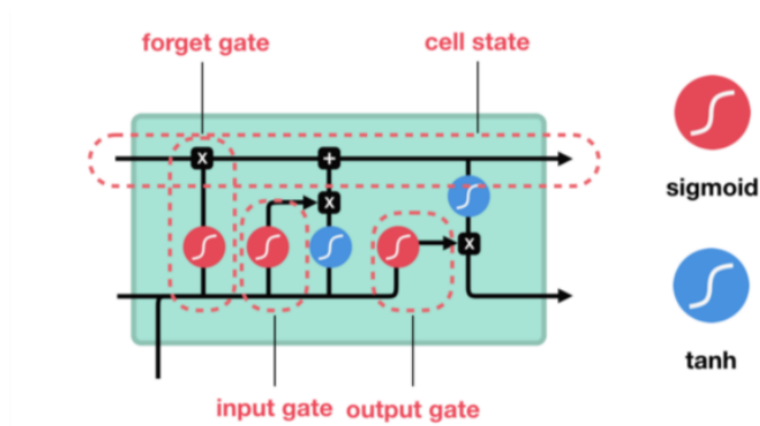
### 2.2.5.2 Long Short Term Memory Δίκτυα

Τα Long Short Term Memory, ή εν συντομία LSTM δίκτυα είναι μια κατηγορία αναδρομικών δικτύων, πολύ δημοφιλή στον κλάδο της Μηχανικής Μάθησης. Όπως προδίδει και το όνομά τους, έχουν την ιδιότητα της μνήμης, την οποία διαθέτουν από προηγούμενα δεδομένα και τη χρησιμοποιούν στην εκπαίδευση των επόμενων. Στα απλά RNNs η πληροφορία όσο πιο παλιά είναι, τόσο περισσότερο εξασθενεί. Για παράδειγμα, η πληροφορία που υπάρχει από το δεύτερο δεδομένο του πίνακα δεδομένων της χρονοσειράς επηρεάζει πολύ λιγότερο τον έκτο κόμβο παρά τον τρίτο.



Σχήμα 2.26: Το πρόβλημα της εξασθένησης των συσχετίσεων στα RNN με την πάροδο του χρόνου. [23]

Τα LSTM με την αρχιτεκτονική τους, έχουν τη δυνατότητα να διατηρούν την πληροφορία ακόμα και για μεγάλα διαστήματα χρόνου. Αυτό επιτυγχάνεται με τις επιπλέον μονάδες που βρίσκονται στο εσωτερικό των κόμβων του. Πιο συγκεκριμένα, κάθε μονάδα LSTM αποτελείται από ένα ή περισσότερα κελιά μνήμης που συνδέονται μεταξύ τους και τρία επιπλέον στοιχεία, τις πύλες εισόδου (input gates), εξόδου (output gates) και τις πύλες επιλεκτικής συγκράτησης (forget gates), οι οποίες είναι υπεύθυνες για λειτουργίες εγγραφής, ανάγνωσης και επαναφοράς κόμβων.



Σχήμα 2.27: Εσωτερική δομή μιας μονάδας LSTM [24]

## Κεφάλαιο 3

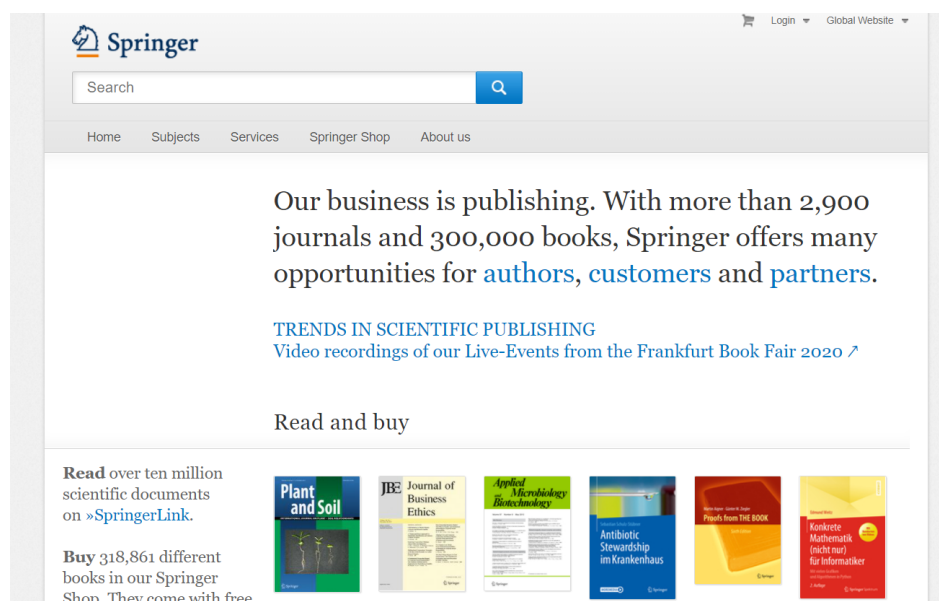
# Παρουσίαση, Επεξεργασία και Ανάλυση Δεδομένων

Σε αυτό το κεφάλαιο γίνεται μια συνοπτική παρουσίαση των δεδομένων τα οποία θα επεξεργαστούν και θα χρησιμοποιηθούν στην εργασία. Τα δεδομένα ανακτήθηκαν από τη βάση δεδομένων του Springer και ο διαχωρισμός τους σε κατηγορίες πραγματοποιήθηκε ανεξάρτητα. Εδώ θα πραγματοποιηθεί όλη η διαδικασία της προεπεξεργασίας των δεδομένων και της κωδικοποίησής τους προκειμένου να χρησιμοποιηθούν ως είσοδος για τη διαδικασία της εκπαίδευσης. Επιπλέον, γίνονται κάποιες οπτικοποιήσεις μέσω των οποίων διακρίνεται κατά πόσο οι διαφορετικές κατηγορίες διαφοροποιούνται ως προς το περιεχόμενό τους.

### 3.1 Βάση δεδομένων Springer

Η βάση δεδομένων που χρησιμοποιείται σε αυτή τη διπλωματική εργασία, αποτελείται από 56401 ηλεκτρονικά συγγράμματα, τα οποία έχουν ανακτηθεί από το ηλεκτρονικό αποθετήριο Springer. Τα βιβλία αυτά, αφού γίνει η κατάλληλη επεξεργασία, θα χρησιμοποιηθούν ως δεδομένα για την εκπαίδευση των μοντέλων που θα υλοποιηθούν.

Το ηλεκτρονικό αποθετήριο Springer είναι ευρέως γνωστό για την τεράστια ποικιλία και τον όγκο συγγραμμάτων που διαθέτει.



Σχήμα 3.1: Αρχική σελίδα του ηλεκτρονικού αποθετηρίου Springer. [9]

Για κάθε ένα βιβλίο υπάρχουν διαθέσιμα πολλά διαφορετικά πεδία που το αφορούν. Αυτά τα πεδία θα πρέπει να επεξεργαστούν και να φιλτραριστούν, έτσι ώστε να είναι εφικτό να χρησιμοποιηθούν ως διανύσματα χαρακτηριστικών για τα βιβλία. Αυτά τα χαρακτηριστικά είναι που θα τροφοδοτούν τα συστήματα Μηχανικής Μάθησης για να βγάλουν αποτελέσματα.

### 3.2 Κατηγοριοποίηση Δεδομένων

Στην επιβλεπόμενη μάθηση τα δεδομένα πρέπει να διαθέτουν κάποια ετικέτα από ένα ορισμένο σύνολο ετικετών. Στην περίπτωση μας, πρέπει να δοθεί στο κάθε σύγγραμμα

μια ετικέτα που το αντιπροσωπεύει. Έτσι, θα κατηγοριοποιήσουμε τα βιβλία σε ένα σύνολο κατηγοριών που σχετίζονται με το θεματικό τους περιεχόμενο.

Τα βιβλία είναι ήδη κατηγοριοποιημένα από τη βάση δεδομένων που διαθέτουμε σε 252 διαφορετικές κατηγορίες. Μπορούμε να ορίσουμε μια διαδικασία με την οποία να ομαδοποιήσουμε τις κατηγορίες αυτές σε λιγότερες και πιο γενικές. Αυτή η διαδικασία πραγματοποιήθηκε από το paper [25] και με τη βοήθεια βιβλιοθηκονόμων του ΕΜΠ. Επομένως, τις νέες 26 κλάσεις για κάθε βιβλίο τις διαθέτουμε έτοιμες και είναι οι εξής:

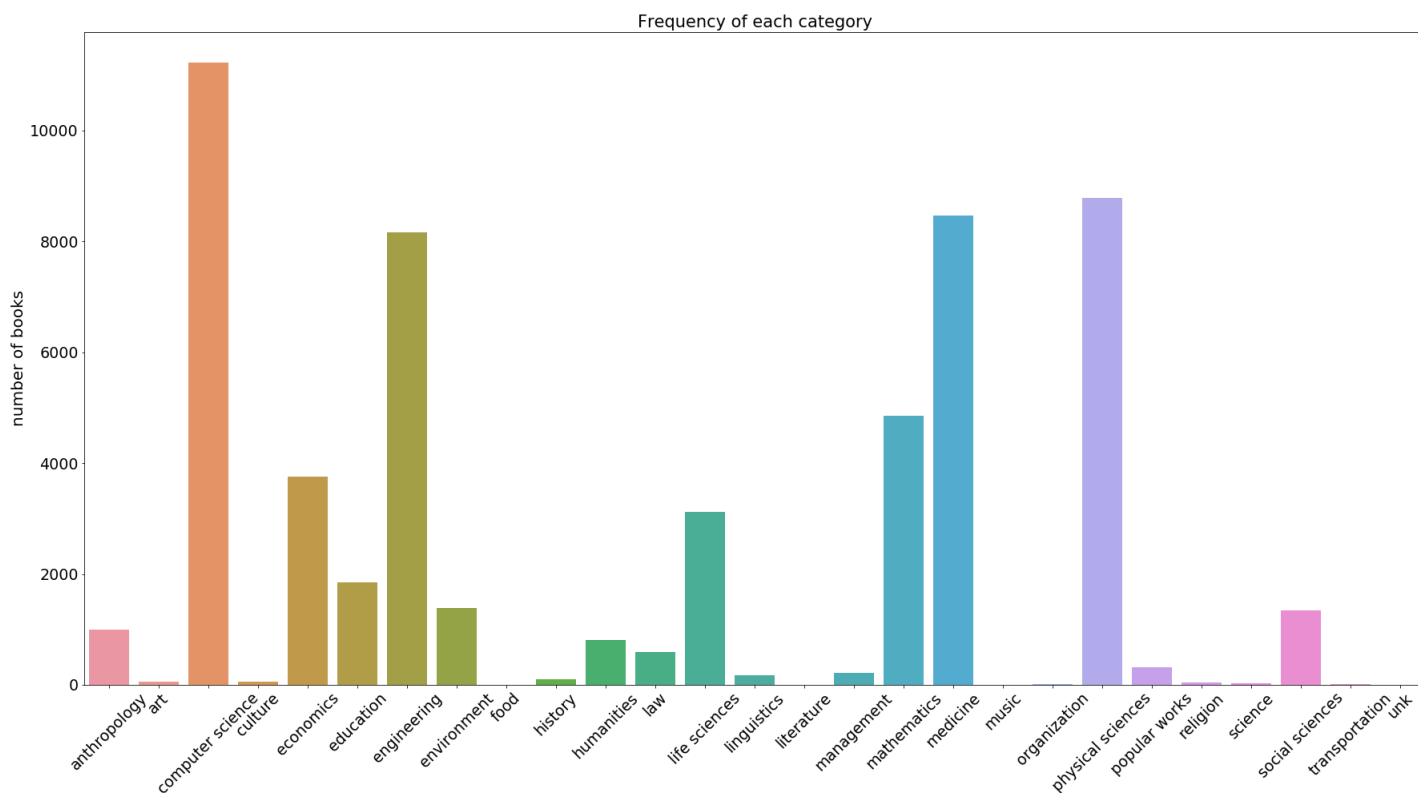
- |  |  |
|--|--|
| 1. Ανθρωπολογία (Anthropology)             | 14. Γλωσσολογία (Linguistics)              |
| 2. Τέχνη (Art)                             | 15. Λογοτεχνία (Literature)                |
| 3. Επιστήμη Υπολογιστών (Computer Science) | 16. Διαχείριση (Management)                |
| 4. Κουλτούρα (Culture)                     | 17. Μαθηματικά (Mathematics)               |
| 5. Οικονομικά (Economics)                  | 18. Ιατρική (Medicine)                     |
| 6. Εκπαίδευση (Education)                  | 19. Μουσική (Music)                        |
| 7. Μηχανική (Engineering)                  | 20. Οργάνωση (Organisation)                |
| 8. Περιβάλλον (Environment)                | 21. Φυσικές Επιστήμες (Physical Sciences)  |
| 9. Φαγητό (Food)                           | 22. Δημοφιλή Έργα (Popular Works)          |
| 10. Ιστορία (History)                      | 23. Θρησκεία (Religion)                    |
| 11. Ανθρωπιστικές Επιστήμες (Humanities)   | 24. Κοινωνικές Επιστήμες (Social Sciences) |
| 12. Νομικά (Law)                           | 25. Επιστήμη (Science)                     |
| 13. Βιοεπιστήμες (Life Sciences)           | 26. Μεταφορά (Transportation)              |

Οι συχνότητες εμφάνισης των 56401 δειγμάτων για τις παραπάνω 26 κλάσεις είναι η εξής:

Κατηγορία	Αριθμός Δειγμάτων	Κατηγορία	Αριθμός Δειγμάτων
Ανθρωπολογία	997	Γλωσσολογία	172
Τέχνη	53	Λογοτεχνία	5
Επιστήμη Υπολογιστών	11219	Διαχείριση	221
Κουλτούρα	57	Μαθηματικά	4852
Οικονομικά	3762	Ιατρική	8471
Εκπαίδευση	1842	Μουσική	1
Μηχανική	8161	Οργάνωση	11
Περιβάλλον	1394	Φυσικές Επιστήμες	8788
Φαγητό	4	Δημοφιλή Έργα	322
Ιστορία	105	Θρησκεία	38
Ανθρωπιστικές Επιστήμες	815	Κοινωνικές Επιστήμες	1344
Νομικά	599	Επιστήμη	34
Βιοεπιστήμες	3124	Μεταφορά	10
<b>Σύνολο</b>		<b>56401</b>	

Πίνακας 3.1: Αριθμός δειγμάτων ανά κατηγορία.

Και το αντίστοιχο διάγραμμα φαίνεται παρακάτω:



Σχήμα 3.2: Συχνότητες εμφάνισης δειγμάτων για τις 26 κλάσεις.

Παρατηρούμε ότι κάποιες κατηγορίες έχουν πολύ περισσότερα βιβλία σε σχέση με κάποιες άλλες. Αυτό μπορεί να οδηγήσει στο φαινόμενο υπερεκπαίδευσης των μοντέλων που θα δημιουργήσουμε στις κλάσεις με τα περισσότερα δεδομένα. Αυτό θα έχει ως αποτέλεσμα το σύστημα να κάνει λάθος προβλέψεις για δείγματα που ανήκουν στις κλάσεις με λιγότερα δείγματα. Επιπρόσθετα, παρατηρούμε ότι κάποιες κατηγορίες συγχέονται, δηλαδή ένα βιβλίο μπορεί να ανήκει και στην κλάση “Μηχανική” και στην κλάση “Μαθηματικά”. Έτσι, θα στραφούμε στις πιο δημοφιλείς και γενικές κατηγορίες με τις οποίες θα ασχοληθούμε, χωρίς όμως να παραβλέψουμε και τις υπόλοιπες και να απορρίψουμε τα αντίστοιχα δείγματα.



Οι κύριες κατηγορίες λοιπόν φαίνονται στον παρακάτω πίνακα:

Κατηγορία	Αριθμός Δειγμάτων
Επιστήμη Υπολογιστών	11219
Φυσικές Επιστήμες	8788
Ιατρική	8471
Μηχανική	8161
Μαθηματικά	4852
Οικονομικά	3762
Βιοεπιστήμες	3124
<b>Σύνολο</b>	<b>48385</b>

Πίνακας 3.2: Αριθμός δειγμάτων ανά κατηγορία για τις 7 επιλεγμένες κατηγορίες.

Ο συνολικός αριθμός των δειγμάτων που ανήκουν στις παραπάνω 7 κατηγορίες είναι 48385, δηλαδή περίπου το 86% όλων των δειγμάτων που διαθέτουμε.

### 3.3 Ανάλυση και Προεπεξεργασία Δεδομένων

Το πρώτο βήμα που πρέπει να γίνει, προτού πραγματοποιηθεί η διαδικασία της επεξεργασίας, είναι η ανάλυση των δεδομένων. Στα προβλήματα μηχανικής Μάθησης η ανάλυση των δεδομένων είναι πολύ σημαντική και μπορεί να δώσει χρήσιμες πληροφορίες για τη γενική εικόνα και τα χαρακτηριστικά τους αλλά και τον τρόπο με τον θα γίνει η επεξεργασία τους.

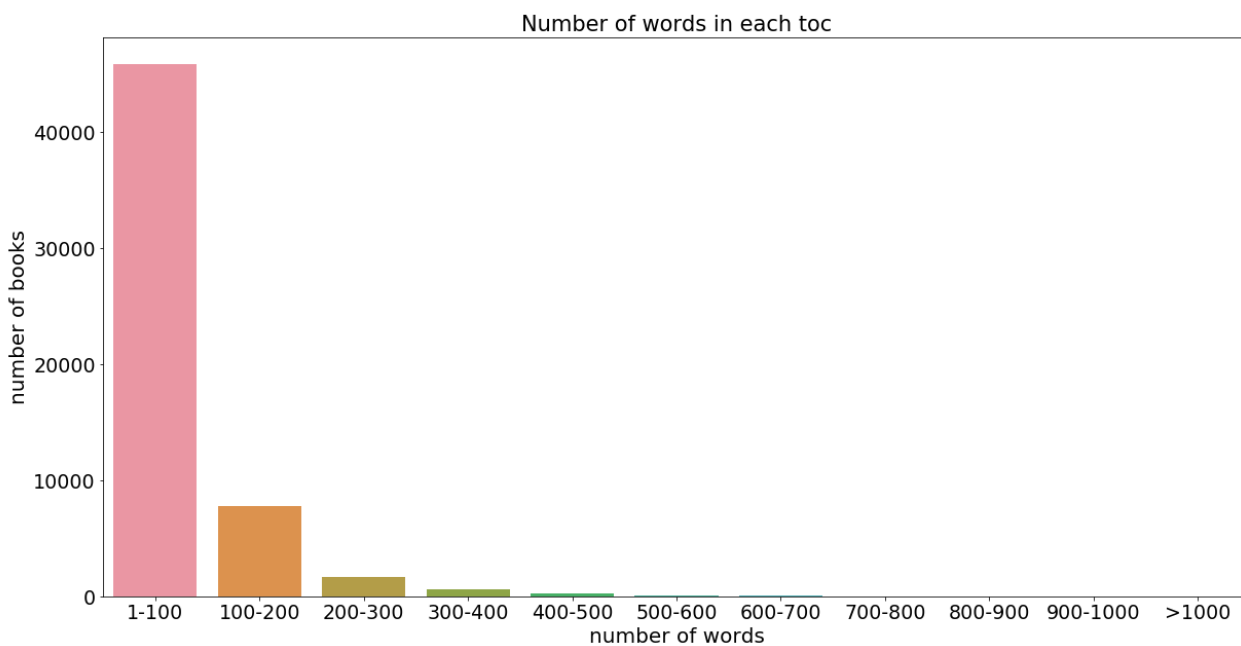
#### 3.3.1 Χαρακτηριστικά Δεδομένων

Τα δεδομένα που διαθέτουμε, όπως αναφέρθηκε, είναι δεδομένα από τη βάση δεδομένων του Springer. Είναι ένα σύνολο 56401 ηλεκτρονικών βιβλίων όπου το κάθε ένα από αυτά έχει κάποια χαρακτηριστικά που το αντιπροσωπεύουν. Τα χαρακτηριστικά λοιπόν που θα ανακτήσουμε από τη βάση δεδομένων για κάθε βιβλίο είναι τα εξής:

- τον τίτλο του βιβλίου
- τον πίνακα περιεχομένων του

- την περίληψη
- τις λέξεις κλειδιά που αφορούν το θεματικό του περιεχόμενο.

Το πιο βασικό χαρακτηριστικό των βιβλίων, το οποίο θα χρησιμοποιήσουμε για την κατηγοριοποίησή τους, είναι ο πίνακας περιεχομένων (table of contents). Επειδή έχει σημασία για τον σχεδιασμό των μοντέλων, θα μελετήσουμε τον αριθμό των λέξεων που περιέχει κάθε ένα σύγγραμμα στον πίνακα περιεχομένων του.



Σχήμα 3.3: Αριθμός λέξεων στον πίνακα περιεχομένων του κάθε βιβλίου.

### 3.3.2 Προεπεξεργασία Δεδομένων

Το στάδιο αυτό αποτελεί το στάδιο της προεπεξεργασίας δεδομένων. Αφού τελειώσει η διαδικασία αυτή, τα δεδομένα μας θα είναι στην κατάλληλη μορφή για να εισαχθούν στα μοντέλα βαθιάς μάθησης που πρόκειται να σχεδιαστούν στο επόμενο κεφάλαιο.

Όπως αναφέρθηκε, στην παρούσα διπλωματική, ως χαρακτηριστικό κάθε βιβλίου θα χρησιμοποιηθεί ο πίνακας περιεχομένων του. Δηλαδή τα μοντέλα επιβλεπόμενης και

μη επιβλεπόμενης μάθησης θα δέχονται ως είσοδο τα περιεχόμενα του ηλεκτρονικού συγγράμματος για να εξάγουν αποτελέσματα για αυτό.

Ας δούμε την αρχική μορφή που έχει ένας πίνακας περιεχομένων ενός βιβλίου το οποίο ανήκει στην κατηγορία “Μαθηματικά”:

*1. Introduction / 2. Lie Groups / 3. Theory of Moving Frames / 4. Euclidean Geometry / 5. Spherical Geometry / 6. Hyperbolic Geometry / 7. Complex Structure / 8. Minimal Immersions in Euclidean Space / 9. Isothermic Immersions / 10. The Bonnet Problem / 11. CMC 1 Surfaces in  $H^3$  / 12. Mbius Geometry / 13. Complex Structure and Mbius Geometry / 14. Isothermic Immersions in Mbius Space / 15. Lie Sphere Geometry / Solutions to Select Problems / References / Index.*

Παρατηρούμε ότι, τα δεδομένα του πίνακα περιεχομένων είναι μια λίστα από τους τίτλους των κεφαλαίων του αντίστοιχου βιβλίου. Τα δεδομένα αυτά θα πρέπει να επεξεργαστούν κατάλληλα έτσι ώστε να αποτυπώνουν το ακριβές περιεχόμενο των κεφαλαίων τους και τη θεματική ενότητα στην οποία εντάσσονται.

### 3.3.2.1 Αλγόριθμος Προεπεξεργασίας Δεδομένων

Θα πρέπει λοιπόν να ορισθεί ένας αλγόριθμος ο οποίος θα επεξεργάζεται τα δεδομένα εισόδου για να τα μετατρέψει στην επιθυμητή μορφή. Βασική ιδέα είναι η έξοδος να είναι μια λίστα από λέξεις κλειδιά που αντιπροσωπεύουν τον αντίστοιχο πίνακα περιεχομένων. Ο αλγόριθμος προεπεξεργασίας δεδομένων ο οποίος ως είσοδο δέχεται τον πίνακα περιεχομένων ενός βιβλίου, ίδιας μορφής με αυτόν που παρουσιάστηκε προηγουμένως, περιγράφεται παρακάτω:

- **Βήμα 1** Το πρώτο βήμα του αλγορίθμου είναι να χωρίσει τον πίνακα περιεχομένων σε λέξεις και να δημιουργήσει μια λίστα από όλες τις λέξεις που περιέχουν όλα τα ξεχωριστά κεφάλαια. Επιπλέον, όλα τα ειδικά σύμβολα, τα σημεία στίξης και οι αριθμητικοί χαρακτήρες θα αφαιρεθούν καθώς δεν προσφέρουν κάποια χρήσιμη πληροφορία. Το αποτέλεσμα του παραπάνω αρχικού μη επεξεργασμένου πίνακα περιεχομένων, μετά την εφαρμογή του πρώτου βήματος, είναι το εξής:

*['Introduction', 'Lie', 'Groups', 'Theory', 'of', 'Moving', 'Frames', 'Euclidean', 'Geometry', 'Spherical', 'Geometry', 'Hyperbolic', 'Ge-*

*ometry', 'Complex', 'Structure', 'Minimal', 'Immersion', 'in', 'Euclidean', 'Space', 'Isothermic', 'Immersion', 'The', 'Bonnet', 'Problem', 'CMC', 'Surfaces', 'in', 'H', 'Mbius', 'Geometry', 'Complex', 'Structure', 'and', 'Mbius', 'Geometry', 'Isothermic', 'Immersion', 'in', 'Mbius', 'Space', 'Lie', 'Sphere', 'Geometry', 'Solutions', 'to', 'Select', 'Problems', 'References', 'Index']*

- **Βήμα 2** Το επόμενο βήμα του αλγορίθμου είναι σε κάθε μια από τις λέξεις της λίστας, όλοι οι χαρακτήρες να μετατραπούν σε πεζούς έτσι ώστε να μην υπάρχει διαφορετική γραφή ανάμεσα σε ίδιες λέξεις αλλά πλήρης ομοιογένεια στους χαρακτήρες από τους οποίους απαρτίζεται, π.χ. το σύστημα θα πρέπει να αντιλαμβάνεται ότι οι τύποι “ανάλυση”, “Ανάλυση” ή “ΑΝΑΛΥΣΗ” αφορούν την ίδια λέξη. Αυτό θα χρησιμεύσει αργότερα όταν σε κάθε λέξη θα δοθεί ένα αναγνωριστικό id προκειμένου να εισέλθει ως είσοδος στο μοντέλο. Επιπλέον, αφαιρούνται και οι λέξεις-χαρακτήρες οι οποίες αποτελούνται από έναν μόνο χαρακτήρα. Το αποτέλεσμα του βήματος 2 φαίνεται παρακάτω:

*['introduction', 'lie', 'groups', 'theory', 'of', 'moving', 'frames', 'euclidean', 'geometry', 'spherical', 'geometry', 'hyperbolic', 'geometry', 'complex', 'structure', 'minimal', 'immersion', 'in', 'euclidean', 'space', 'isothermic', 'immersion', 'the', 'bonnet', 'problem', 'cmc', 'surfaces', 'in', 'mbius', 'geometry', 'complex', 'structure', 'and', 'mbius', 'geometry', 'isothermic', 'immersion', 'in', 'mbius', 'space', 'lie', 'sphere', 'geometry', 'solutions', 'to', 'select', 'problems', 'references', 'index']*

- **Βήμα 3** Το τρίτο και τελευταίο βήμα του αλγορίθμου έχει σκοπό να αφαιρέσει λέξεις οι οποίες δεν έχουν σχέση με το περιεχόμενο του βιβλίου, είναι γενικές και ενδεχομένως να επηρεάσουν αρνητικά την πρόβλεψη των μοντέλων που θα αναπτυχθούν. Οι απαγορευτικές λέξεις λοιπόν που θα αφαιρεθούν από τη λίστα είναι:

- κάποιες γενικές λέξεις που περιέχει ένας πίνακας περιεχομένων ανεξαρτήτως του θεματικού περιεχομένου του βιβλίου

*['introduction', 'conclusion', 'chapter', 'section', 'part']*

– και κάποιες γενικές συνδετικές λέξεις

[’doing’, ’in’, ’ma’, ”mustn’t”, ’was’, ’too’, ’this’, ’they’, ”you’re”, ’yours’, ’once’, ’most’, ’re’, ’as’, ”hasn’t”, ’theirs’, ’now’, ’were’, ’each’, ’more’, ’will’, ’is’, ’has’, ’aren’, ’it’, ’themselves’, ”you’ve”, ’which’, ’her’, ”needn’t”, ’from’, ’between’, ”you’ll”, ’nor’, ”won’t”, ’your’, ’ain’, ’for’, ’some’, ’wasn’, ’being’, ’their’, ’ourselves’, ’me’, ’you’, ’herself’, ’before’, ’few’, ’shan’, ’against’, ’his’, ’shouldn’, ’such’, ”you’d”, ’o’, ’he’, ’am’, ’an’, ’only’, ’above’, ’ve’, ’been’, ’mightn’, ’s’, ’to’, ’hers’, ’into’, ”mightn’t”, ’by’, ”aren’t”, ’didn’, ’m’, ’of’, ’haven’, ’no’, ’again’, ’wouldn’, ’that’, ’below’, ’himself’, ’under’, ”shouldn’t”, ’the’, ’its’, ’so’, ’off’, ”don’t”, ’couldn’, ’y’, ’both’, ’very’, ’and’, ’why’, ’i’, ’weren’, ’are’, ’on’, ’those’, ”didn’t”, ’doesn’, ’d’, ’when’, ’ours’, ’same’, ’further’, ’own’, ’don’, ’hadn’, ’won’, ”wasn’t”, ’t’, ”wouldn’t”, ’did’, ’over’, ’other’, ’a’, ’him’, ’then’, ’any’, ’my’, ’mustn’, ”it’s”, ’be’, ’after’, ’just’, ”isn’t”, ’during’, ’hasn’, ’needn’, ”should’ve”, ’there’, ’where’, ’at’, ’have’, ’should’, ’through’, ”haven’t”, ”weren’t”, ’what’, ’up’, ’does’, ’while’, ’them’, ’how’, ’yourself’, ’all’, ’ll’, ’we’, ’myself’, ’with’, ”she’s”, ’yourselves’, ’isn’, ’itself’, ’because’, ’here’, ’she’, ’had’, ’can’, ’than’, ”shan’t”, ’our’, ’who’, ’if’, ’these’, ”that’ll”, ’having’, ’out’, ’do’, ’down’, ”doesn’t”, ’but’, ’about’, ’until’, ”couldn’t”, ”hadn’t”, ’not’, ’or’, ’whom’]

Έτσι, μετά και την εφαρμογή αυτού του βήματος, έχουμε την τελική μορφή του επεξεργασμένου πίνακα περιεχομένων:

[’lie’, ’groups’, ’theory’, ’moving’, ’frames’, ’euclidean’, ’geometry’, ’spherical’, ’geometry’, ’hyperbolic’, ’geometry’, ’complex’, ’structure’, ’minimal’, ’immersions’, ’euclidean’, ’space’, ’isothermic’, ’immersions’, ’bonnet’, ’problem’, ’cmc’, ’surfaces’, ’mbius’, ’geometry’, ’complex’, ’structure’, ’mbius’, ’geometry’, ’isothermic’, ’immersions’, ’mbius’, ’space’, ’lie’, ’sphere’, ’geometry’, ’solutions’, ’select’, ’problems’, ’references’, ’index’]

### 3.3.3 Word Embeddings

Τα μοντέλα βαθιάς μάθησης και τα Νευρωνικά Δίκτυα δεν μπορούν να δεχθούν ως είσοδο λέξεις και χαρακτήρες. Τα συστήματα αυτά δέχονται ως είσοδο αριθμητικές τιμές. Εφόσον τα δεδομένα εισόδου είναι λέξεις θα πρέπει να βρεθεί ένας τρόπος προκειμένου οι λέξεις να μετατραπούν σε αριθμούς.

Τα word embeddings είναι ουσιαστικά διανύσματα λέξεων, κάθε λέξη δηλαδή μπορεί να αναπαρασταθεί με ένα διάνυσμα. Γιατί όμως να χρησιμοποιηθούν διανύσματα για κάθε λέξη ενώ θα μπορούσε να χρησιμοποιηθεί απλά ένα αναγνωριστικό id και να μετασχηματιστεί έτσι ο πίνακας περιεχομένων σε αριθμητικό διάνυσμα εισόδου;

Σκοπός είναι να μετατρέψουμε τις λέξεις σε αριθμητικές τιμές έτσι ώστε να μεταφερθεί όλη η πληροφορία που περιέχουν οι πίνακες περιεχομένων. Τα word embeddings κωδικοποιούν την έννοια της κάθε λέξης σε ένα διάνυσμα έτσι ώστε οι λέξεις με παρόμοια σημασία ή με κοντινό θεματικό περιεχόμενο να βρίσκονται πιο κοντά στον διανυσματικό χώρο. Δηλαδή το διάνυσμα θέσης της λέξης “χρώμα” είναι πιο κοντά με το αντίστοιχο διάνυσμα θέσης της λέξης “κόκκινο” παρά με αυτό της λέξης “μαθηματικά”.



Σχήμα 3.4: Παράδειγμα αναπαράστασης διανυσμάτων λέξεων σε δισδιάστατο χώρο.

Αν πάρουμε για παράδειγμα τον πίνακα περιεχομένων που επεξεργάστηκε ο αλγόριθμος προεπεξεργασίας δεδομένων, το αποτέλεσμα περιέχει μια λίστα από 41 λέξεις. Εφαρμόζοντας την κωδικοποίηση με χρήση των διανυσμάτων λέξεων, έστω διανύσματα 20 διαστάσεων, τότε ο δισδιάστατος πίνακας με διαστάσεις  $41 \times 20$ , δηλαδή [αριθμός λέξεων]  $\times$  [διαστάσεις διανύσματος λέξης], εμπεριέχει όλη την πληροφορία των λέξεων και μπορεί να αποτελέσει το διάνυσμα εισόδου που αντιπροσωπεύει το αντίστοιχο

βιβλίο.

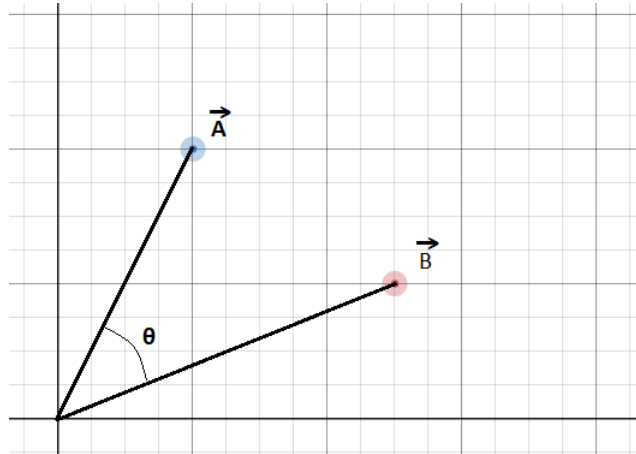
Στην εργασία αυτή θα χρησιμοποιήσουμε τον αλγόριθμο word2vec της Google [26] ο οποίος περιέχει ένα εκπαιδευμένο Νευρωνικό Δίκτυο το οποίο δέχεται ως είσοδο μία λέξη και ως έξοδο δίνει έναν πίνακα μήκους 300 θέσεων με αριθμητικές τιμές. Το δίκτυο αυτό έχει εκπαιδευτεί πάνω σε πολλά δεδομένα και περιέχει 3 εκατομμύρια λέξεις στις οποίες μπορεί να δώσει τον αντίστοιχο πίνακα. Ο πίνακας αυτός αποτελεί και το διάνυσμα της αντίστοιχης λέξης. Επιλέχθηκε ο αλγόριθμος αυτός διότι τα αποτελέσματα που δίνει είναι θετικά και η βάση γνώσης του είναι τεράστια. Έτσι, μπορεί να μεταφέρει όλη την απαραίτητη πληροφορία στο δίκτυό μας.

Το μοντέλο word2vec έχει πολλές δυνατότητες με τις οποίες είναι εφικτό να ελεγχθεί και η απόδοσή του. Δίνοντάς του μια λέξη, για παράδειγμα την λέξη “car”, και ζητώντας του να βρει τις πέντε πιο “κοντινές”, δίνει τα εξής αποτελέσματα:

Λέξη	Ομοιότητα
vehicle	0.7821096181869507
cars	0.7423830032348633
SUV	0.7160962820053101
minivan	0.6907036304473877
truck	0.6735789775848389

Πίνακας 3.3: Αποτελέσματα όμοιων λέξεων της λέξης “car”.

Στην πρώτη στήλη βρίσκεται η λέξη και στη δεύτερη η ομοιότητα συνημιτόνων του διανύσματος της λέξης αυτής με αυτό της αρχικής λέξης. Η ομοιότητα δύο διανυσμάτων λέξεων μετράται με τον υπολογισμό της ομοιότητας συνημιτόνων (cosine similarity). Η ομοιότητα συνημιτόνων είναι μια μετρική ομοιότητας μεταξύ δύο μη μηδενικών διανυσμάτων. Έστω τα διανύσματα  $\vec{A}$  και  $\vec{B}$ . Η ομοιότητα συνημιτόνων των δύο αυτών διανυσμάτων είναι το συνημίτονο της γωνίας  $\theta$  που σχηματίζουν μεταξύ τους. Όσο μεγαλύτερη είναι η ομοιότητα τόσο περισσότερο συγκλίνουν και τα διανύσματα. Είναι φανερό ότι η μέγιστη τιμή είναι 1 και η ελάχιστη 0 καθώς είναι συνημίτονο γωνίας.



Σχήμα 3.5: Συχνότητες εμφάνισης δειγμάτων για τις 26 κλάσεις.

Ο γενικός τύπος της ομοιότητας συνημιτόνων των διανυσμάτων  $\vec{A} = (A_1, A_2, \dots, A_n)$  και  $\vec{B} = (B_1, B_2, \dots, B_n)$  μπορεί να υπολογιστεί με τη βοήθεια του τύπου του εσωτερικού τους γινομένου:

$$\vec{A} \cdot \vec{B} = \|\vec{A}\| \cdot \|\vec{B}\| \cdot \cos(\theta) \quad (3.1)$$

και είναι ο εξής:

$$\text{cosine\_similarity} = \cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| \cdot \|\vec{B}\|} = \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n A_i^2} \cdot \sqrt{\sum_{i=1}^n B_i^2}} \quad (3.2)$$

### 3.3.4 Οπτικοποιήσεις Δεδομένων

Πλέον με τη βοήθεια του εκπαιδευμένου μοντέλου word2vec της Google, είναι εφικτό κάθε πίνακας περιεχομένων  $N$  λέξεων να μετατραπεί σε έναν διδιάστατο αριθμητικό πίνακα  $N \times 300$  ο οποίος είναι ικανός να τον αντιπροσωπεύσει. Περαιτέρω, εάν δημιουργηθεί ένα διάνυσμα για κάθε βιβλίο, αντίστοιχα με το διάνυσμα λέξης, τότε θα μπορούσε να γίνει μια οπτικοποίηση των βιβλίων, με βάση τα διανύσματά τους, και να εξαχθούν χρήσιμα συμπεράσματα προτού γίνει ο σχεδιασμός των μοντέλων.

Ένας τρόπος που μπορεί να επιτευχθεί αυτό είναι με την εφαρμογή μέσου όρου των διανυσμάτων όλων των λέξεων που περιέχει ένας πίνακας περιεχομένων. Έστω

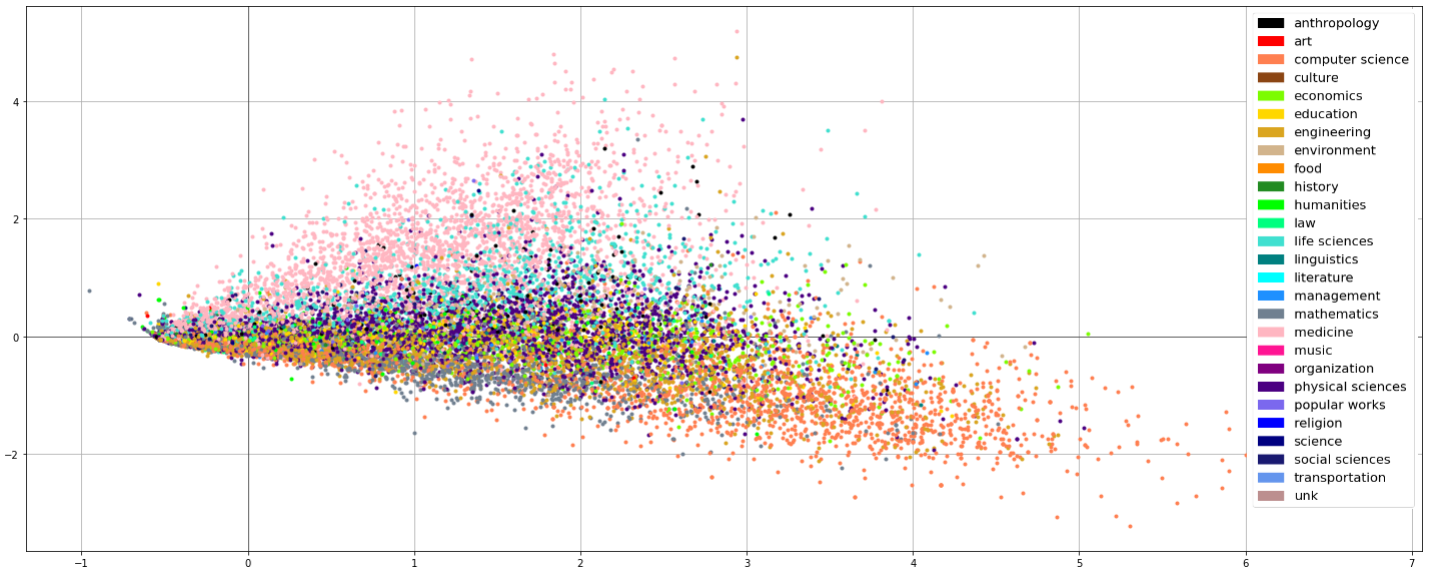


ότι  $A$  είναι ο πίνακας  $N \times 300$  ενός βιβλίου. Επομένως, ο μονοδιάστατος πίνακας  $A[i]$  είναι και το διάνυσμα λέξης  $\vec{A}[i]$  της  $i$ -στής λέξης του πίνακα περιεχομένων. Ορίζεται  $\vec{C} = (C_1, C_2, \dots, C_{300})$  ως το διάνυσμα του αντίστοιχου βιβλίου και μπορεί να υπολογιστεί ως εξής:

$$C_i = \frac{\sum_{j=1}^N A[j][i]}{N} \quad (3.3)$$

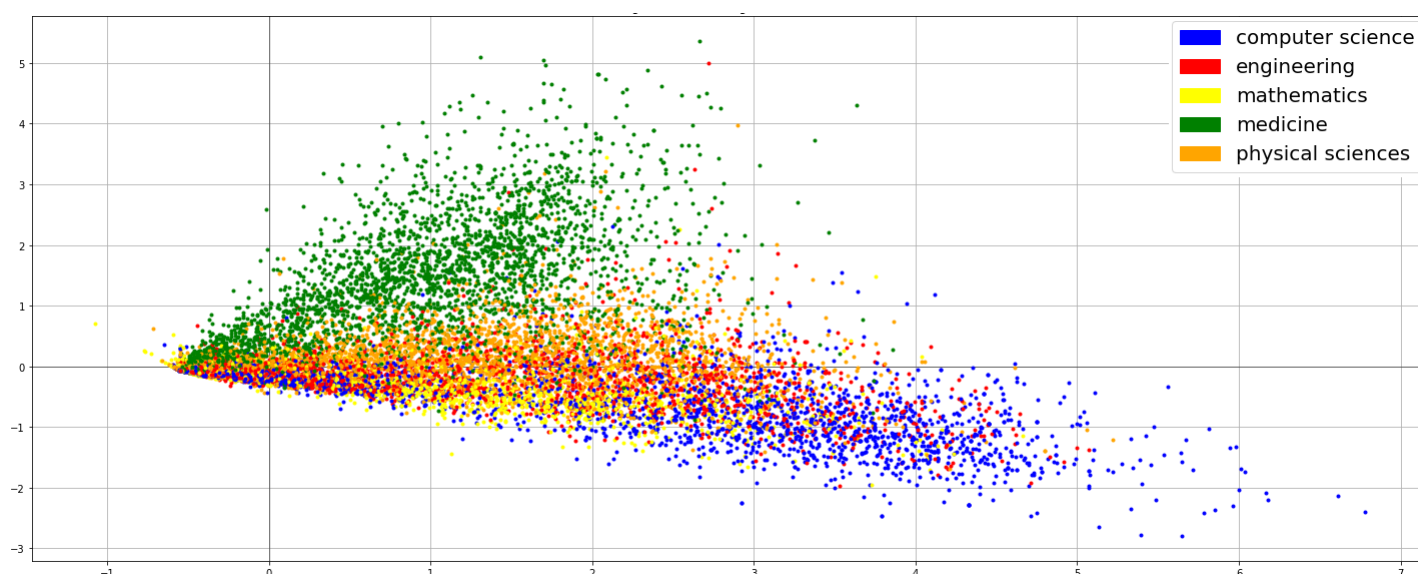
Μέχρι στιγμής λοιπόν κάθε βιβλίο διαθέτει το δικό του διάνυσμα, έναν μονοδιάστατο πίνακα 300 θέσεων. Παρόλα αυτά, είναι αδύνατο να οπτικοποιηθούν όλα τα βιβλία στον 300-στατο χώρο. Θα πρέπει λοιπόν να χρησιμοποιηθεί μια τεχνική με την οποία οι 300 διαστάσεις του κάθε διανύσματος θα μετατραπούν σε 2 και έτσι θα είναι εφικτή η οπτικοποίηση των βιβλίων στον δισδιάστατο χώρο.

Θα χρησιμοποιηθεί η μέθοδος Principal Component Analysis (PCA) [27] της βιβλιοθήκης SKlearn της python. Η PCA πραγματοποιεί μείωση της γραμμικής διάστασης με χρήση της μεθόδου Αποσύνθεσης Ιδιαζουσών Τιμών (Singular Value Decomposition ή εν συντομία SVD). Εφαρμόζοντας την PCA μέθοδο, πλέον το κάθε βιβλίο αποτελείται από ένα διάνυσμα 2 διαστάσεων και έτσι είναι εφικτή η αναπαράσταση αυτών των διανυσμάτων στον δισδιάστατο χώρο.



Σχήμα 3.6: Οπτικοποίηση όλων των βιβλίων στον δισδιάστατο χώρο.

Στο παραπάνω διάγραμμα κάθε κατηγορία έχει το δικό της χρώμα. Παρόλο που η οπτικοποίηση έγινε με 56401 βιβλία, δηλαδή σημεία στο καρτεσιανό επίπεδο, τα οποία ανήκουν σε μια από τις 26 κατηγορίες, με 26 διαφορετικά χρώματα, μπορούμε να διακρίνουμε μια ομαδοποίηση των βιβλίων με κοινό θεματικό περιεχόμενο. Αυτό είναι θετικό διότι με τον τρόπο που κατανέμονται οι διαφορετικές κλάσεις στον χώρο με βάση τα χαρακτηριστικά τους, είναι εφικτός ο σχεδιασμός κάποιου μοντέλου επιβλεπόμενης μάθησης με πρόβλεψη κατηγορίας αλλά και μη επιβλεπόμενης μάθησης για δημιουργία επιπλέον ομάδων (clusters). Επειδή οι 26 κατηγορίες είναι πολλές σε συνδυασμό με τον μεγάλο όγκο δεδομένων, θα γίνει και μια οπτικοποίηση των πιο “δημοφιλών” πέντε κλάσεων.



Σχήμα 3.7: Οπτικοποίηση των βιβλίων που ανήκουν σε μια από τις 5 πιο δημοφιλείς κατηγορίες στον δισδιάστατο χώρο.

Στο παραπάνω διάγραμμα βλέπουμε τα βιβλία που ανήκουν στις επιμέρους κλάσεις: Επιστήμη Υπολογιστών, Μηχανική, Μαθηματικά, Ιατρική και Φυσικές Επιστήμες. Το σύνολο των βιβλίων αυτών είναι 41.491, δηλαδή περίπου το 74% των συνολικών δεδομένων. Εδώ διακρίνεται καλύτερα ο διαχωρισμός των κατηγοριών, με τις κλάσεις της “Μηχανικής” και των “Μαθηματικών” να αλληλοκαλύπτονται με κάποιες άλλες, όπως είναι αναμενόμενο, διότι για παράδειγμα ένα βιβλίο της κλάσης “Επιστήμη των Υπολογιστών” μπορεί εύλογα να ανήκει και στην κλάση της “Μηχανικής”.

## Κεφάλαιο 4

# Εφαρμογές Βαθιάς Μάθησης και Κατασκευή Μοντέλων

Στο τρέχον κεφάλαιο θα μελετηθεί το κύριο θέμα της διπλωματικής εργασίας. Περιέχει όλη την ανάλυση, την έρευνα και τη μελέτη της Επιβλεπόμενης Μάθησης, τη διερεύνηση των διάφορων αρχιτεκτονικών μέχρι και την επιλογή των τελικών μοντέλων για ταξινόμηση ηλεκτρονικών συγγραμμάτων. Στη συνέχεια, με τη βοήθεια των τελικών αυτών μοντέλων γίνονται οπτικοποιήσεις των ενδιάμεσων επιπέδων και επιλέγονται με βάση αυτά, νέα χαρακτηριστικά για κάθε βιβλίο. Κατόπιν, έρχεται η σειρά της Μη Επιβλεπόμενης Μάθησης η οποία δέχεται τα βιβλία με τα νέα χαρακτηριστικά τους, τα αναλύει και δημιουργεί νέα clusters. Τέλος, αξιοποιώντας τα νέα χαρακτηριστικά των βιβλίων και με τη βοήθεια κάποιων τεχνικών, δημιουργείται ένα σύστημα εύρεσης όμοιων βιβλίων για ένα συγκεκριμένο σύγγραμμα το οποίο θα χρησιμοποιηθεί στην web εφαρμογή του Μέρους II της διπλωματικής εργασίας.

## 4.1 Σχεδιασμός Απλών Μοντέλων

Στην πρώτη εφαρμογή βαθιάς μάθησης στα δεδομένα θα δοκιμαστούν κάποια απλά μοντέλα Επιβλεπόμενης Μάθησης τα οποία ως είσοδο θα δέχονται τον πίνακα περιεχομένων ενός βιβλίου και θα το ταξινομούν σε μια από τις γνωστές κατηγορίες που περιγράφηκαν στο προηγούμενο κεφάλαιο.

### 4.1.1 Προετοιμασία Δεδομένων

Τα δεδομένα θα εισαχθούν στην αρχική τους μορφή, δηλαδή σαν μια λίστα από λέξεις χωρίς την προεπεξεργασία που πραγματοποιήθηκε σε αυτά. Με αυτόν τον τρόπο θα διαπιστωθεί στη συνέχεια αν αυτή η προεπεξεργασία βοήθησε τελικά τα μοντέλα ή όχι. Κάθε λέξη έχει το δικό της μοναδικό id και επομένως η είσοδος είναι μια λίστα από αριθμούς (τα αντίστοιχα ids των λέξεων).

Το σύνολο των δεδομένων θα διαιρευθεί στα παρακάτω μέρη:

#### 1. Δεδομένα εκπαίδευσης (train)

Τα δεδομένα αυτά θα χρησιμοποιηθούν από το μοντέλο για τη διαδικασία της εκπαίδευσης. Αποτελούν συνήθως το 80% των συνολικών δεδομένων.

#### 2. Δεδομένα validation

Τα δεδομένα αυτά θα χρησιμοποιηθούν από το μοντέλο για τον έλεγχο της απόδοσης του κάθε εποχής. Αποτελούν συνήθως το 10% των συνολικών δεδομένων.

#### 3. Δεδομένα δοκιμής (test)

Τα δεδομένα αυτά θα χρησιμοποιηθούν για τη δοκιμή της απόδοσης του τελικού εκπαιδευμένου μοντέλου. Αποτελούν συνήθως το 10% των συνολικών δεδομένων.

Επομένως τα 56401 βιβλία θα χωριστούν τυχαία ως εξής:

1. **train:** 45121 βιβλία

2. **validation:** 5640 βιβλία

3. **test:** 5640 βιβλία

Στα προβλήματα Μηχανικής Μάθησης καλό είναι αρχικά να δοκιμάζονται απλά μοντέλα για να ορισθούν οι βάσεις των μετρικών απόδοσης και να φαίνεται σε πιο σύνθετα μοντέλα αν υπάρχει πραγματική βελτίωση ή όχι.

#### 4.1.2 Δοκιμή Απλών Μοντέλων

Το πρώτο μοντέλο είναι ένα απλό μοντέλο το οποίο περιέχει μόνο ένα Dense ενδιάμεσο επίπεδο και το επίπεδο εξόδου.

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Dense	(None, 100)	Relu	100100
Dense	(None, 26)	Softmax	2727

Πίνακας 4.1: Classification Μοντέλο 1

Το τελευταίο dense επίπεδο του μοντέλου έχει μήκος 26, όσες και οι κλάσεις του προβλήματος, και χρησιμοποιεί για συνάρτηση ενεργοποίησης την softmax για να βρει ποια ταιριάζει περισσότερο στο αντίστοιχο βιβλίο. Η έξοδος κάθε επιπέδου περιέχει και μια επιπλέον διάσταση μήκους None η οποία αφορά το μέγεθος δέσμης (batch size) που χρησιμοποιείται στην εκπαίδευση. Εκπαιδύοντας το συγκεκριμένο μοντέλο για 5 εποχές και batch size 32 παρατηρείται ότι δεν υπάρχει βελτίωση στις μετρικές απόδοσης ακρίβειας και κόστους. Η ακρίβεια του μοντέλου μετά την εκπαίδευσή του είναι μόλις 19%.

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	Ακρίβεια
Adam	categorical crossentropy	19%
RMSProp	categorical crossentropy	19%

Πίνακας 4.2: Αποτελέσματα Μοντέλου 1

Το μοντέλο αυτό δε διαθέτει ούτε την απαραίτητη πληροφορία των λέξεων μιας και δε χρησιμοποιεί τα word embeddings που περιγράφηκαν στο Κεφάλαιο 3.

Στο επόμενο μοντέλο δε θα χρησιμοποιηθεί το word2vec της Google για δημιουργία των word embeddings αλλά ένα απλό επίπεδο το οποίο δίνει τυχαία αρχικά διανύσματα για κάθε διαφορετική λέξη.

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 64)	-	100100
LSTM	(None, 64)	-	33024
Dense	(None, 32)	{Relu, Sigmoid}	2727
Dense	(None, 26)	Softmax	914

Πίνακας 4.3: Classification Μοντέλο 2

Σε κάθε βιβλίο δίνονται 64 χαρακτηριστικά μέσω του Embedding επιπέδου. Παρατηρείται ότι οι διαστάσεις του κάθε δείγματος αυξάνονται κατά μια. Για να μειωθούν οι διαστάσεις και να μπορέσουν να εισαχθούν τα Dense επίπεδα, θα εισαχθεί ένα RNN και πιο συγκεκριμένα LSTM επίπεδο στο μοντέλο. Εκτός από την Relu θα δοκιμασθεί και η σιγμοειδής συνάρτηση ενεργοποίησης. Το μοντέλο εκπαιδεύεται για 8 εποχές και μέγεθος δέσμης 32 και τα αποτελέσματα είναι εμφανώς καλύτερα.

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	Συνάρτηση Ενεργοποίησης	Ακρίβεια
Adam	categorical crossentropy	Sigmoid	51%
RMSProp	categorical crossentropy	Sigmoid	49%
Adam	categorical crossentropy	Relu	57%
RMSProp	categorical crossentropy	Relu	54%

Πίνακας 4.4: Αποτελέσματα Μοντέλου 2

## 4.2 Δοκιμές και Πειραματικά Αποτελέσματα Σύνθετων Μοντέλων

Σε αυτό το κεφάλαιο θα δοκιμαστούν μοντέλα με περισσότερα και πιο σύνθετα επίπεδα μέχρι να καταλήξουμε στο τελικό. Στη θέση του RNN θα τοποθετηθεί ένα απλό pooling επίπεδο. Εφόσον η έξοδος ενός τέτοιου επιπέδου δε ρίχνει τις διαστάσεις των δεδομένων εισόδου κατά μια, θα χρησιμοποιηθεί ένα pooling επίπεδο. Επιπλέον, θα χρησιμοποιηθεί το word2vec της Google για τα word embeddings. Να σημειωθεί ότι εφεξής τα δεδομένα εισόδου είναι οι προεπεξεργασμένοι πίνακες περιεχομένων του Κεφαλαίου 3.

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Global Max Pooling 1D	(None, 300)	-	0
Dense	(None, 128)	Relu	38528
Dense	(None, 64)	Relu	8256
Dense	(None, 26)	Softmax	1755

Πίνακας 4.5: Classification Μοντέλο 3

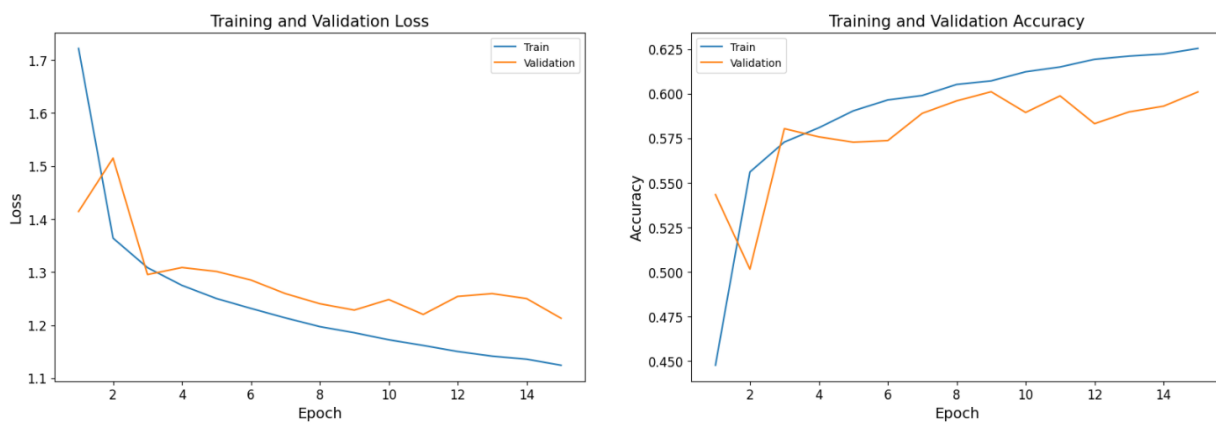
Το Embedding επίπεδο είναι ήδη εκπαιδευμένο (pretrained). Περιέχει 900.000.000 παραμέτρους οπότε φαίνεται πόσο μεγάλη είναι εσωτερικά η αρχιτεκτονική του και πόση βάση γνώσης διαθέτει. Το επίπεδο αυτό δίνει σε κάθε ένα από τα δεδομένα, 300 χαρακτηριστικά, όπως αναλύθηκε και στο προηγούμενο κεφάλαιο. Στην εκπαίδευση θα απενεργοποιηθεί η εκπαίδευση αυτών των παραμέτρων καθώς τα βάρη τους είναι ήδη μορφοποιημένα και επιπλέον τόσες πολλές μεταβλητές προς εκπαίδευση θα καθυστερήσουν αρκετά την όλη διαδικασία της μάθησης.

Το μοντέλο εκπαιδεύεται για 15 εποχές με batch size 32 και τα αποτελέσματα της εκπαίδευσης φαίνονται παρακάτω:

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	Ακρίβεια
Adam	categorical crossentropy	61%
RMSProp	categorical crossentropy	56%

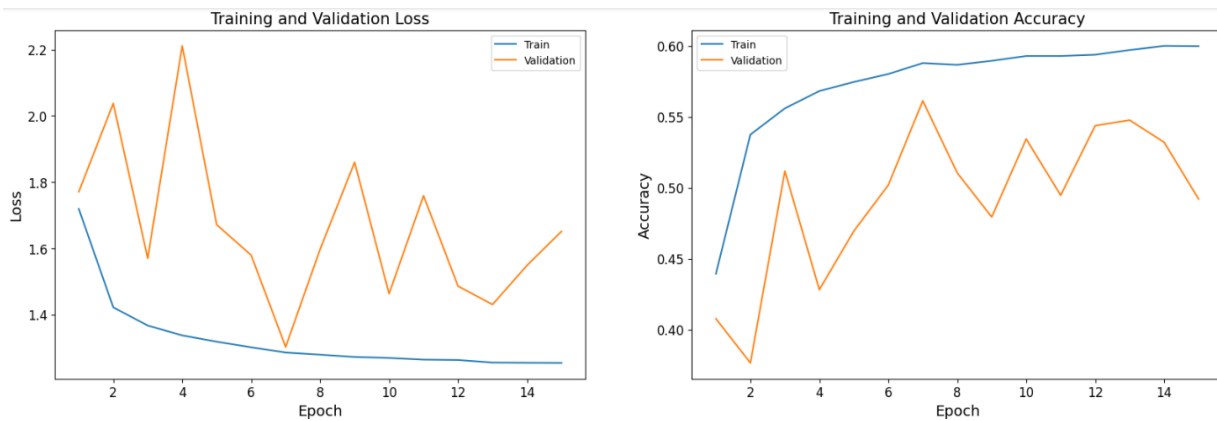
Πίνακας 4.6: Αποτελέσματα Μοντέλου 3

Στα παρακάτω διαγράμματα μελετάται η διακύμανση των μετρικών απόδοσης του μοντέλου.



Σχήμα 4.1: Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 3 με Αλγόριθμο Βελτιστοποίησης Adam





Σχήμα 4.2: Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 3 με Αλγόριθμο Βελτιστοποίησης RMSProp

Χωρίς το RNN επίπεδο παρατηρείται μια ανωμαλία στην εκπαίδευση του μοντέλου, η οποία είναι μεγαλύτερη όταν χρησιμοποιείται RMSProp ως αλγόριθμος βελτιστοποίησης.

Ένα ενδιαφέρον μοντέλο προκύπτει αν προστεθεί ένα μονοδιάστατο Convolutional επίπεδο. Η αρχιτεκτονική του μοντέλου φαίνεται παρακάτω:

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Conv1D	(None, None, 128)	Relu	192128
Global Max Pooling 1D	(None, 128)	-	0
Dense	(None, 64)	Relu	8256
Dense	(None, 26)	Softmax	1755

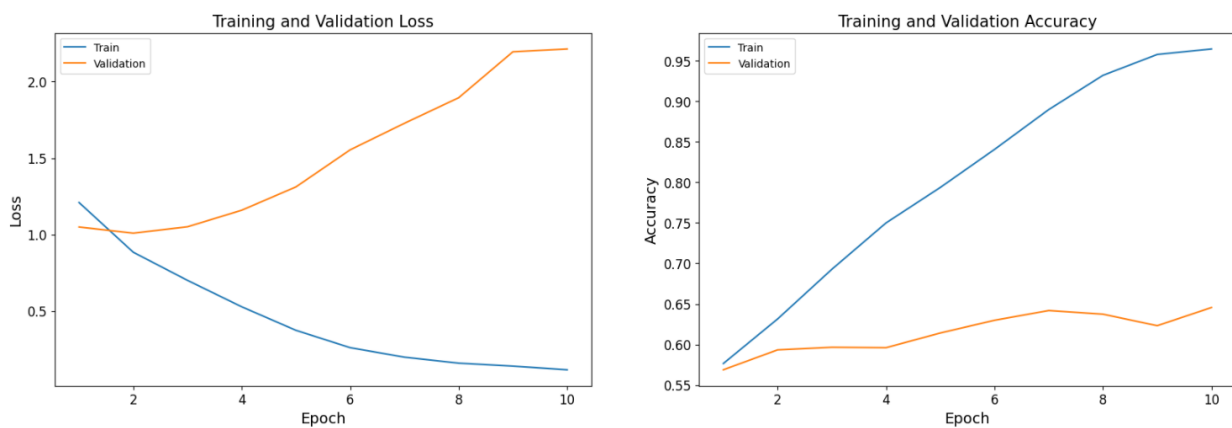
Πίνακας 4.7: Classification Μοντέλο 4

Και στην περίπτωση αυτή οι παράμετροι του Embedding επιπέδου θα μείνουν σταθερές κατά τη διάρκεια της εκπαίδευσης. Το μοντέλο εκπαιδεύεται για 10 εποχές με batch size 32 και τα αποτελέσματα της εκπαίδευσης καταγράφονται παρακάτω:

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	Ακρίβεια
Adam	categorical crossentropy	64%
RMSProp	categorical crossentropy	58%

Πίνακας 4.8: Αποτελέσματα Μοντέλου 4

Η διαδικασία εκπαίδευσης με Adam αλγόριθμο βελτιστοποίησης παρουσιάζεται παρακάτω:



Σχήμα 4.3: Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 4

Τα νούμερα είναι αρκετά αισιόδοξα και αυτό οφείλεται στο Embedding επίπεδο το οποίο διαθέτει μια μεγάλη βάση γνώσης, για αυτό φαίνεται ότι από την πρώτη κιόλας εποχή η ακρίβεια ξεκινάει από μια ικανοποιητική τιμή. Παρατηρείται μια ένδειξη overfitting για την οποία θα γίνουν προσπάθειες για να εξαλειφθεί στη συνέχεια.

Παρακάτω θα δοκιμαστεί ένα παρόμοιο μοντέλο, όμως αντί για max pooling θα υπάρχει ένα επίπεδο average pooling, και θα μελετηθούν οι διαφορές στα αποτελέσματα.

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Conv1D	(None, None, 512)	Relu	192128
Dropout (rate = 0.1)	(None, None, 512)	-	0
Global Average Pooling 1D	(None, 512)	-	0
Dense	(None, 256)	Relu	131328
Dense	(None, 128)	Relu	32896
Dense	(None, 64)	Relu	8256
Dense	(None, 26)	Softmax	1755

Πίνακας 4.9: Classification Μοντέλο 5

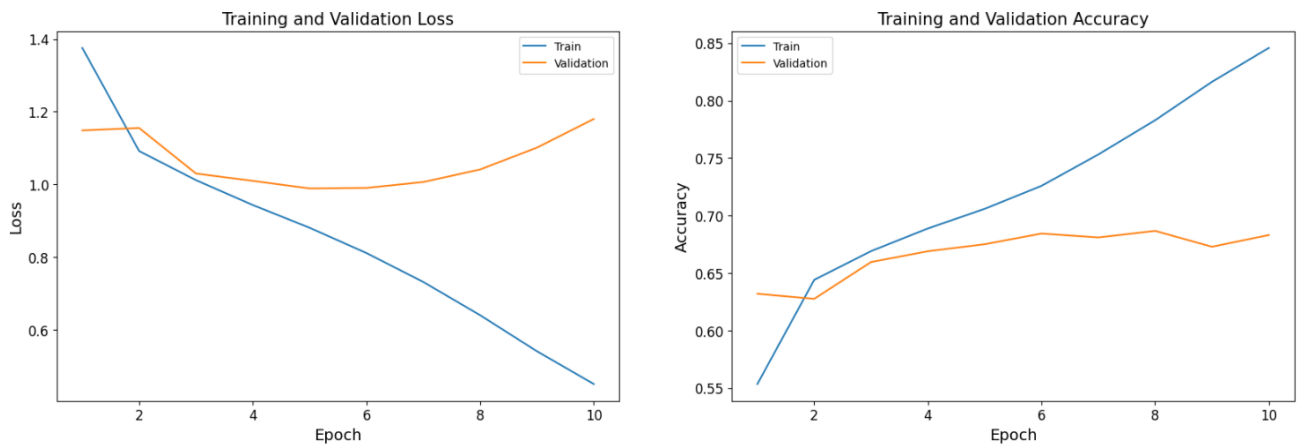
Ομοίως, σε αυτή την εκπαίδευση θα απενεργοποιηθεί η εκπαίδευση αυτών των παραμέτρων καθώς τα βάρη τους είναι ήδη μορφοποιημένα και επιπλέον τόσες πολλές μεταβλητές προς εκπαίδευση θα καθυστερήσουν αρκετά την όλη διαδικασία της μάθησης. Στο παραπάνω μοντέλο προστέθηκε και το επίπεδο Dropout για να αποφευχθεί το overfitting και να προστεθούν μερικά ακόμα Dense επίπεδα για να εξεταστεί αν θα βελτιωθεί η απόδοση.

Το μοντέλο εκπαιδεύεται για 10 εποχές με batch size 16 και τα αποτελέσματα της εκπαίδευσης φαίνονται παρακάτω:

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	Ακρίβεια
Adam	categorical crossentropy	66%
RMSProp	categorical crossentropy	60%

Πίνακας 4.10: Αποτελέσματα Μοντέλου 5

Όλη η διαδικασία εκπαίδευσης με Adam αλγόριθμο βελτιστοποίησης παρουσιάζεται παρακάτω:



Σχήμα 4.4: Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 5

Είναι εμφανές ότι το overfitting έχει ελαττωθεί και η απόδοση του μοντέλου έχει αυξηθεί.

Δεδομένου ότι η απόδοση έχει βελτιωθεί, απομένει να δημιουργηθεί ένα ακόμη πιο σύνθετο μοντέλο με περισσότερα επίπεδα. Κρατώντας τον αλγόριθμο Adam και αλλάζοντας το max pooling με ένα average pooling, θα μελετηθεί αν η απόδοσή του έχει βελτιωθεί ή όχι με αυτές τις αλλαγές.

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Conv1D	(None, None, 300)	Relu	540300
Conv1D	(None, None, 300)	Relu	270300
Dropout (rate = 0.1)	(None, None, 300)	-	0
Global Average Pooling 1D	(None, 300)	-	0
Dense	(None, 256)	Relu	77056
Dense	(None, 256)	Relu	65792
Dropout (rate = 0.1)	(None, None, 300)	-	0
Dense	(None, 128)	Relu	32896
Dense	(None, 128)	Relu	16512
Dropout (rate = 0.1)	(None, None, 300)	-	0
Dense	(None, 64)	Relu	8256
Dense	(None, 64)	Relu	4160
Dense	(None, 26)	Softmax	1755

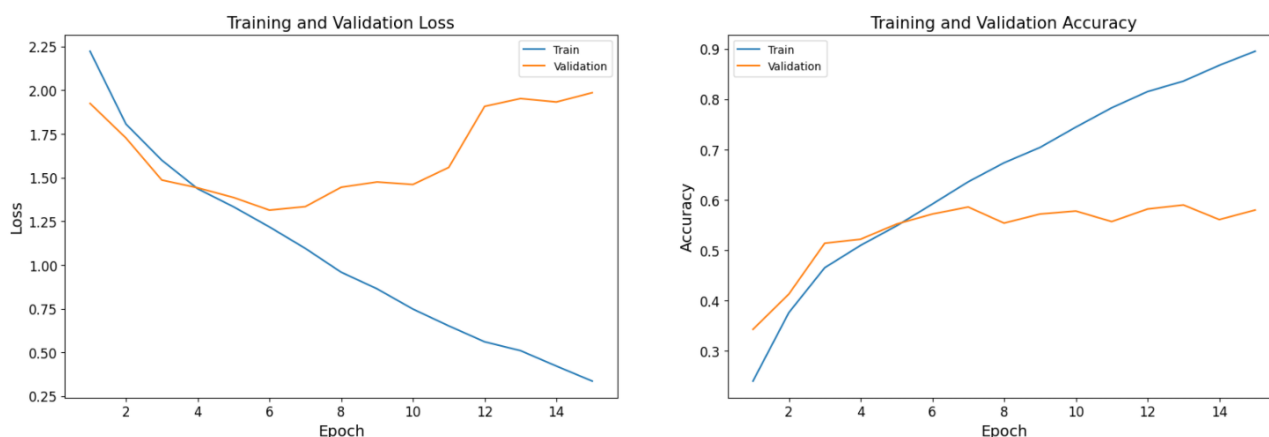
Πίνακας 4.11: Classification Μοντέλο 6

Το μοντέλο εκπαιδεύεται για 15 εποχές με batch size 16 και τα αποτελέσματα της εκπαίδευσης φαίνονται παρακάτω:

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	Ακρίβεια
Adam	categorical crossentropy	61%

Πίνακας 4.12: Αποτελέσματα Μοντέλου 6

Η διαδικασία εκπαίδευσης κατά τη διάρκεια των εποχών παρουσιάζεται παρακάτω:



Σχήμα 4.5: Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 6

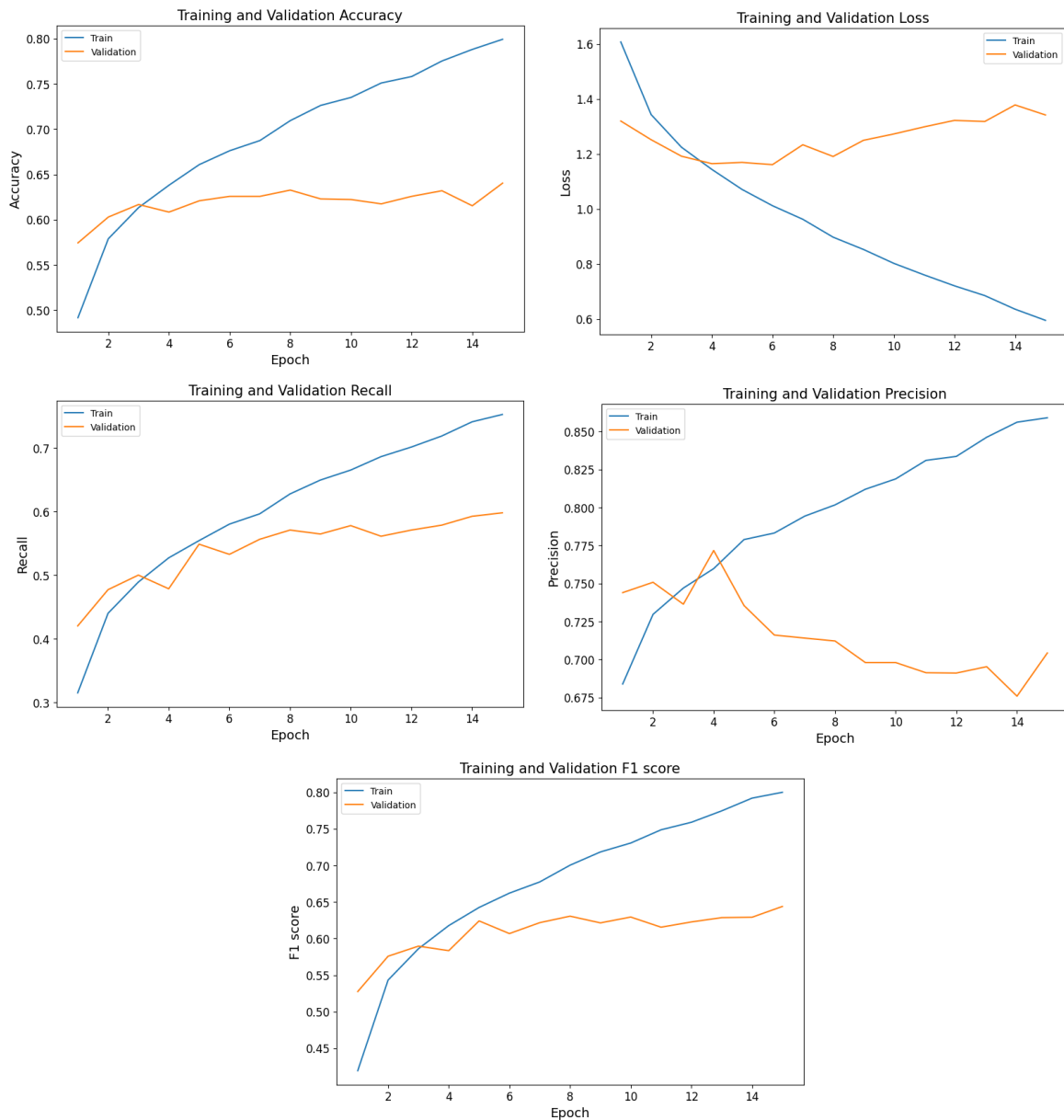
Από τα διαγράμματα προκύπτει πως η μεγαλύτερη πολυπλοκότητα της αρχιτεκτονικής του μοντέλου όχι μόνο δε βελτίωσε αλλά αντίθετα υποβάθμισε την απόδοσή του.

Στη συνέχεια, θα μελετηθεί ένα μοντέλο με ένα επίπεδο Bidirectional LSTM, δηλαδή ένα RNN που προσπελάζει τα δεδομένα κάθε βιβλίου με κατεύθυνση από την αρχή προς το τέλος και ένα παράλληλο το οποίο τα προσπελάζει από το τέλος προς την αρχή. Εδώ θα διαπιστωθεί αν υπάρχουν μοτίβα τα οποία, τα μέχρι τώρα σχεδιασμένα μοντέλα, δεν μπόρεσαν να εντοπίσουν.

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Bidirectional LSTM	(None, 128)	Relu	186880
Dropout (rate = 0.1)	(None, 128)	-	0
Batch Normalization	(None, 128)	-	512
Dense	(None, 256)	Relu	33024
Dense	(None, 128)	Relu	32896
Dense	(None, 64)	Relu	8256
Dense	(None, 26)	Softmax	1755

Πίνακας 4.13: Classification Μοντέλο 7

Στο παραπάνω μοντέλο έχει προστεθεί και το επίπεδο Batch Normalization το οποίο εκτελεί μια κανονικοποίηση σε κάθε δέσμη (batch) κατά τη διάρκεια της εκπαίδευσης. Το μοντέλο εκπαιδεύεται για 15 εποχές με batch size 12. Παρακάτω θα μελετηθούν περισσότερες μετρικές απόδοσης στα αποτελέσματα.



Σχήμα 4.6: Διαγράμματα μετρικών εκπαίδευσης Μοντέλου 7

Τα αντίστοιχα αποτελέσματα με τα train δεδομένα καταγράφονται παρακάτω:

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	F1 Score
Adam	categorical crossentropy	64%
Precision	Recall	Ακρίβεια
70%	59%	64%

Πίνακας 4.14: Αποτελέσματα Μοντέλου 7

### 4.3 Επιλογή Τελικού Μοντέλου

Τα μοντέλα με το RNN επίπεδο φαίνεται να λειτουργούν καλά και να αποδίδουν ικανοποιητικά αποτελέσματα. Ωστόσο, το Convolutional επίπεδο αποδεικνύεται αποτελεσματικότερο. Η υπερβολική πολυπλοκότητα δεν ωφέλησε, όπως διαπιστώθηκε στην προηγούμενη ενότητα, ενώ όσον αφορά τις συναρτήσεις ενεργοποίησης, η Relu έδωσε τα καλύτερα αποτελέσματα. Ο αλγόριθμος βελτιστοποίησης Adam φάνηκε πιο ικανός στο πρόβλημα αυτό καθώς επίσης και το average pooling επίπεδο αντί του max pooling. Η αρχιτεκτονική λοιπόν του τελικού μοντέλου παρουσιάζεται παρακάτω:

Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Conv1D	(None, None, 512)	Relu	768512
Dropout (rate = 0.1)	(None, None, 512)	-	0
Global Average Pooling 1D	(None, 512)	-	0
Dense	(None, 256)	Relu	131328
Dense	(None, 128)	Relu	32896
Dense	(None, 64)	Relu	8256
Dense	(None, 26)	Softmax	1690

Πίνακας 4.15: Τελικό Classification Μοντέλο 26 κλάσεων

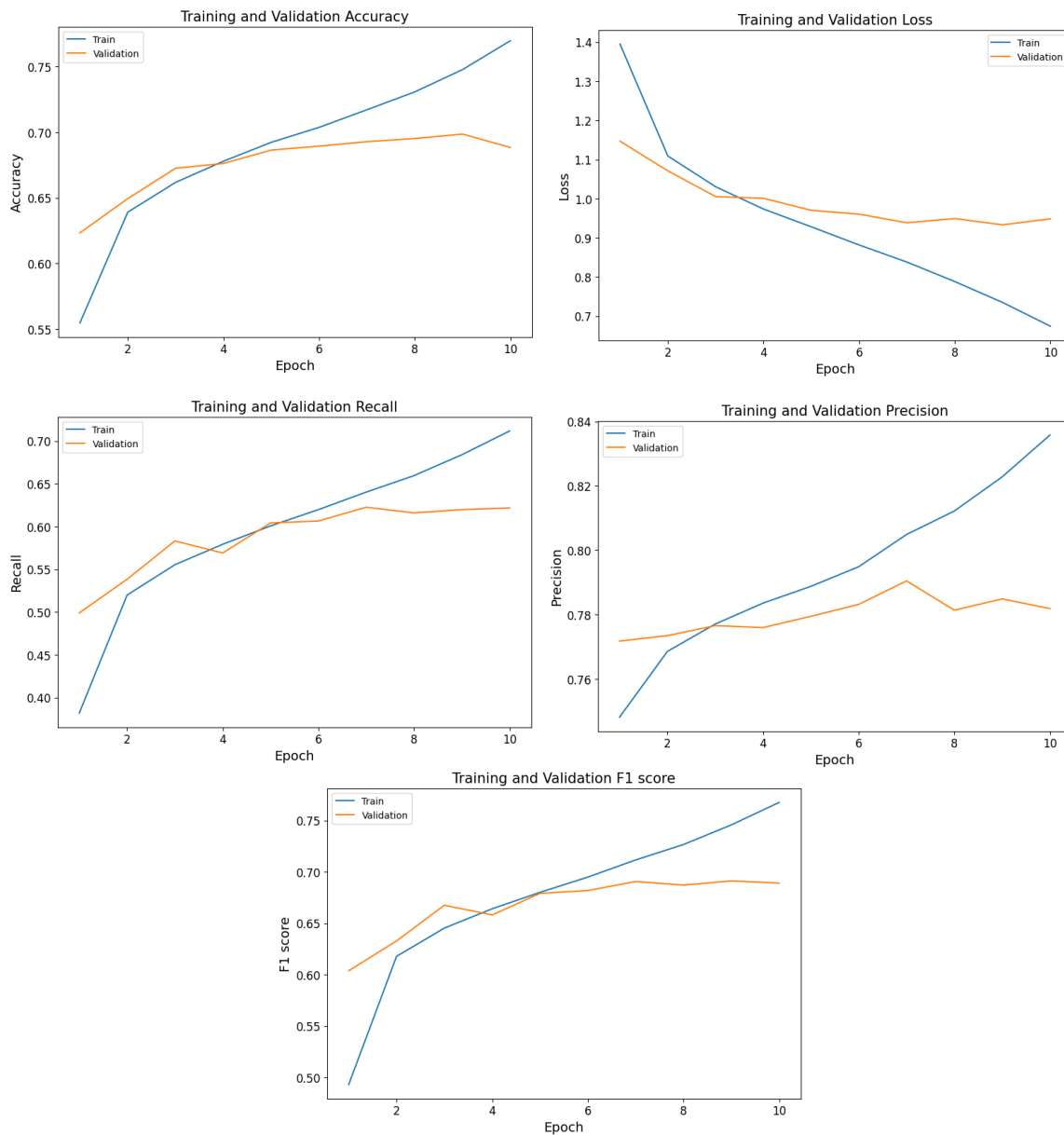
Δεδομένου ότι στις 15 εποχές δεν παρατηρείται τόσο μεγάλη βελτίωση, το τελικό



μοντέλο θα εκπαιδευτεί για 10 εποχές. Το μέγεθος δέσμης (batch size) ορίζεται 32 διότι με 16 δεν παρουσιάστηκαν βελτιώσεις και δε συντρέχει λόγος να σπαταληθεί υπολογιστική ισχύς για να το μειώσουμε.

<b>Αλγόριθμος Βελτιστοποίησης</b>	<b>Συνάρτηση Κόστους</b>	<b>F1 Score</b>
Adam	categorical crossentropy	68%
<b>Precision</b>	<b>Recall</b>	<b>Ακρίβεια</b>
74%	62%	70%

Πίνακας 4.16: Αποτελέσματα Τελικού Μοντέλου 26 κλάσεων



Σχήμα 4.7: Διαγράμματα μετρικών εκπαίδευσης Τελικού Μοντέλου 26 κλάσεων

Τα αποτελέσματα επαληθεύουν τις παραπάνω παρατηρήσεις και καταδεικνύουν ότι η συγκεκριμένη αρχιτεκτονική είναι καταλληλότερη για να χρησιμοποιηθεί και να ταξινομήσει βιβλία με βάση τον πίνακα περιεχομένων τους. Οι μετρικές απόδοσης κυμαίνονται γύρω στο 70%, ένα πολύ υψηλό ποσοστό συγκρινόμενο με τα αρχικά αποτελέσματα.

Λαμβάνοντας υπόψη ότι οι κλάσεις δεν είναι ισότιμες καθώς κάποιες περιέχουν πολύ περισσότερα δεδομένα σε σχέση με κάποιες άλλες, θα δοκιμαστεί το ίδιο μοντέλο, αλλά αυτή τη φορά θα εκπαιδευτεί μόνο με βιβλία που ανήκουν στις 7 πιο δημοφιλείς κατηγορίες. Αυτές, όπως έχει ήδη επισημανθεί στο προηγούμενο κεφάλαιο είναι οι εξής:

1. Επιστήμη Υπολογιστών (11219 βιβλία)
2. Φυσικές Επιστήμες (8788 βιβλία)
3. Ιατρική (8471 βιβλία)
4. Μηχανική (8161 βιβλία)
5. Μαθηματικά (4852 βιβλία)
6. Οικονομικά (3762 βιβλία)
7. Βιοεπιστήμες (3124 βιβλία)

Το σύνολο των βιβλίων αυτών είναι 48385 και ακολούθως τα δεδομένα χωρίζονται τυχαία ως εξής:

1. **train:** 38708 βιβλία
2. **validation:** 4838 βιβλία
3. **test:** 4839 βιβλία

Το μοντέλο αυτό διαφοροποιείται μόνο ως προς το τελευταίο Dense επίπεδο το οποίο πλέον έχει μήκος 7 (όσο και οι κλάσεις).

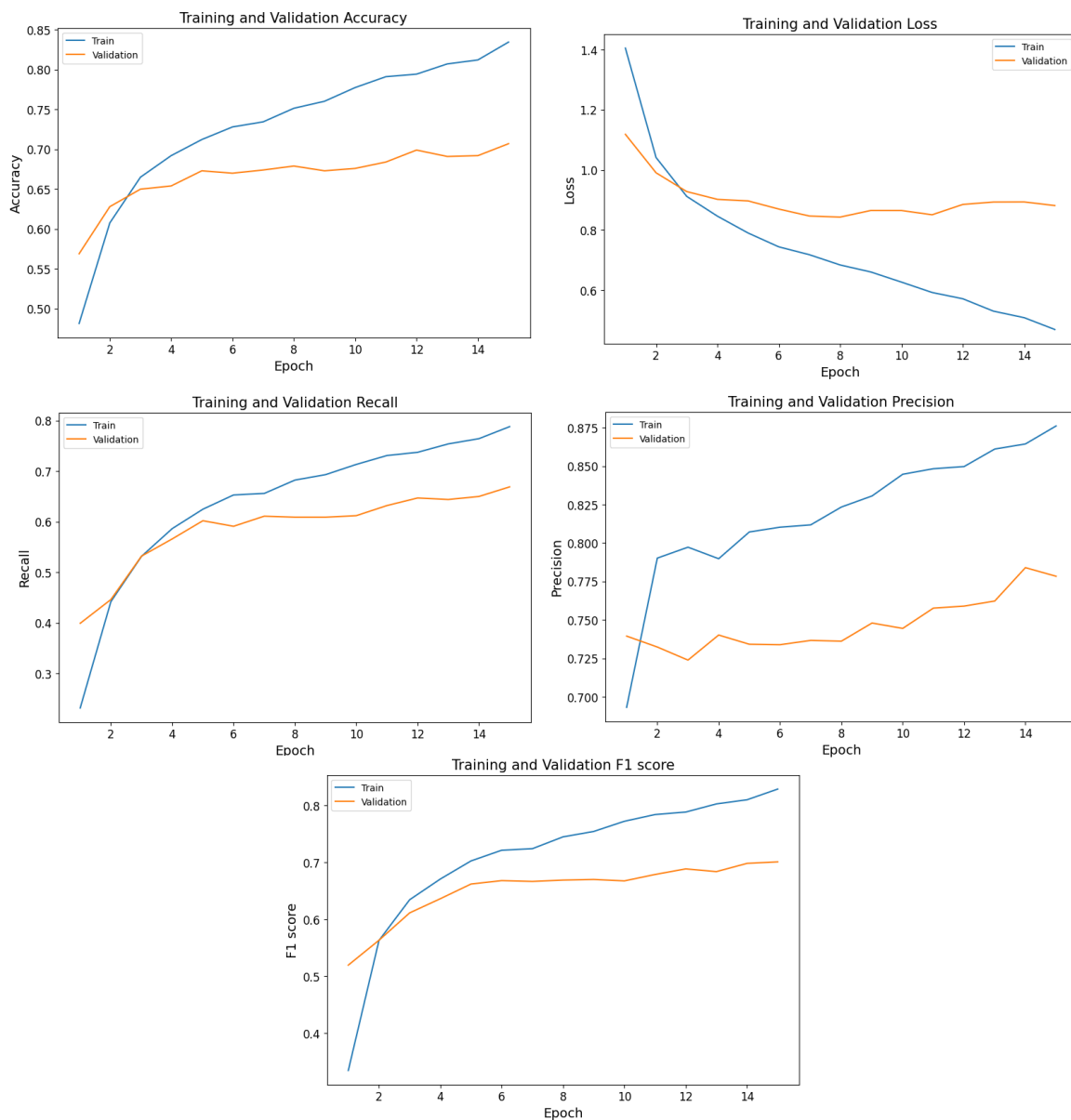
Επίπεδο	Έξοδος	Συνάρτηση Ενεργοποίησης	Παράμετροι
Embedding	(None, None, 300)	-	900000000
Conv1D	(None, None, 512)	Relu	768512
Dropout (rate = 0.1)	(None, None, 512)	-	0
Global Average Pooling 1D	(None, 512)	-	0
Dense	(None, 256)	Relu	131328
Dense	(None, 128)	Relu	32896
Dense	(None, 64)	Relu	8256
Dense	(None, 7)	Softmax	455

Πίνακας 4.17: Τελικό Classification Μοντέλο 7 κλάσεων

Το μοντέλο εκπαιδεύεται για 15 εποχές και το μέγεθος δέσμης (batch size) ορίζεται κι εδώ 32. Τα τελικά αποτελέσματα παρουσιάζονται παρακάτω:

Αλγόριθμος Βελτιστοποίησης	Συνάρτηση Κόστους	F1 Score
Adam	categorical_crossentropy	71%
<b>Precision</b>	<b>Recall</b>	<b>Ακρίβεια</b>
77%	67%	72%

Πίνακας 4.18: Αποτελέσματα Τελικού Μοντέλου 6 κλάσεων



Σχήμα 4.8: Διαγράμματα μετρικών εκπαίδευσης Τελικού Μοντέλου 6 κλάσεων

Από τα διαγράμματα συνάγεται ότι υπήρξε σημαντική βελτίωση στις μετρικές απόδοσης. Αυτό δικαιολογείται από το γεγονός ότι ο μέσος όρος δειγμάτων ανά κατηγορία είναι μεγαλύτερος και έτσι το μοντέλο κατόρθωσε να διακρίνει αποτελεσματικότερα τα χαρακτηριστικά της κάθε κατηγορίας. Το εκπαιδευμένο αυτό μοντέλο θα

χρησιμοποιηθεί και στην εφαρμογή του Μέρους II της διπλωματικής εργασίας.

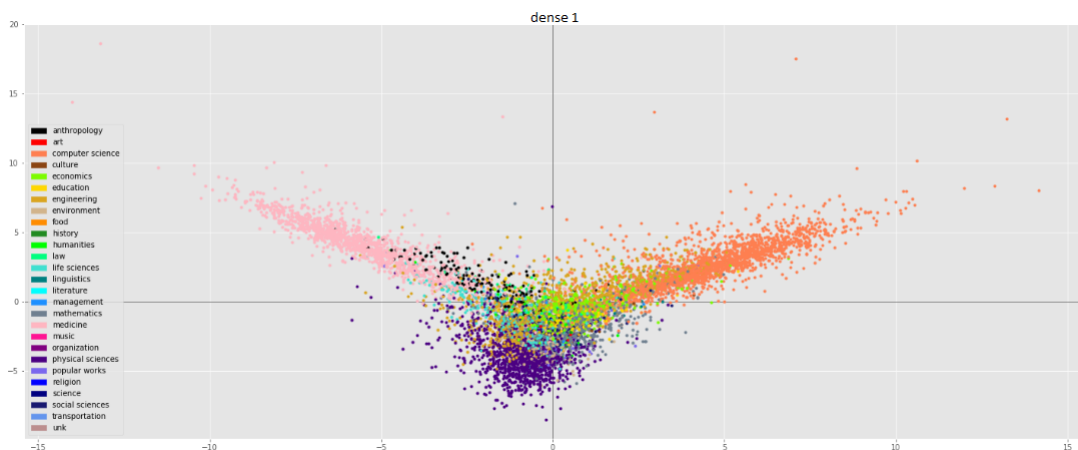
## 4.4 Οπτικοποιήσεις Δεδομένων

Στο προηγούμενο κεφάλαιο επιχειρήθηκε μια οπτικοποίηση των δεδομένων στον διδιάστατο χώρο. Σε αυτή την ενότητα θα προστεθούν νέα χαρακτηριστικά σε κάθε βιβλίο, τα οποία θα χρησιμοποιηθούν και στις εφαρμογές της Μη Επιβλεπόμενης Μάθησης. Τα χαρακτηριστικά αυτά θα ανακτηθούν από τα ενδιάμεσα επίπεδα ενεργοποίησης του τελικού μοντέλου επιβλεπόμενης μάθησης.

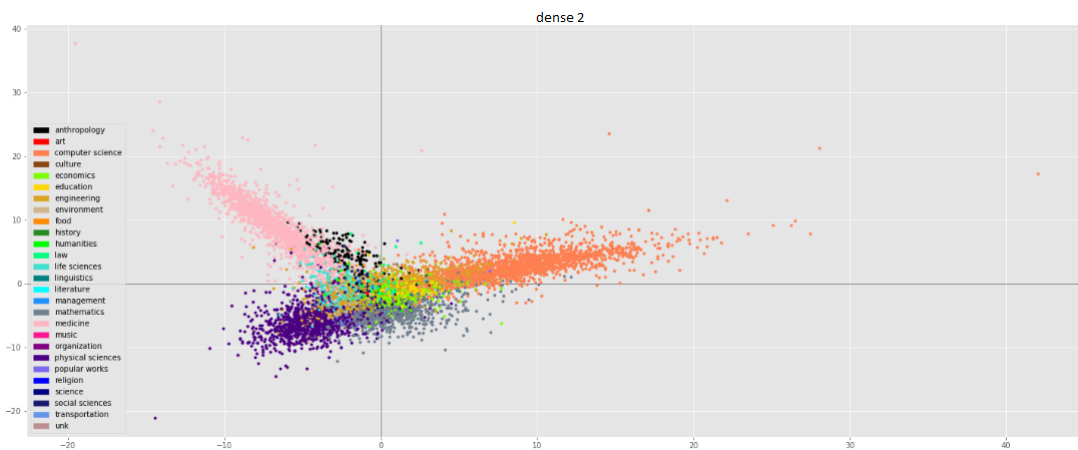
Για κάθε βιβλίο που εισάγεται στο Νευρωνικό Δίκτυο, θα μελετηθούν τα 3 Dense επίπεδά του και οι ενεργοποιήσεις τους.

1. **dense 1:** Dense επίπεδο μήκους 256
2. **dense 2:** Dense επίπεδο μήκους 128
3. **dense 3:** Dense επίπεδο μήκους 64

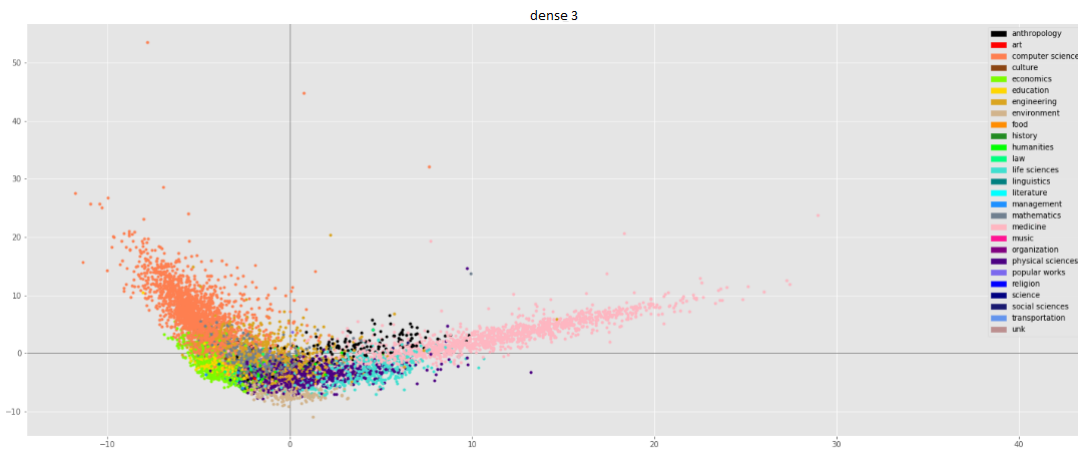
Συνεπώς, τα δεδομένα θα έχουν 256, 128 και 64 χαρακτηριστικά αντίστοιχα, βγαίνοντας από κάθε επίπεδο. Για να πραγματοποιηθεί η διαδικασία της οπτικοποίησης στον διδιάστατο χώρο θα χρησιμοποιηθεί η συνάρτηση PCA. Τα αποτελέσματα της φαίνονται παρακάτω:



Σχήμα 4.9: Οπτικοποίηση δεδομένων μετά την έξοδο τους από το επίπεδο dense 1.



Σχήμα 4.10: Οπτικοποίηση δεδομένων μετά την έξοδο τους από το επίπεδο dense 2.



Σχήμα 4.11: Οπτικοποίηση δεδομένων μετά την έξοδο τους από το επίπεδο dense 3.

Τα παραπάνω διαγράμματα αξιολογούνται πολύ θετικά καθώς αποτυπώνεται ευδι-  
άκριτα η διατήρηση της πληροφορίας στο εσωτερικό του δικτύου και η ομαδοποίηση  
των κλάσεων ανάλογα τα χαρακτηριστικά τους. Για τις επόμενες ενότητες της μη επι-  
βλεπόμενης μάθησης θα χρησιμοποιηθεί το επίπεδο dense 3 και πλέον το κάθε βιβλίο  
θα έχει 64 νέα χαρακτηριστικά.

## 4.5 Δημιουργία Clusters

Έχοντας πλέον διαθέσιμα τα νέα χαρακτηριστικά για κάθε βιβλίο, θα εφαρμοστούν τεχνικές μη επιβλεπόμενης μάθησης με σκοπό τη δημιουργία ομάδων (clusters). Τα βιβλία που ανήκουν στο ίδιο cluster θα έχουν κάποια κοινά χαρακτηριστικά.

### 4.5.1 Επίπεδα και Αλγόριθμοι Clustering

Παρακάτω θα σχεδιαστούν και θα μελετηθούν διάφορες αρχιτεκτονικές μοντέλων. Τα εσωτερικά επίπεδα των μοντέλων μπορούν να διακριθούν σε τρεις κατηγορίες, ανάλογα με την εργασία που εκτελούν:

#### 1. Επίπεδο επιλογής χαρακτηριστικών

Σε αυτό το επίπεδο, χρησιμοποιώντας τη συνάρτηση **variance threshold**, αφαιρούνται από τα δεδομένα εισόδου όλα εκείνα τα χαρακτηριστικά τα οποία έχουν μηδενικές ή μικρές διακυμάνσεις στις διάφορες τιμές τους στα αντίστοιχα δείγματα. Όταν η διακύμανση είναι μικρότερη από ένα άνω όριο (threshold), τότε το αντίστοιχο χαρακτηριστικό αφαιρείται από τα διανύσματα χαρακτηριστικών των δειγμάτων.

#### 2. Επίπεδο προεπεξεργασίας δεδομένων

Σε αυτό το επίπεδο, όπως είναι γνωστό και από την επιβλεπόμενη μάθηση, γίνεται προεπεξεργασία των δεδομένων εισόδου. Εδώ θα χρησιμοποιηθεί η συνάρτηση **MinMax**. Η συνάρτηση αυτή κανονικοποιεί τα δεδομένα εισόδου και μετασχηματίζει όλες τις τιμές των χαρακτηριστικών σε μια κλίμακα από το 0 έως το 1. Σε αυτό το επίπεδο θα χρησιμοποιηθεί και η συνάρτηση **PCA**, η οποία έχει αναλυθεί και χρησιμοποιηθεί ήδη στο κεφάλαιο της προεπεξεργασίας και ανάλυσης δεδομένων. Η συνάρτηση αυτή μειώνει τα χαρακτηριστικά των δεδομένων με χρήση γραμμικής άλγεβρας. Ο αριθμός των νέων χαρακτηριστικών δίνεται ως είσοδος (number of components).

#### 3. Επίπεδο ομαδοποίησης (clustering)

Πρόκειται για το επίπεδο στο οποίο πραγματοποιείται η διαδικασία της ομαδοποίησης δεδομένων με τη δημιουργία clusters. Για να επιτευχθεί αυτό, χρησιμο-

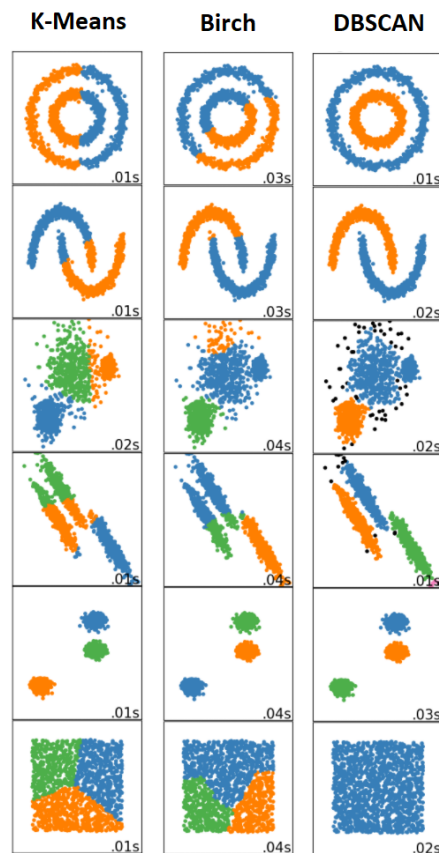


ποιούνται αλγόριθμοι clustering. Οι αλγόριθμοι που θα εισαχθούν στα παρακάτω μοντέλα είναι οι εξής:

- **K-Means** Ο αλγόριθμος αυτός ομαδοποιεί τα δεδομένα με κριτήριο την ελαχιστοποίηση μιας μεταβλητής, γνωστής ως *inertia*. Αυτή η μεταβλητή είναι το άθροισμα των αποστάσεων κάθε σημείου από το κοντινότερο κέντρο από τα  $C$  κέντρα.

$$inertia = \sum_{i=1}^n \min_{m \in C} (\|x_i - m_i\|^2) \quad (4.1)$$

- **Birch** Ο αλγόριθμος Birch δημιουργεί ένα δέντρο για τα δεδομένα εισόδου, το CFT (Clustering Feature Tree). Τα δεδομένα εμφωλεύονται μέσα σε κόμβους, οι οποίοι είναι παιδιά κόμβων και περιέχουν υποψήφιες ομάδες clusters.
- **DBSCAN** Ο αλγόριθμος αυτός έχει τη δυνατότητα αφενός να ομαδοποιεί τα δεδομένα και αφετέρου να βρίσκει τα σημεία τα οποία αποκλίνουν από το σύνολο (outliers), τα σημεία δηλαδή που είναι κάπως απομακρυσμένα, και δεν ταιριάζουν σε κάποιο cluster. Τον αλγόριθμο αυτόν θα τον χρησιμοποιήσουμε για αυτήν του τη δυνατότητα.



Σχήμα 4.12: Ομαδοποίηση δεδομένων με τους 3 clustering αλγορίθμους και οπτικοποίηση των clusters [10]

Τα μοντέλα που θα σχεδιαστούν πρέπει να αξιολογηθούν ως προς την απόδοσή τους. Παρακάτω θα οριστεί η μετρική απόδοσης Silhouette:

$$silhouette = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (4.2)$$

όπου  $a(i)$ : η μέση απόδοση ενός σημείου  $i$  και όλων των υπόλοιπων σημείων που ανήκουν στο ίδιο cluster

$b(i)$ : η μικρότερη μέση απόσταση του σημείου  $i$  από όλα τα σημεία τα οποία ανήκουν στα υπόλοιπα clusters εκτός του cluster του  $i$ . Το γειτονικό cluster επιλέγεται για τον υπολογισμό του  $b(i)$  και είναι αυτό με τη μικρότερη τιμή.

Η μετρική silhouette παίρνει τιμές στο διάστημα  $[-1, 1]$ .

### 4.5.2 Δοκιμές Μοντέλων και Αποτελέσματα

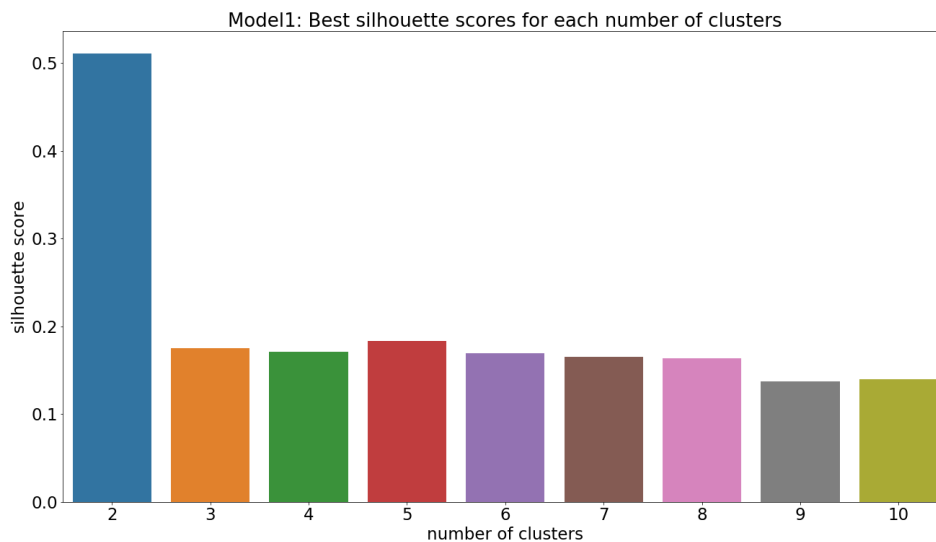
Κατά τις πρώτες δοκιμές που έγιναν στα δεδομένα, καταγράφηκε αρκετά χαμηλή τιμή στη μετρική silhouette. Με μια διερεύνηση στα δεδομένα, διαπιστώνεται ότι κάποιοι πίνακες περιεχομένων περιέχουν μόνο λίγες λέξεις οι οποίες δεν είναι επαρκείς για να ταξινομηθεί το βιβλίο σε κάποιο cluster. Έτσι, φιλτράροντας τα βιβλία και αφαιρώντας εκείνα που στον πίνακα περιεχομένων διαθέτουν πλήθος λέξεων μικρότερο από 10, το σύνολο των βιβλίων που θα ομαδοποιηθούν είναι πλέον 50904.

Αρχικά, θα μελετηθεί το πιο απλό μοντέλο που αποτελείται μόνο από τον αλγόριθμο kmeans.

CLUSTERER
<i>kmeans</i> (number of clusters)

Πίνακας 4.19: Clustering Μοντέλο 1

Παρακάτω απεικονίζεται το διάγραμμα με τα αποτελέσματα για κάθε διαφορετικό αριθμό clusters:



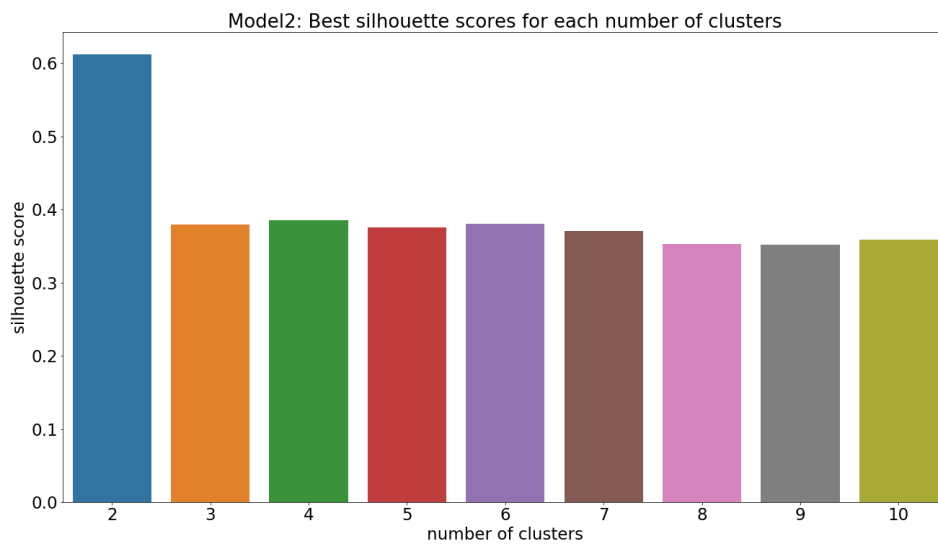
Σχήμα 4.13: Αποτελέσματα της μετρικής silhouette για κάθε αριθμό clusters

Στο δεύτερο μοντέλο θα προστεθεί και το επίπεδο προεπεξεργασίας.

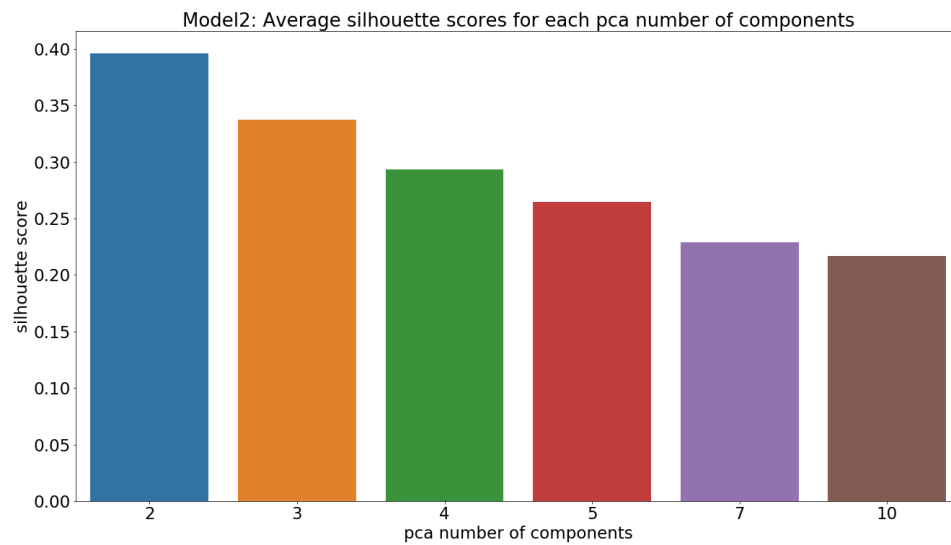
<b>PREPROCESSOR</b>
<i>PCA</i> (number of components)
<b>CLUSTERER</b>
<i>kmeans</i> (number of clusters)

Πίνακας 4.20: Clustering Μοντέλο 2

Τα αποτελέσματα απεικονίζονται παρακάτω:



Σχήμα 4.14: Αποτελέσματα της μετρικής silhouette για κάθε αριθμό clusters

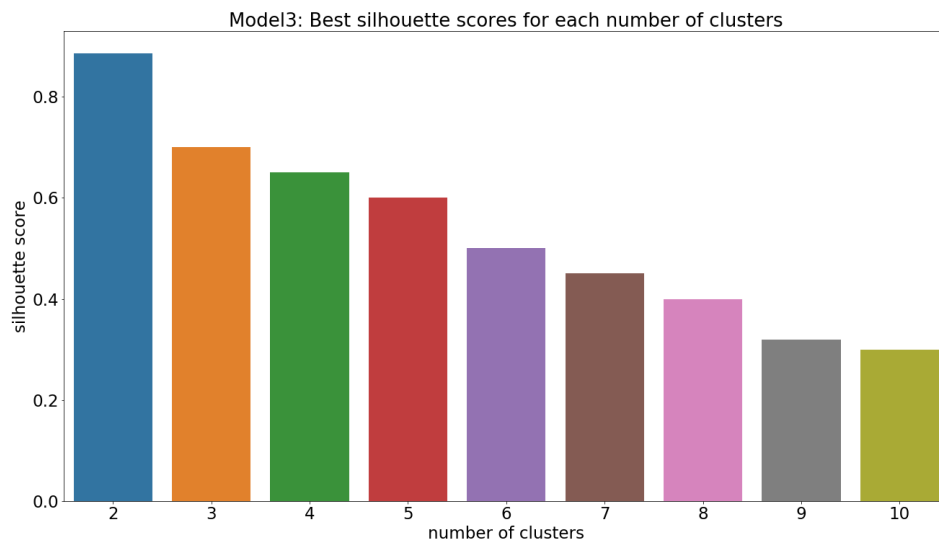


Σχήμα 4.15: Μέση απόδοση silhouette για κάθε αριθμό pca components

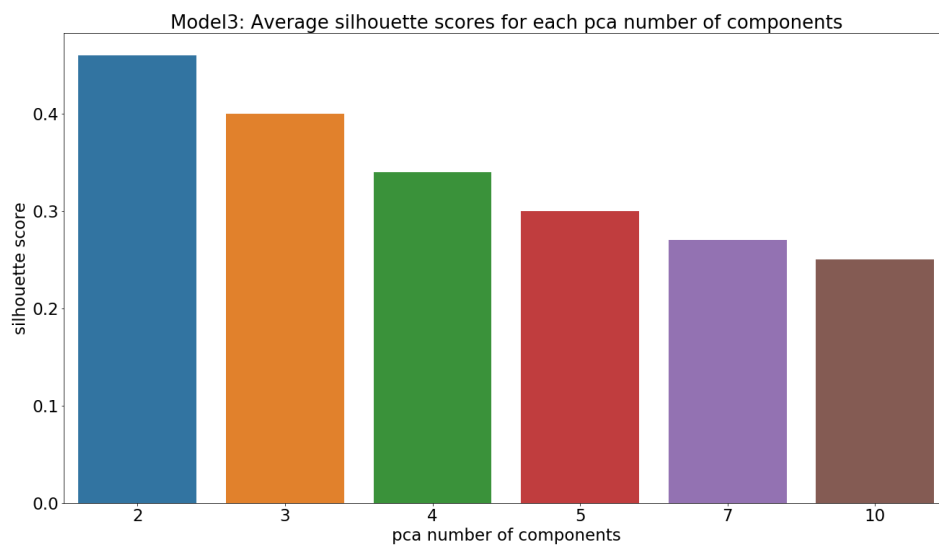
Στη συνέχεια, θα προστεθεί και το επίπεδο επιλογής χαρακτηριστικών με τη συνάρτηση variance threshold.

<b>FEATURE SELECTOR</b>
<i>variance threshold(threshold)</i>
<b>PREPROCESSOR</b>
<i>PCA(number of components)</i>
<b>CLUSTERER</b>
<i>kmeans(number of clusters)</i>

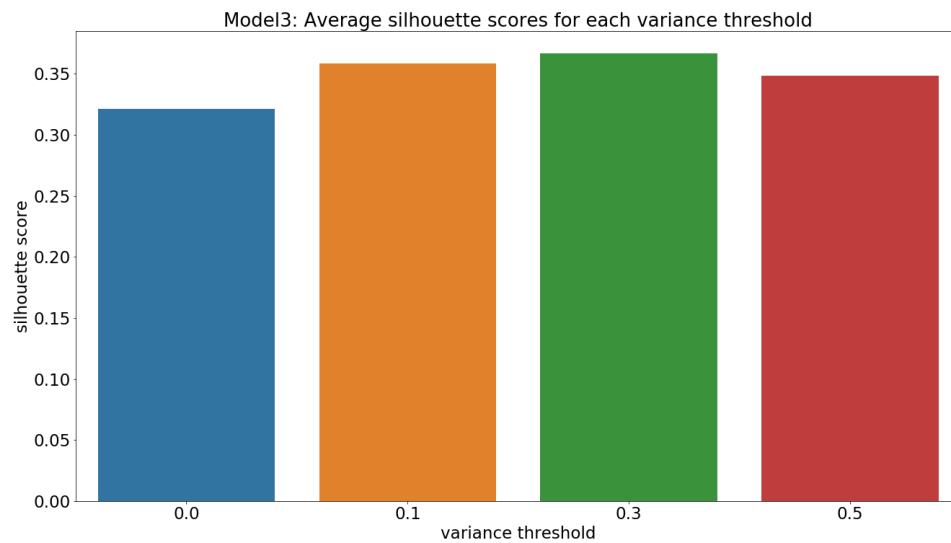
Πίνακας 4.21: Clustering Μοντέλο 3



Σχήμα 4.16: Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5]



Σχήμα 4.17: Μέση απόδοση silhouette για κάθε αριθμό pca components

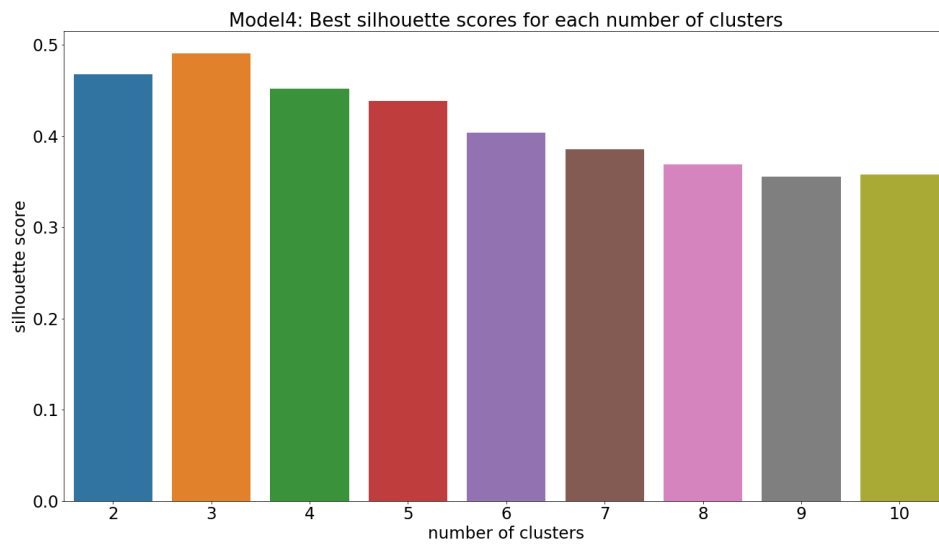


Σχήμα 4.18: Μέση απόδοση silhouette για κάθε αριθμό threshold

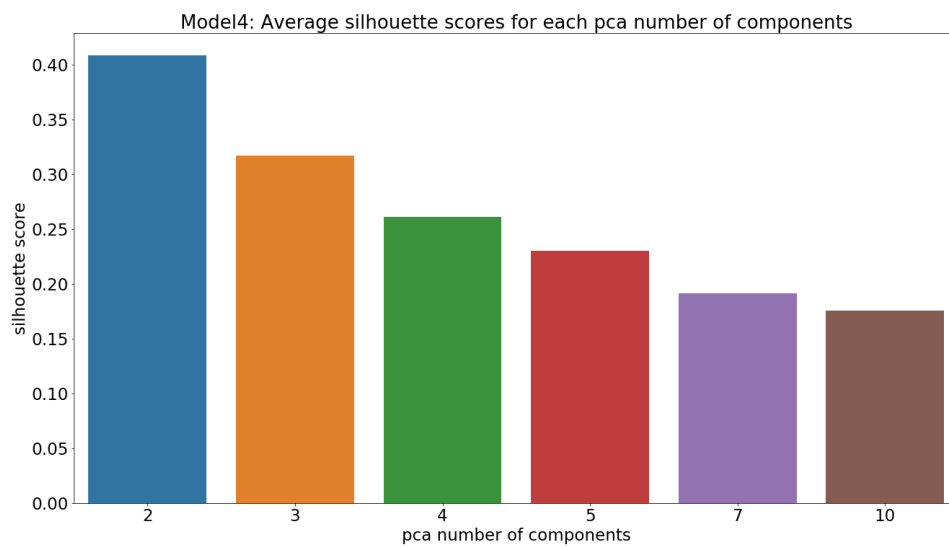
Σε εξέλιξη του προηγούμενου μοντέλου θα προστεθεί σε αυτό και η συνάρτηση MinMax στο επίπεδο της προεπεξεργασίας.

<b>FEATURE SELECTOR</b>
<i>variance threshold(threshold)</i>
<b>PREPROCESSOR</b>
<i>MinMax Scaler()</i>
<i>PCA(number of components)</i>
<b>CLUSTERER</b>
<i>kmeans(number of clusters)</i>

Πίνακας 4.22: Clustering Μοντέλο 4

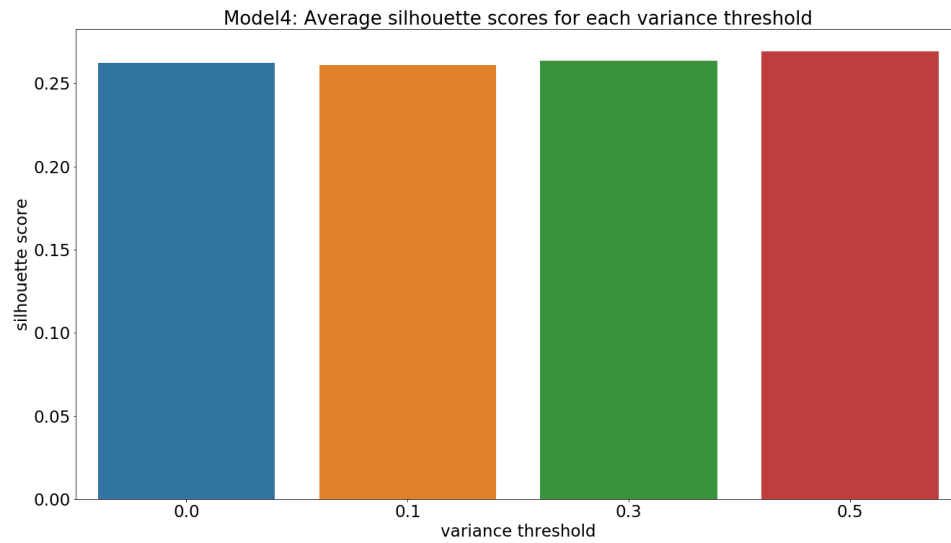


Σχήμα 4.19: Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5]



Σχήμα 4.20: Μέση απόδοση silhouette για κάθε αριθμό pca components



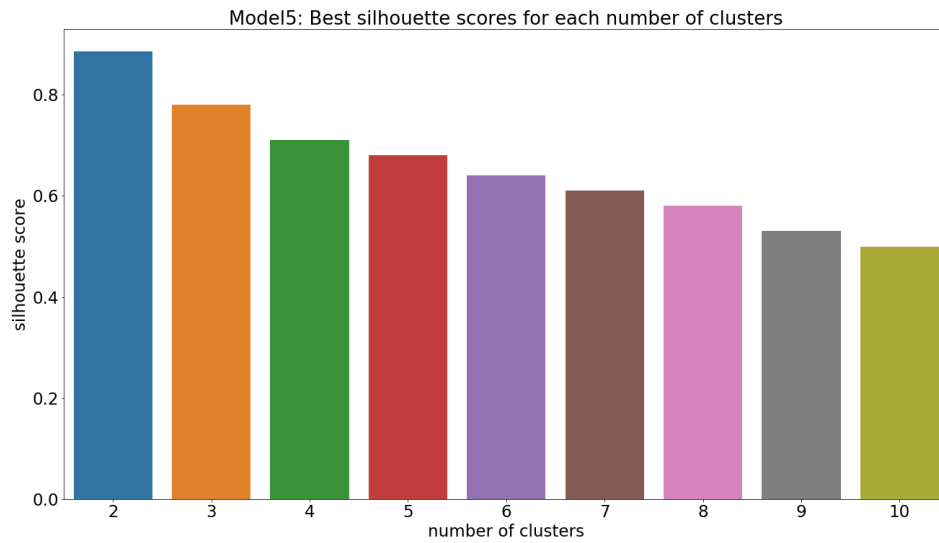


Σχήμα 4.21: Μέση απόδοση silhouette για κάθε αριθμό threshold

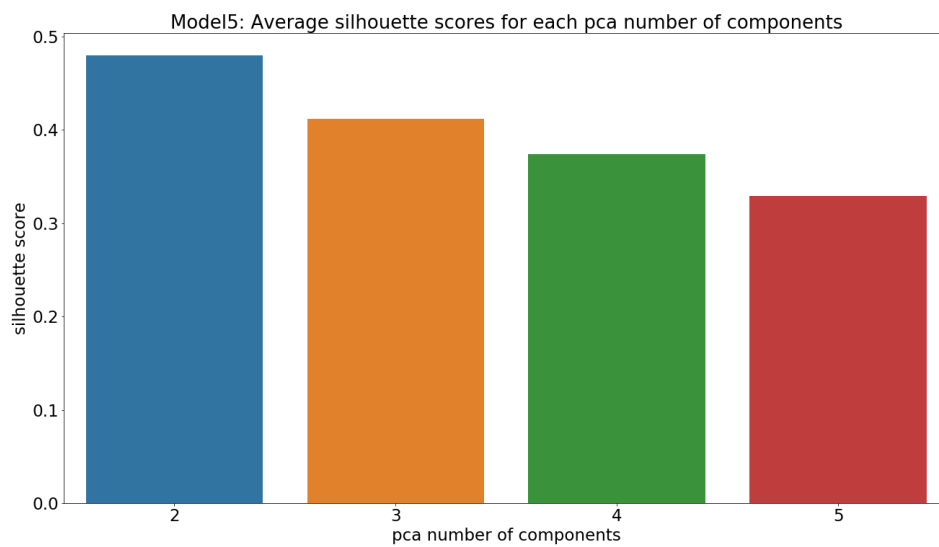
Τέλος, θα διερευνηθεί και ο clustering αλγόριθμος Birch στο τελευταίο μοντέλο που θα αναλυθεί.

<b>FEATURE SELECTOR</b>
<i>variance threshold(threshold)</i>
<b>PREPROCESSOR</b>
<i>PCA(number of components)</i>
<b>CLUSTERER</b>
<i>Birch(number of clusters)</i>

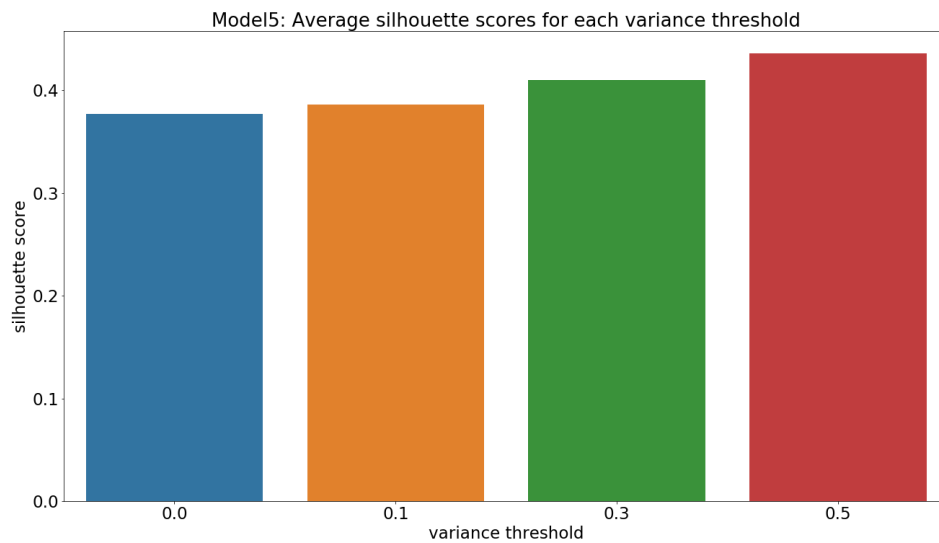
Πίνακας 4.23: Clustering Μοντέλο 5



Σχήμα 4.22: Απεικόνιση επιπέδων ενός συνελικτικού δικτύου [5]



Σχήμα 4.23: Μέση απόδοση silhouette για κάθε αριθμό pca components



Σχήμα 4.24: Μέση απόδοση silhouette για κάθε αριθμό threshold

### 4.5.3 Σχεδιασμός Τελικού Μοντέλου και Παρουσίαση των Clusters

Προτού τα δεδομένα εισαχθούν στο τελικό μοντέλο, θα εφαρμοστεί ο αλγόριθμος DBSCAN για να εντοπίσει τα απομακρυσμένα σημεία προκειμένου να απαλειφθούν τα αντίστοιχα βιβλία από το σύνολο. Εκτελώντας τον αλγόριθμο, βρίσκουμε 401 βιβλία τα οποία αποκλίνουν από το σύνολο. Επομένως, στο Μοντέλο 5 θα εισαχθούν 50503 βιβλία για ομαδοποίηση.

Ο αλγόριθμος Birch παράγει σαφώς καλύτερα αποτελέσματα και, επομένως, είναι ο καταλληλότερος για να χρησιμοποιηθεί, ενώ το Μοντέλο 5 είναι το τελικό μοντέλο clustering. Πραγματοποιήθηκαν δοκιμές για αριθμό clusters ίσο με 4, 5 και 6. Τα βέλτιστα αποτελέσματα όσον αφορά το περιεχόμενο τους τα πήραμε για τα 4 clusters.

Έτσι, το τελικό μοντέλο θα αποτελείται από τις ακόλουθες παραμέτρους:

- **number of clusters** = 5
- **pca components** = 2

- **threshold** = 0.5

Η απόδοση silhouette για το παραπάνω μοντέλο σύμφωνα με τα τελικά δεδομένα είναι **0.71**. Για κάθε cluster θα διερευνηθούν το ποσοστό των συνολικών βιβλίων που περιέχουν, η πλειοψηφία των κατηγοριών στις οποίες εντάσσονται τα δεδομένα τους και οι πιο συχνές λέξεις που εμφανίζονται στους πίνακες περιεχομένων τους.

<b>Cluster 1</b>		
<b>Ποσοστό βιβλίων</b>	<b>Συχνότερες κατηγορίες</b>	<b>Κοινές λέξεις</b>
3031 βιβλία (6%)	Ιατρική (91%) Βιοεπιστήμες (5%) Φυσικές Επιστήμες (2.5%) Νομικά (0.5%)	surgery disease treatment management

Πίνακας 4.24: Περιεχόμενα του Cluster 1

<b>Cluster 2</b>		
<b>Ποσοστό βιβλίων</b>	<b>Συχνότερες κατηγορίες</b>	<b>Κοινές λέξεις</b>
19191 βιβλία (38%)	Επιστήμη Υπολογιστών (48%) Μηχανική (25%) Μαθηματικά (14%) Οικονομικά (6%)	model data software systems management applications

Πίνακας 4.25: Περιεχόμενα του Cluster 2

<b>Cluster 3</b>		
<b>Ποσοστό βιβλίων</b>	<b>Συχνότερες κατηγορίες</b>	<b>Κοινές λέξεις</b>
4041 βιβλία (8%)	Ιατρική (80%) Βιοεπιστήμες (8%) Φυσικές Επιστήμες (7%) Μηχανική (1%)	cell treatment cancer clinical disease therapy

Πίνακας 4.26: Περιεχόμενα του Cluster 3

<b>Cluster 4</b>		
<b>Ποσοστό βιβλίων</b>	<b>Συχνότερες κατηγορίες</b>	<b>Κοινές λέξεις</b>
21717 βιβλία (43%)	Φυσικές Επιστήμες (35%) Μηχανική (17%) Μαθηματικά (14%) Οικονομικά (10%)	analysis system research development

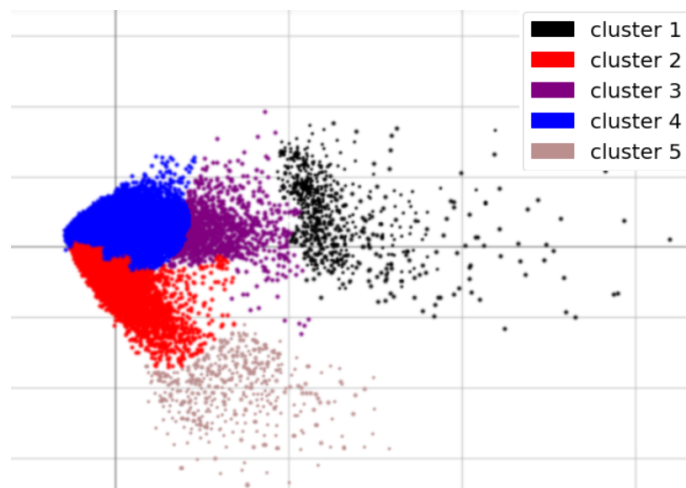
Πίνακας 4.27: Περιεχόμενα του Cluster 4

<b>Cluster 5</b>		
<b>Ποσοστό βιβλίων</b>	<b>Συχνότερες κατηγορίες</b>	<b>Κοινές λέξεις</b>
2523 βιβλία (5%)	Επιστήμη Υπολογιστών (55%) Μηχανική (44%)	algorithm neural network semantic query web

Πίνακας 4.28: Περιεχόμενα του Cluster 5

Μετά το στάδιο της προεπεξεργασίας, και συγκεκριμένα μετά την εκτέλεση της συνάρτησης PCA, τα δεδομένα διαθέτουν μόνο δύο χαρακτηριστικά - διαστάσεις. Έτσι,

εάν τοποθετήσουμε τα βιβλία στον δισδιάστατο χώρο και τα χρωματίσουμε με βάση το cluster στο οποίο ανήκουν, διαμορφώνεται το παρακάτω διάγραμμα:



Σχήμα 4.25: Οπτικοποίηση των 5 clusters

## 4.6 Δημιουργία Συστήματος Εύρεσης Συναφών Συγγραμμάτων

Σε αυτή την ενότητα θα περιγραφεί η διαδικασία υλοποίησης του συστήματος εύρεσης συναφών συγγραμμάτων για ένα σύγγραμμα της βάσης δεδομένων. Το σύστημα αυτό θα εισαχθεί και θα χρησιμοποιηθεί στην εφαρμογή, η οποία θα παρουσιαστεί στο Μέρος II της διπλωματικής εργασίας.

Όπως και στις προηγούμενες ενότητες, έτσι κι εδώ θα χρησιμοποιηθεί ο πίνακας των 64 χαρακτηριστικών για κάθε βιβλίο. Αυτά τα χαρακτηριστικά, τα οποία προέκυψαν από τα ενδιάμεσα επίπεδα των Νευρωνικών Δικτύων ταξινόμησης, μελετήθηκαν, αναλύθηκαν και συνέβαλαν στη δημιουργία των clusters. Με αυτά τα χαρακτηριστικά λοιπόν θα δουλέψει το σύστημα αυτό.

Έστω τα βιβλία  $B_1, B_2$  με διανύσματα χαρακτηριστικών:

$$B_1 = (b_{1,1}, b_{1,2}, \dots, b_{1,64})$$

$$B_2 = (b_{2,1}, b_{2,2}, \dots, b_{2,64})$$

Η ευκλείδεια απόσταση των δύο βιβλίων στον 64-διάστατο χώρο είναι:

$$d_{12} = \sqrt{\sum_{i=1}^{64} (b_{1,i} - b_{2,i})^2} \quad (4.3)$$

Για ένα επιλεγμένο βιβλίο, τα 6 εγγύτερά του είναι αυτά με την μικρότερη ευκλείδεια απόσταση από αυτό στο σύνολο των βιβλίων. Παρακάτω, παρουσιάζονται δύο παραδείγματα συναφών βιβλίων σύμφωνα με την τεχνική που αναλύθηκε.

Επιλεγμένο βιβλίο	
Τίτλος	Κατηγορία
Wireless Networking and Mobile Data Management	Επιστήμη Υπολογιστών

Συναφή βιβλία		
Τίτλος	Κατηγορία	Ευκλείδεια Απόσταση
The Safety of Systems	Επιστήμη Υπολογιστών	1.31
Modeling and Analysis of Voice and Data in Cognitive Radio Networks	Επιστήμη Υπολογιστών	1.37
Practical .NET 2.0 Networking Projects	Επιστήμη Υπολογιστών	1.48
Network Coding at Different Layers in Wireless Networks	Μηχανική	1.52
Scheduling and Congestion Control for Wireless Internet	Επιστήμη Υπολογιστών	1.55
Guide to Voice and Video over IP	Επιστήμη Υπολογιστών	1.60

Επιλεγμένο βιβλίο	
Τίτλος	Κατηγορία
Real Analysis	Μαθηματικά

Συναφή βιβλία		
Τίτλος	Κατηγορία	Ευκλείδεια Απόσταση
Invariant Random Fields on Spaces with a Group Action	Μαθηματικά	1.87
Real Analysis via Sequences and Series	Μαθηματικά	1.88
Modern Real Analysis	Μαθηματικά	2.60
Translational Recurrences	Μαθηματικά	2.78
Singular Stochastic Differential Equations	Μαθηματικά	2.80
An Introduction to Markov Processes	Μαθηματικά	2.86

Τα συναφή βιβλία στις δύο αυτές περιπτώσεις που παρουσιάστηκαν, είναι αρκετά στοχευμένα και συγκλίνουν ως προς το περιεχόμενό τους. Στο πρώτο παράδειγμα τα συναφή βιβλία έχουν ως κοινό θέμα τα δίκτυα ενώ στο δεύτερο παράδειγμα κοινό τους θέμα είναι μαθηματική ανάλυση. Επιπρόσθετα, όλα υπάγονται στην ίδια ή σε περιμετρική κατηγορία.

Επειδή δεν είναι εφικτό να εξεταστούν για κάθε βιβλίο χωριστά τα 6 συναφή του, θα μελετηθεί συλλογικά η απόδοση του συστήματος. Έτσι, για κάθε βιβλίο που εντάσσεται στις 7 πιο δημοφιλείς κατηγορίες, αφού πρώτα εντοπιστούν τα 6 συναφή με αυτό, στη συνέχεια θα προσδιοριστεί το ποσοστό με το οποίο συμμετέχει το καθένα σε κάθε κατηγορία. Με αυτόν τον τρόπο θα προκύψει μια γενική εικόνα για την αποτελεσματικότητα του συστήματος εύρεσης συναφών βιβλίων. Τα αποτελέσματα αποτυπώνονται παρακάτω:



Κατηγορία	Κατηγορίες συναφών βιβλίων	Ποσοστό
Επιστήμη Υπολογιστών (11219 βιβλία)	Επιστήμη Υπολογιστών	72%
	Μηχανική	16%
	Μαθηματικά	6%
	Οικονομικά	3%
Φυσικές Επιστήμες (8788 βιβλία)	Φυσικές Επιστήμες	59%
	Μηχανική	15%
	Μαθηματικά	14%
	Βιοεπιστήμες	5%
Ιατρική (8471 βιβλία)	Ιατρική	76%
	Βιοεπιστήμες	6%
	Φυσικές Επιστήμες	5%
	Ανθρωπολογία	3%
Μηχανική (8161 βιβλία)	Μηχανική	52%
	Επιστήμη Υπολογιστών	22%
	Φυσικές Επιστήμες	15%
	Μαθηματικά	8%
Μαθηματικά (4852 βιβλία)	Μαθηματικά	56%
	Μηχανική	11%
	Μαθηματικά	9%
	Επιστήμη Υπολογιστών	6%
Οικονομικά (3762 βιβλία)	Οικονομικά	57%
	Μηχανική	14%
	Μαθηματικά	11%
	Επιστήμη Υπολογιστών	6%
Βιοεπιστήμες (3124 βιβλία)	Βιοεπιστήμες	51%
	Φυσικές Επιστήμες	15%
	Ιατρική	15%
	Περιβάλλον	6%

Πίνακας 4.31: Κατηγορίες συναφών βιβλίων για τα βιβλία που ανήκουν στις 7 πιο δημοφιλείς κλάσεις.

Από τον παραπάνω πίνακα φαίνεται η ευστοχία του συστήματος να εντοπίζει βιβλία τα οποία ανήκουν στην ίδια ή σε αντίστοιχη κατηγορία. Οι δοκιμές πραγματοποιήθηκαν μόνο για τις 7 κατηγορίες, διότι για τις υπόλοιπες δεν υπάρχουν διαθέσιμα επαρκή δεδομένα.

## Μέρος II

# Υλοποίηση Web Application

## Κεφάλαιο 5

### Θεωρητικό Υπόβαθρο

Η έρευνα και οι επιστημονικές μελέτες που πραγματοποιούνται στον κλάδο της πληροφορικής, γίνονται με κάποιο σκοπό. Μεγάλο ποσοστό αυτών επιδιώκεται να βρει πρακτική εφαρμογή στην καθημερινότητα των ανθρώπων και να χρησιμοποιηθεί από πολλούς χρήστες, επιλύοντας κάποιο πρόβλημα ή ψυχαγωγώντας τους. Επομένως, η σχεδίαση και υλοποίηση συστημάτων με σκοπό την εφαρμογή και διάθεσή τους στο ευρύ κοινό είναι πολύ σημαντική για τη σωστή χρήση και αξιοποίηση των μοντέλων Τεχνητής Νοημοσύνης που έχουν σχεδιαστεί μετά από πολλή έρευνα, μελέτη και πειραματικές δοκιμές.

Για το λόγο αυτόν, θα κατασκευαστεί μια ολοκληρωμένη web εφαρμογή η οποία θα μπορεί να υποστηρίξει πλήρως όλα τα μοντέλα και τις αναλύσεις που έγιναν στο πρώτο μέρος της διπλωματικής και παράλληλα θα αποτελέσει ένα περιβάλλον στο οποίο ο χρήστης θα μπορεί να πλοηγηθεί άνετα, να εξερευνήσει και να αναζητήσει βιβλία, να ενημερωθεί για αυτά και να βρει αυτό που ψάχνει χωρίς κόπο και πολύωρη προσπάθεια.

## 5.1 Αλληλεπίδραση Ανθρώπου-Υπολογιστή

Ο κλάδος της επιστήμης που μελετά τον τρόπο, τις μεθόδους και τις τεχνικές με τις οποίες οι άνθρωποι αλληλεπιδρούν με τους υπολογιστές ονομάζεται Αλληλεπίδραση Ανθρώπου-Υπολογιστή (Human-Computer Interaction ή για συντομία HCI).

*“Η Αλληλεπίδραση Ανθρώπου-Υπολογιστή, γνωστή και ως επικοινωνία ανθρώπου-μηχανής, είναι το επιστημονικό πεδίο της πληροφορικής που μελετά την αλληλεπίδραση μεταξύ ανθρώπων (χρηστών) και υπολογιστών.”*  
[28]

Αυτή η επιστήμη θεωρείται ως το σημείο τομής μεταξύ πληροφορικής, επιστημών ψυχολογίας, όπως η γνωστική και κοινωνική ψυχολογία, γλωσσολογίας και βιομηχανικού σχεδιασμού. Η αλληλεπίδραση αυτή, λοιπόν, μεταξύ ανθρώπου και υπολογιστή πραγματοποιείται στο επίπεδο της διεπαφής χρήστη (user interface), μέσω κατάλληλου λογισμικού και υλικού και εν προκειμένω μέσω της web εφαρμογής.

## 5.2 React Framework

Το 2011 η ομάδα προγραμματιστών του Facebook υλοποίησαν μια νέα τεχνολογία για την ανάπτυξη web και native εφαρμογών, την React (ReactJS ή React.js). Τα βασικότερα χαρακτηριστικά του framework αυτού είναι:

- η διαίρεση του κώδικα σε μικρότερα στοιχεία, γνωστά ως components
- η μονοσήμαντη ροή δεδομένων μέσω των ιδιοτήτων (props) των στοιχείων
- η δυνατότητα αποθήκευσης της κατάστασης ενός στοιχείου (stateful component) και η μετάδοσής της στα στοιχεία-παιδιά (child-component) από τα στοιχεία-γονείς (parent-component) μέσω των ιδιοτήτων (props)
- η JSX σύνταξη η οποία είναι παρόμοια με την HTML
- η δυναμική ανανέωση μόνο του περιεχομένου της σελίδας το οποίο μεταβάλλεται (virtual DOM)
- η δυνατότητα εκτέλεσης τμημάτων κώδικα σε δεδομένα χρονικά σημεία ζωής του αντίστοιχου στοιχείου που αφορούν

Στη web εφαρμογή που θα σχεδιαστεί, ο frontend server θα αναπτυχθεί με το react framework.

### 5.3 Βάση Δεδομένων MongoDB

Η MongoDB είναι μια μη-σχεσιακή (non-relational) βάση δεδομένων (database) με υποστήριξη σε πολλές πλατφόρμες, που χρησιμοποιεί έγγραφα για την καταγραφή των δεδομένων (document-oriented). Η αποθήκευση δεδομένων γίνεται σε συλλογές (collections), όπου κάθε ένα έχει την μορφή JSON κειμένου.

Τα κυριότερα χαρακτηριστικά της MongoDB είναι τα εξής:

- **Queries**

Η MongoDB υποστηρίζει διάφορες μορφές αναζήτησης δεδομένων εντός της βάσης με διάφορα κριτήρια και επιλογές. Το αποτέλεσμα μπορεί να είναι ένα συγκεκριμένο πεδίο εντός εγγράφου, μια ομάδα πεδίων ή ένα σύνολο εγγράφων δεδομένου μεγέθους.

- **Replication**

Η MongoDB παρέχει υψηλή διαθεσιμότητα σε replica sets, κάθε ένα από τα οποία περιέχει αντίγραφα των δεδομένων. Όλες οι ενέργειες εγγραφής και ανάγνωσης γίνονται στο κύριο αντίγραφο από προεπιλογή. Όταν ένα κύριο αντίγραφο αποτύχει, τα δευτερεύοντα παρέχουν τα δεδομένα που χρειάζονται.

- **Indexing**

Το κάθε πεδίο στη βάση έχει έναν μοναδικό δείκτη (index).

- **File Storage**

Η MongoDB μπορεί να χρησιμοποιηθεί ως ένα σύστημα αρχείων που ονομάζεται GridFS με δυνατότητες εξισορρόπησης φορτίου και αναπαραγωγής δεδομένων σε πολλές μηχανές ταυτόχρονα, με αποδοτικές μεθόδους. Τα αρχεία διαιρούνται σε τμήματα και κάθε ένα από αυτά αποθηκεύονται ως ξεχωριστά έγγραφα.

Αυτές είναι κάποιες από τις δυνατότητες και τα χαρακτηριστικά της βάσης δεδομένων MongoDB, η οποία θα αποτελέσει τη βάση δεδομένων της εφαρμογής που θα χρησιμοποιήσουμε.

## Κεφάλαιο 6

# Σχεδιασμός και Υλοποίηση Συστήματος

Σε αυτό το τελευταίο κεφάλαιο θα προσδιοριστούν και θα αναλυθούν οι απαιτήσεις της web εφαρμογής που θα υλοποιηθεί. Αυτές χωρίζονται σε λειτουργικές και μη λειτουργικές και οι διαφορές τους θα αναδειχθούν στην ανάλυση. Στην ανάλυση που ακολουθεί, θα παρουσιαστούν οι ανάγκες της εφαρμογής, η αρχιτεκτονική και ο σχεδιασμός της. Στο τελευταίο μέρος του κεφαλαίου θα παρουσιαστεί η εφαρμογή στην τελική της μορφή, οι δυνατότητες και οι διάφορες οθόνες που αντιστοιχούν σε κάθε σενάριο χρήσης είτε χρησιμοποιείται υπολογιστής (desktop) είτε κινητό (smartphone).

## 6.1 Απαιτήσεις Συστήματος

Σε αυτή την ενότητα θα αναφερθούν οι απαιτήσεις του συστήματος, οι οποίες εξασφαλίζουν την ολοκληρωμένη λειτουργία της εφαρμογής και την προσφορά της επιθυμητής εμπειρίας στον χρήστη. Περαιτέρω, θα αναλυθούν οι λειτουργικές απαιτήσεις οι οποίες είναι βασικές και απαραίτητες για την εφαρμογή και οι μη λειτουργικές απαιτήσεις οι οποίες δεν έχουν σχέση με την υλοποίηση της εφαρμογής, αλλά γενικότερα με τη λειτουργικότητά της.

### 6.1.1 Λειτουργικές Απαιτήσεις Συστήματος

Η διεπαφή του χρήστη της εφαρμογής θα πρέπει να προσφέρει τις παρακάτω λειτουργίες:

#### 1. Προβολή όλων των ηλεκτρονικών βιβλίων

Ο χρήστης θα έχει τη δυνατότητα να δει μια λίστα με όλα τα βιβλία που έχουν καταχωρηθεί στο σύστημα συνοπτικά. Η λίστα των βιβλίων περιέχει μερικές βασικές πληροφορίες που βοηθούν τον χρήστη να κατανοήσει τα βασικά τους χαρακτηριστικά, να διαπιστώσει αν ενδιαφέρεται για κάποιο και να ανακτήσει περισσότερες πληροφορίες για αυτό.

#### 2. Προβολή ενός ηλεκτρονικού βιβλίου

Ο χρήστης θα έχει τη δυνατότητα να ανακτήσει περισσότερες πληροφορίες για κάποιο σύγγραμμα του ενδιαφέροντός του. Θα μπορεί δηλαδή να δει όλες τις επιπλέον λεπτομέρειες και τα χαρακτηριστικά του βιβλίου αυτού, όπως τα αναλυτικά περιεχόμενα, την περίληψη και τις λέξεις-κλειδιά.

#### 3. Προβολή παρόμοιων συγγραμμάτων για κάθε βιβλίο

Ο χρήστης, για κάθε βιβλίο που ανοίγει, προκειμένου να δει όλες τις πληροφορίες που περιέχει το σύστημα για αυτό, θα μπορεί να δει και παρόμοια συγγράμματα με αυτό, τα οποία έχει εντοπίσει το σύστημα.

#### 4. Προβολή των clusters και των βιβλίων που ανήκουν σε αυτά

Ο χρήστης θα έχει πρόσβαση στα περιεχόμενα του κάθε cluster, αποτέλεσμα εφαρμογής Μη Επιβλεπόμενης Μάθησης, όπως αναπτύχθηκε στο πρώτο μέρος της διπλωματικής εργασίας.



### 5. Μπάρα αναζήτησης

Ο χρήστης θα μπορεί να πληκτρολογήσει κάποια λέξη-κλειδί και να πραγματοποιήσει αναζήτηση σχετικά με αυτή.

### 6. Δυνατότητα εισαγωγής πίνακα περιεχομένων ενός νέου βιβλίου

Ο χρήστης θα μπορεί να εισαγάγει τον πίνακα περιεχομένων ενός νέου βιβλίου και το σύστημα να το κατηγοριοποιήσει σε μια από τις διαθέσιμες κατηγορίες στις οποίες έχει εκπαιδευτεί το αντίστοιχο Νευρωνικό Δίκτυο. Επιπλέον, θα μπορεί να δει τα παρόμοια με αυτό βιβλία.

## 6.1.2 Μη Λειτουργικές Απαιτήσεις Συστήματος

Για την ολοκληρωμένη εμπειρία του χρήστη, η εφαρμογή θα πληροί και κάποιες μη λειτουργικές απαιτήσεις, οι οποίες απαριθμούνται παρακάτω:

### 1. Απόδοση

Ο χρόνος απόκρισης του συστήματος σε αιτήματα που δέχεται από τους χρήστες θα πρέπει να είναι σύντομος και η επεξεργασία τους να γίνεται με αποδοτικό τρόπο.

### 2. Εύχρηστο Γραφικό Περιβάλλον Χρήστη

Το γραφικό περιβάλλον κατά τη διάρκεια της πλοήγησης του χρήστη, θα πρέπει να είναι σχεδιασμένο έτσι ώστε να πληροί τους δείκτες ευχρηστίας. Ο χρήστης πρέπει να μπορεί εύκολα και απλά να βρει αυτό που θέλει στην εφαρμογή και να λάβει όλη την απαραίτητη πληροφορία που προβλέπεται σε κάθε διαφορετική οθόνη.

### 3. Επεκτασιμότητα

Η υλοποίηση της εφαρμογής θα πρέπει να στηρίζεται σε τεχνολογίες που επιτρέπουν τη μελλοντική επέκτασή τους.

## 6.2 Αρχιτεκτονική Εφαρμογής

Η web εφαρμογή αποτελείται από τρεις βασικούς συντελεστές: τον frontend server, ο οποίος είναι η άμεση διεπαφή του χρήστη (client) με την εφαρμογή, τον backend

server, ο οποίος εξυπηρετεί τα αιτήματα του frontend server και η βάση δεδομένων, που αποτελεί τον αποθηκευτικό χώρο της εφαρμογής.

### 6.2.1 Υποδομή Client (Frontend)

Η εφαρμογή δομείται σε τρία κύρια μέρη: την αρχική σελίδα, τη σελίδα εξερεύνησης και αναζήτησης συγγραμμάτων και τη σελίδα εισαγωγής και κατηγοριοποίησης ενός νέου βιβλίου. Οι αντίστοιχες οθόνες για αυτές τις λειτουργίες αναλύονται παρακάτω:

#### 1. Αρχική σελίδα

Είναι η σελίδα στην οποία μεταφέρεται ο χρήστης μόλις μπει στην εφαρμογή.

#### 2. Οθόνη εμφάνισης πολλών βιβλίων

Εδώ ο χρήστης μπορεί να δει όλα τα βιβλία που περιέχει η βάση δεδομένων και να επιλέξει για ποιο θέλει να ανακτήσει περισσότερες πληροφορίες.

#### 3. Οθόνη αναλυτικής εμφάνισης ενός βιβλίου

Σε αυτή την οθόνη ο χρήστης μπορεί να δει αναλυτικά όλες τις πληροφορίες που περιέχει η βάση για ένα βιβλίο καθώς και insights τα οποία έχουν προκύψει από την ανάλυση των βιβλίων με τεχνικές βαθιάς μάθησης.

#### 4. Οθόνη εισαγωγής νέου βιβλίου

Σε αυτή την οθόνη ο χρήστης έχει τη δυνατότητα να εισαγάγει ένα νέο βιβλίο και έπειτα να δει τα αποτελέσματα των Νευρωνικών Δικτύων βαθιάς μάθησης που έχουν υλοποιηθεί σε αυτή τη διπλωματική εργασία.

Τα πιθανά σενάρια χρήσης της εφαρμογής είναι τα εξής:

1. Ο χρήστης όταν βρίσκεται στην αρχική σελίδα έχει δύο επιλογές: είτε να μεταφερθεί στη σελίδα εξερεύνησης όλων των βιβλίων είτε στη σελίδα εισαγωγής ενός νέου.
2. Όταν ο χρήστης είναι στη σελίδα εμφάνισης όλων των βιβλίων, θα βλέπει μόνο τις λίγες και απαραίτητες πληροφορίες για το κάθε ένα για να μπορέσει να καταλάβει την κατηγορία του και το θεματικό του περιεχόμενο.

3. Όταν ο χρήστης έχει ανοίξει ένα συγκεκριμένο σύγγραμμα, τότε θα βλέπει όλες τις πληροφορίες για αυτό, οι οποίες περιέχονται στη βάση, καθώς και κάποιες ετικέτες και κατηγορίες στις οποίες ανήκει. Επιπλέον, θα μπορεί να δει και συναφή βιβλία τα οποία έχουν εντοπιστεί με τεχνικές βαθιάς μάθησης.
4. Όταν ο χρήστης το επιθυμεί, θα του παρέχεται η δυνατότητα να εισαγάγει ένα νέο βιβλίο για να διαπιστώσει σε ποια κατηγορία θα το κατατάξει το σύστημα.

## 6.2.2 Υποδομή Server και Βάσης Δεδομένων (Backend)

Στην προηγούμενη ενότητα, παρουσιάστηκαν οι βασικές λειτουργικότητες που προσφέρει η εφαρμογή στους χρήστες. Αναλύθηκαν υπηρεσίες και σενάρια χρήσης της εφαρμογής, από την πλευρά της διεπαφής του χρήστη. Σε αυτή την ενότητα θα ασχοληθούμε με την υλοποίηση και το σχεδιασμό του backend server και της βάσης δεδομένων της εφαρμογής. Θα αναλυθούν, δηλαδή, οι ενέργειες που πρέπει να γίνουν από την πλευρά του backend server και ο τρόπος που αλληλεπιδρούν με το υπόλοιπο σύστημα.

### 6.2.2.1 Backend Server

Κάθε αίτημα που κάνει ο χρήστης αρχικά κωδικοποιείται από τον frontend server έτσι ώστε να μπορέσει να δρομολογηθεί προς τον backend server για επεξεργασία. Όμως, αφού οι frontend και backend είναι δύο διαφορετικοί servers, θα χρειαστούμε ένα μέσο με το οποίο θα επικοινωνούν αμφίδρομα. Τα APIs (Application Programming Interface) είναι ένα ενδιάμεσο λογισμικό που επιτρέπει την επικοινωνία μεταξύ δύο εφαρμογών. Έτσι, με τα κατάλληλα APIs μπορεί να επιτευχθεί η επικοινωνία αυτή, ενώ κάθε API αφορά ένα διαφορετικό αίτημα του χρήστη. Το API διαθέτει κάποιες παραμέτρους και μια τιμή επιστροφής. Επομένως, ο frontend server αν θέλει να δει τις αναλυτικές πληροφορίες για ένα σύγγραμμα, θα χρησιμοποιήσει το αντίστοιχο API δίνοντας ως παράμετρο το αναγνωριστικό id του βιβλίου και ως επιστροφή θα περιμένει τις αναλυτικές πληροφορίες του βιβλίου. Τα APIs λοιπόν που υλοποιήθηκαν είναι τα εξής:

- **AllIDs**

Επιστρέφει μια λίστα με όλα τα αναγνωριστικά IDs όλων των βιβλίων

- **short info**

Επιστρέφει μερικές απαραίτητες πληροφορίες για ένα βιβλίο

- **multiple short info**

Επιστρέφει μερικές απαραίτητες πληροφορίες για ένα σύνολο βιβλίων

- **cover image**

Επιστρέφει τη φωτογραφία εξωφύλλου ενός βιβλίου

- **all info**

Επιστρέφει όλες τις πληροφορίες για ένα βιβλίο

- **classifyBook**

Με είσοδο έναν πίνακα περιεχομένων ενός νέου βιβλίου, επιστρέφει την κατηγορία αυτού, χρησιμοποιώντας τα νευρωνικά εκπαιδευμένα μοντέλα

- **similarBooks**

Χρησιμοποιώντας μοντέλα που αναπτύχθηκαν στη διπλωματική, εντοπίζει τα συναφή συγγράμματα ενός βιβλίου

Όλες οι τιμές που μεταδίδονται μέσω των APIs έχουν τη μορφή αρχείου JSON, το οποίο είναι γνωστό και ως δομή dictionary στις γλώσσες προγραμματισμού. Ο backend server θα υλοποιηθεί σε γλώσσα python με τη χρήση του flask framework λόγω της ευκολίας και απλότητάς του στην χρήση αλλά και των δυνατοτήτων που προσφέρει στην ανάπτυξη εφαρμογών.

#### 6.2.2.2 Βάση Δεδομένων

Η βάση δεδομένων Springer που χρησιμοποιούμε για την επεξεργασία και προβολή ηλεκτρονικών συγγραμμάτων διαθέτει τη δική της βάση δεδομένων, η οποία είναι μια SQL database. Επομένως, για όλες τις πληροφορίες που θα χρειαστούμε για κάθε βιβλίο έχουμε ήδη διαθέσιμη τη βάση αυτή στην οποία έχουμε πρόσβαση, χρησιμοποιώντας τα κατάλληλα credentials που μας έχουν δοθεί. Πιο συγκεκριμένα, από τη βάση αυτή ανακτούμε τις ακόλουθες πληροφορίες για κάθε διαφορετικό ηλεκτρονικό σύγγραμμα:

- το μοναδικό αναγνωριστικό id του

- τον τίτλο
- τον πίνακα περιεχομένων
- την περίληψη
- το θεματικό του περιεχόμενο

Για επιπλέον δεδομένα που θα χρειαστεί να αποθηκεύσουμε, θα δημιουργήσουμε μια βάση δεδομένων MongoDB, η οποία θα φιλοξενηθεί σε κάποιο cluster στο cloud. Σε αυτήν τη βάση οι πληροφορίες θα υπάρχουν σε μορφή JSON και θα περιέχει δεδομένα και πληροφορίες, τα οποία είναι αποτέλεσμα έρευνας και επεξεργασίας της διπλωματικής αυτής εργασίας. Πιο συγκεκριμένα, στη βάση αυτή για κάθε ηλεκτρονικό σύγγραμμα έχουμε τις εξής πληροφορίες:

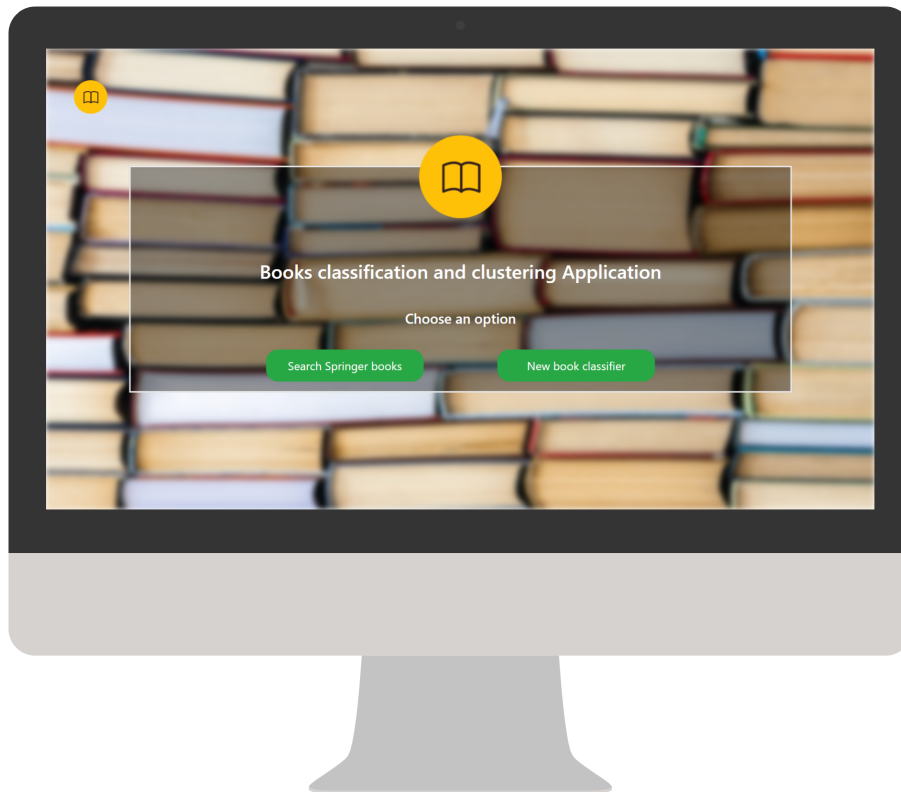
- το μοναδικό αναγνωριστικό id του
- την κατηγορία που ανήκει
- το αναγνωριστικό id της κατηγορίας που ανήκει

## 6.3 Χαρακτηριστικά και Λειτουργικότητα Οθονών Εφαρμογής

Σε αυτή την ενότητα θα παρουσιαστούν οι οθόνες χρήστη της web εφαρμογής από desktop και από smartphone. Όλη η εφαρμογή έχει δημιουργηθεί εξ αρχής (from scratch) με React Framework. Έχει δοθεί ιδιαίτερη προσοχή στο γραφικό περιβάλλον και στην εμφάνιση με σκοπό την καλύτερη δυνατή εμπειρία του χρήστη κατά την πλοήγησή του στην εφαρμογή.

### 6.3.1 Αρχική Σελίδα

Η παρουσίαση της εφαρμογής θα ξεκινήσει με την αρχική σελίδα. Παρακάτω φαίνεται η αρχική σελίδα όταν ο χρήστης πλοηγείται από υπολογιστή.



Σχήμα 6.1: Αρχική σελίδα εφαρμογής από Desktop.

Η σελίδα αυτή έχει δύο βασικές επιλογές:

### 1. Search Springer books

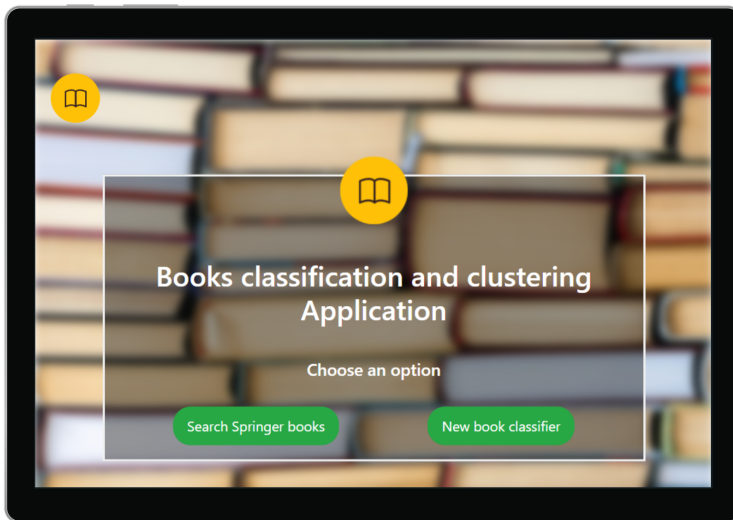
Επιλέγοντας αυτό το κουμπί, ο χρήστης μεταφέρεται στη σελίδα πλοήγησης, αναζήτησης και εξερεύνησης των ηλεκτρονικών βιβλίων.

### 2. New book classifier

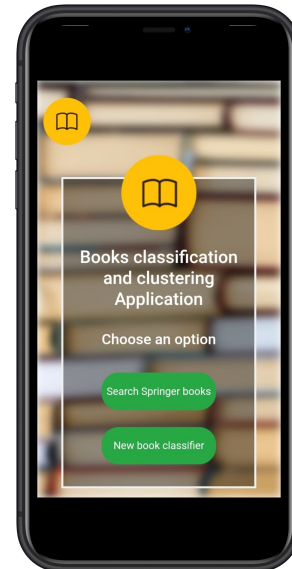
Με αυτό το κουμπί, ο χρήστης μεταφέρεται στη σελίδα εισαγωγής και ταξινόμησης ενός νέου βιβλίου με τη βοήθεια του μοντέλου ταξινόμησης που έχει υλοποιηθεί στο πρώτο μέρος της εργασίας.

Στην ανάπτυξη του λογισμικού της εφαρμογής, έχουν χρησιμοποιηθεί τεχνικές με τις οποίες η σελίδα αλλάζει και προσαρμόζεται με βάση το μέγεθος της οθόνης του χρήστη και τη συσκευή από την οποία πλοηγείται. Κατά την πλοήγηση από υπολογιστή

μπορούν να προβληθούν πολλές πληροφορίες ταυτόχρονα στην οθόνη ενώ από κινητό κάθε γραμμή θα πρέπει να περιέχει ένα στοιχείο και τα κουμπιά θα πρέπει να είναι μεγάλα εφόσον το κινητό θα είναι αφής. Παρακάτω φαίνεται η αρχική σελίδα της εφαρμογής, όταν ο χρήστης χρησιμοποιεί tablet ή smartphone.



Σχήμα 6.2: Αρχική σελίδα εφαρμογής από Tablet.

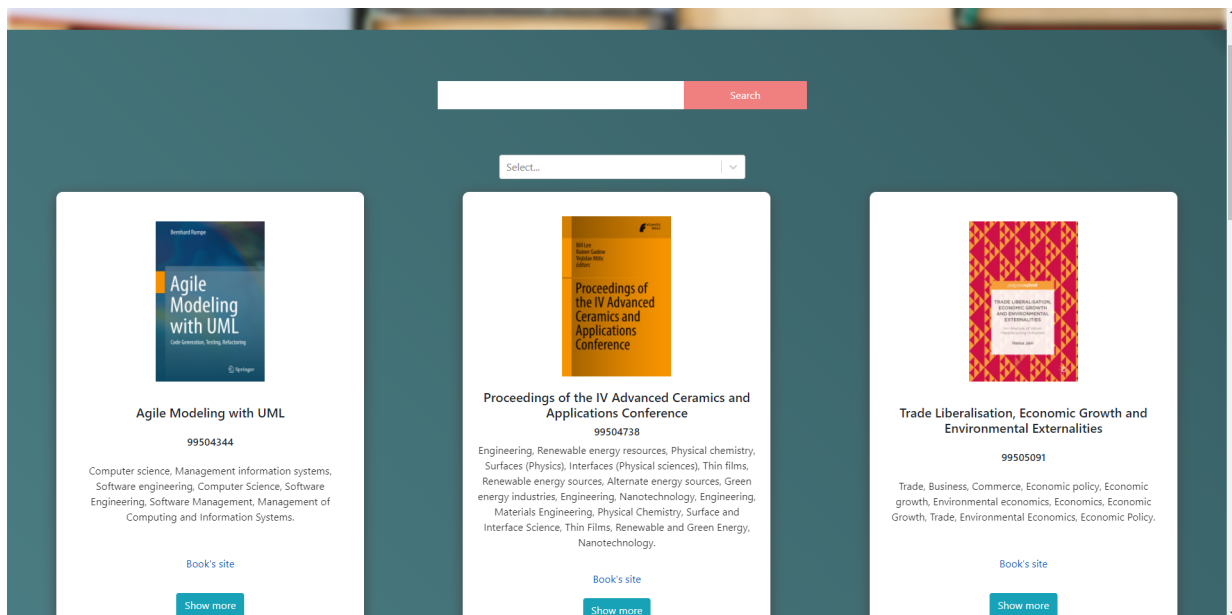


Σχήμα 6.3: Αρχική σελίδα εφαρμογής από Smartphone.

### 6.3.2 Προβολή πολλαπλών βιβλίων

Η σελίδα αυτή εμφανίζεται στην περίπτωση που ο χρήστης επιλέξει το εικονίδιο “Search Springer books”. Η σελίδα αυτή προβάλλει όλα τα βιβλία που περιέχει η βάση δεδομένων. Για κάθε βιβλίο παρέχει τις εξής πληροφορίες:

- την εικόνα εξωφύλλου του (εάν αυτή είναι διαθέσιμη)
- το ξεχωριστό αναγνωριστικό id του βιβλίου, το οποίο είναι καταχωρημένο στη βάση Springer
- τις λέξεις-κλειδιά του βιβλίου που αφορούν το θεματικό του περιεχόμενο.

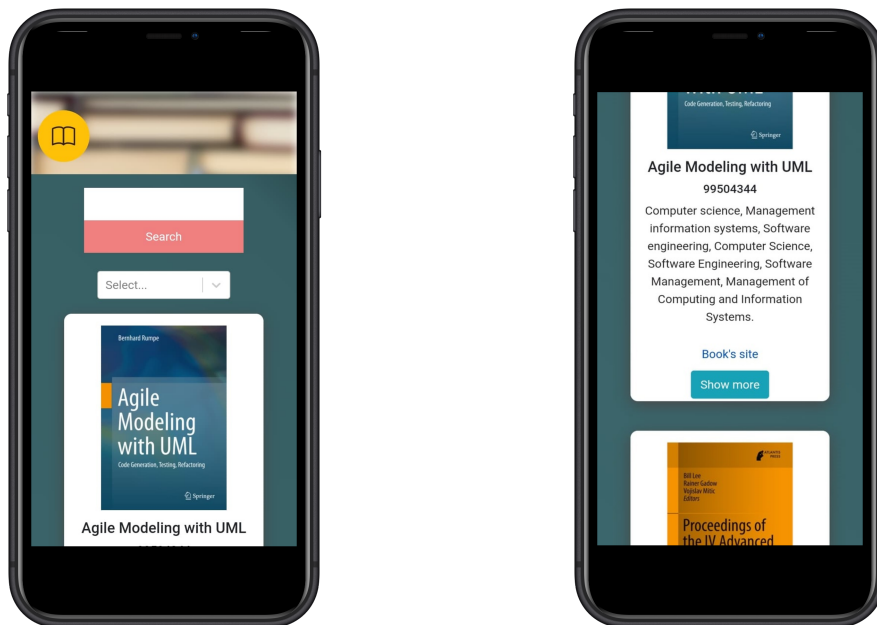


Σχήμα 6.4: Σελίδα προβολής πολλαπλών βιβλίων από Desktop.

Για κάθε βιβλίο, ο χρήστης έχει την επιλογή είτε να μεταφερθεί στην παρουσίασή του στην επίσημη ιστοσελίδα του αποθετηρίου Springer (“Book’s site”) είτε να μεταφερθεί στη σελίδα της εφαρμογής στην οποία παρουσιάζονται όλες οι πληροφορίες για αυτό (“Show more”). Η μπάρα αναζήτησης και το κουμπί επιλογής Cluster δεν έχουν αναπτυχθεί στα πλαίσια αυτής της εργασίας και αποτελούν δύο απαραίτητες μελλοντικές υλοποιήσεις που επιβάλλεται να πραγματοποιηθούν προκειμένου να λειτουργήσει η εφαρμογή.

Η αντίστοιχη σελίδα, στην περίπτωση που ο χρήστης χρησιμοποιεί κινητό, προβάλλεται όπως φαίνεται παρακάτω:



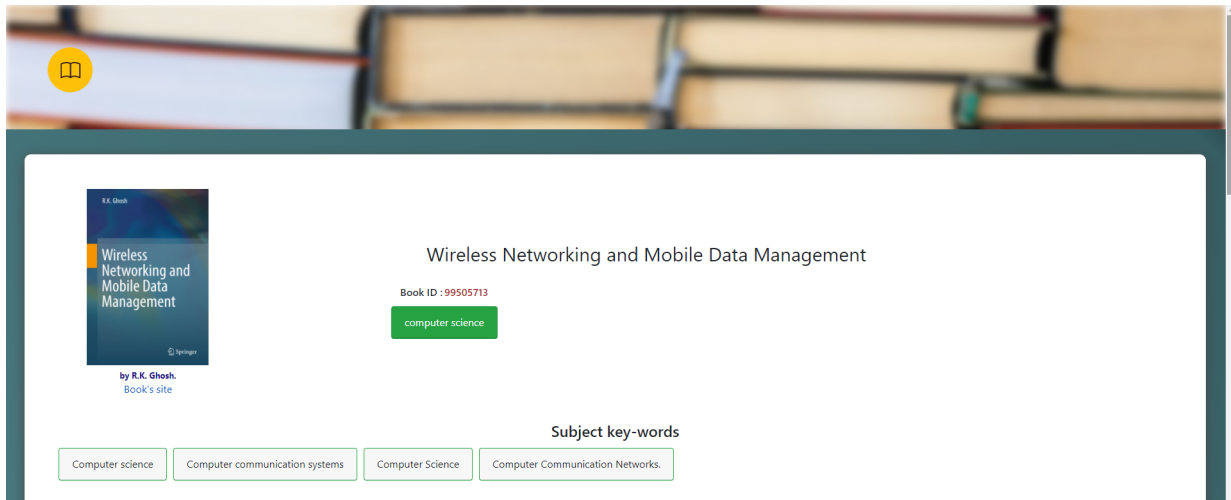


Σχήμα 6.5: Σελίδα προβολής πολλαπλών βιβλίων από Smartphone.

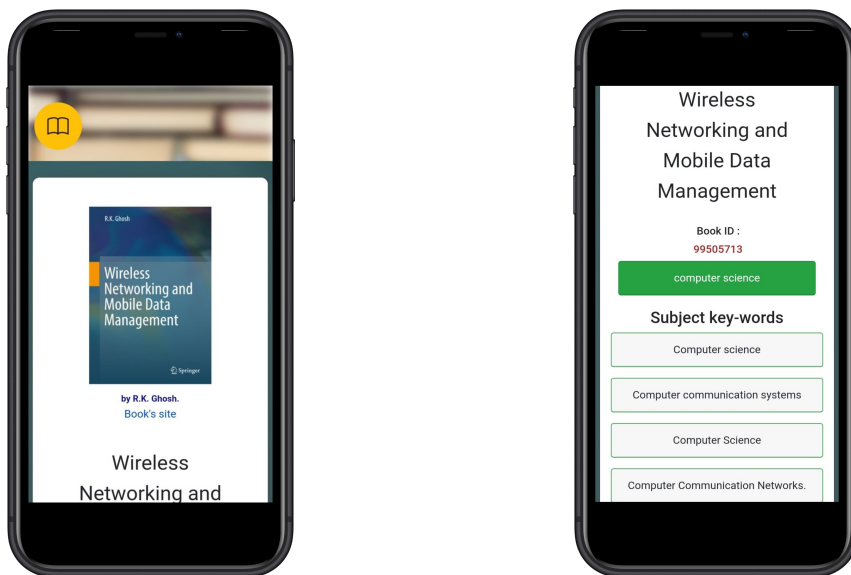
### 6.3.3 Αναλυτική προβολή βιβλίου

Η σελίδα αυτή εμφανίζεται μόλις ο χρήστης πατήσει την επιλογή “Show more”. Η σελίδα αυτή προβάλλει όλες τις διαθέσιμες πληροφορίες για ένα συγκεκριμένο ηλεκτρονικό βιβλίο. Αναλυτικότερα, οι πληροφορίες που εμπεριέχει η σελίδα είναι οι εξής:

- Οι πληροφορίες που δείχνει και η σελίδα προβολής πολλαπλών βιβλίων, δηλαδή:
  - την εικόνα εξωφύλλου του (εάν αυτή είναι διαθέσιμη)
  - το ξεχωριστό αναγνωριστικό id του βιβλίου, το οποίο είναι καταχωρημένο στη βάση Springer
  - τις λέξεις κλειδιά του βιβλίου που αφορούν το θεματικό του περιεχόμενο
- Ο συγγραφέας του βιβλίου
- Η κατηγορία στην οποία ανήκει το βιβλίο, με βάση την κατηγοριοποίηση που πραγματοποιήθηκε στο Κεφάλαιο 3



Σχήμα 6.6: Αναλυτική προβολή βιβλίου (βασικά στοιχεία) από Desktop.

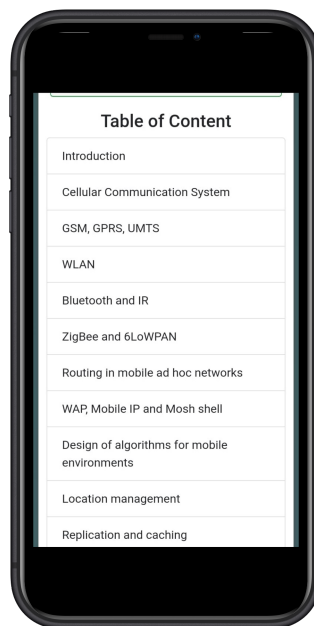


Σχήμα 6.7: Αναλυτική προβολή βιβλίου (βασικά στοιχεία) από Smartphone.

- Ο πίνακας περιεχομένων (table of contents) του βιβλίου

Introduction
Cellular Communication System
GSM, GPRS, UMTS
WLAN
Bluetooth and IR
ZigBee and 6LoWPAN
Routing in mobile ad hoc networks
WAP, Mobile IP and Mosh shell
Design of algorithms for mobile environments
Location management
Replication and caching
Broadcast disk and data dissemination
Indexing in air
File system for mobile environment
Smart environments.

Σχήμα 6.8: Αναλυτική προβολή βιβλίου (πίνακα περιεχομένων) από Desktop.



The image shows a smartphone displaying the same Table of Content as in Figure 6.8. The text is scaled down to fit the smaller screen of the device.

Introduction
Cellular Communication System
GSM, GPRS, UMTS
WLAN
Bluetooth and IR
ZigBee and 6LoWPAN
Routing in mobile ad hoc networks
WAP, Mobile IP and Mosh shell
Design of algorithms for mobile environments
Location management
Replication and caching

Σχήμα 6.9: Αναλυτική προβολή βιβλίου (πίνακα περιεχομένων) από Smartphone.

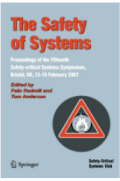
- Η περιγραφή (abstract) του βιβλίου

- τα παρόμοια, με αυτό, βιβλία τα οποία εντόπισε το μοντέλο που αναπτύχθηκε στο πρώτο μέρος της εργασίας

**Abstract**

This book examines two main topics, namely, Wireless Networking and Mobile Data Management. It is designed around a course the author began teaching to senior undergraduate and masters students at the Department of Computer Science & Engineering of the Indian Institute of Technology Kanpur. The first part of the book, consisting of eight chapters, including the introduction, focuses exclusively on wireless networking aspects. It begins with cellular communication systems, which provided the foundation of wireless networking principles. Three subsequent chapters are devoted to the Global System for Mobile communication (GSM), Wireless Local Area Network (WLAN), Bluetooth, infrared (IR), ZigBee and 6LoWPAN protocols. There is also a chapter on routings in ad hoc networks, an area that is currently being intensively researched due to its potential applications in areas of vehicular network, traffic management, tactical and military systems. Furthermore, the book discusses mobile operating systems and wireless network application level protocols such as Wireless Application Protocols (WAP), Mobile IP and Mosh. The second part highlights mobile data management. It addresses the issues like location management, the importance of replication and caching in mobile environments, the concept of broadcast disk and indexing in air, storage systems for sharing data in mobile environments, and building smart environments. Given that the design of algorithms is the key to applications in data management this part begins with a chapter on the type of paradigm shift that has been introduced in the design of algorithms, especially due to asymmetry in mobile environments. Lastly, the closing chapter of the book explores smart environments, showing the readers how wireless technology and mobile data management can be combined to provide optimum comfort for human life. Though the book has been structured as a monograph, it can be used both as a textbook and as a reference material for researchers and developers working in the area.

**Similar Books**

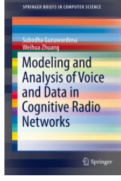


**The Safety of Systems**  
99537149

Computer science, Computer system failures, Software engineering, Computer Science, Software Engineering, System Performance and Evaluation.

[Book's site](#)

[Show more](#)

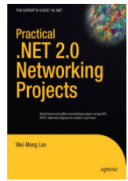


**Modeling and Analysis of Voice and Data in Cognitive Radio Networks**  
99522236

Computer science, Computer communication systems, Electrical engineering, Computer Science, Computer Communication Networks, Communications Engineering, Networks.

[Book's site](#)

[Show more](#)



**Practical .NET 2.0 Networking Projects**  
99541557

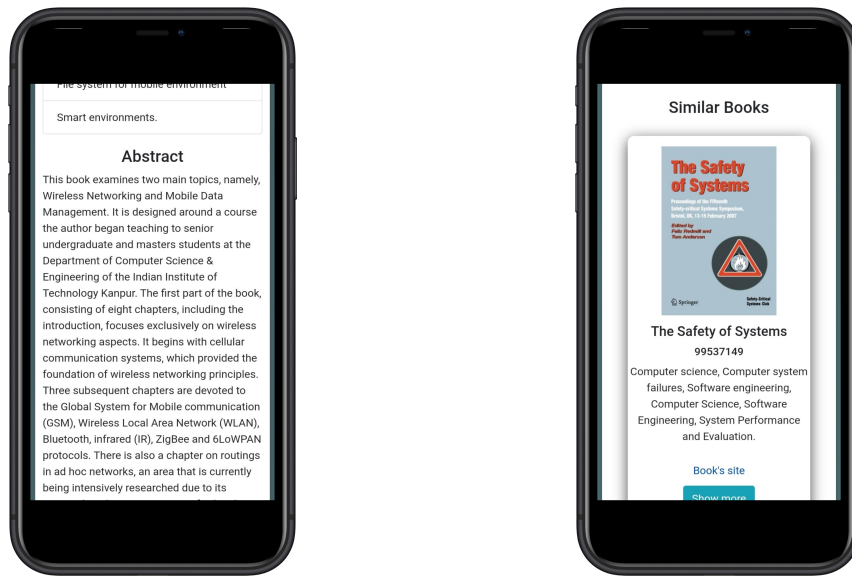
Computer science, Software engineering, Computer Science, Software Engineering/Programming and Operating Systems.

[Book's site](#)

[Show more](#)

Σχήμα 6.10: Αναλυτική προβολή βιβλίου (περίληψη και παρόμοια βιβλία) από Desktop.

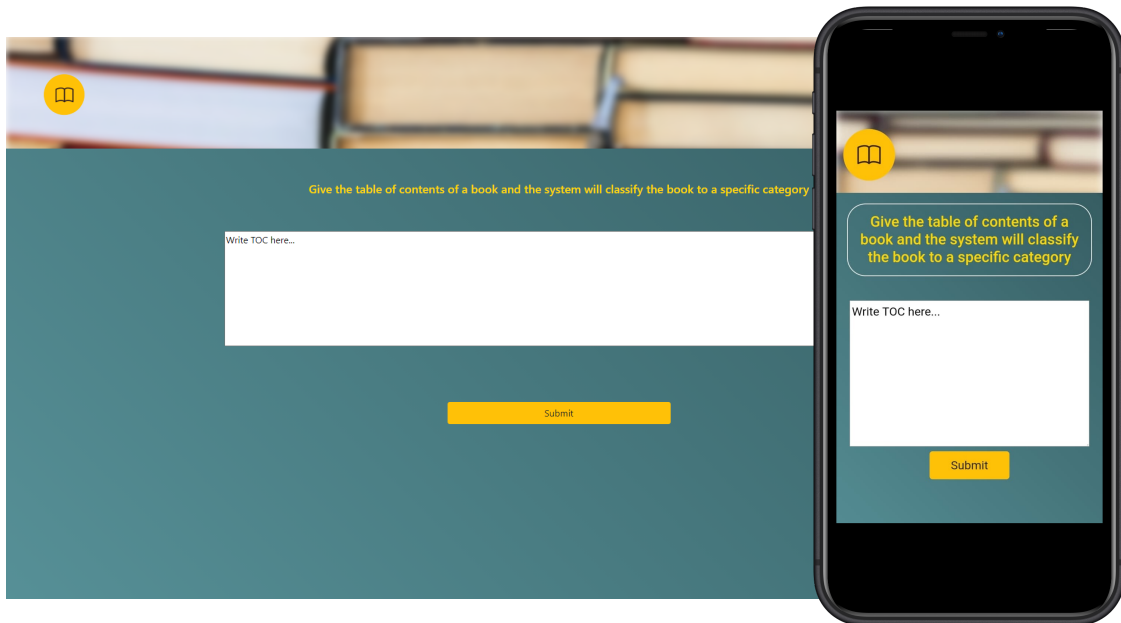
Στο σημείο αυτό μπορεί να επιβεβαιωθεί η αποτελεσματικότητα του μοντέλου εύρεσης συναφών συγγραμμάτων. Διαπιστώνεται ότι το μοντέλο έχει εντοπίσει βιβλία συναφή με το βιβλίο “Wireless Networking and Mobile Management” ανιχνεύοντας συγγράμματα με θεματικό περιεχόμενο σχετικό με τα δίκτυα.



Σχήμα 6.11: Αναλυτική προβολή βιβλίου (περίληψη και παρόμοια βιβλία) από Smartphone.

#### 6.3.4 Ταξινόμηση νέου βιβλίου

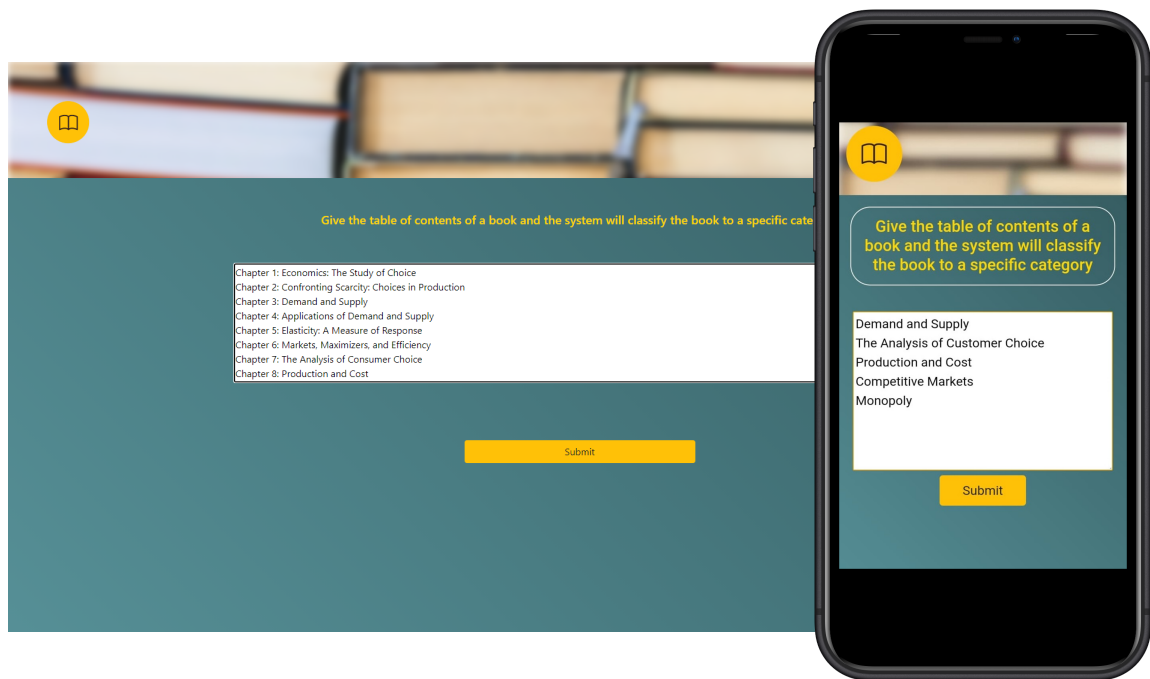
Εάν ο χρήστης πατήσει την επιλογή New book classifier, τότε θα βρεθεί σε αυτήν την σελίδα της εφαρμογής. Σε αυτήν την υποενότητα θα παρουσιαστούν όλα τα πιθανά σενάρια και οι αντίστοιχες οθόνες χρήστη αυτής της σελίδας αφενός από υπολογιστή και αφετέρου από κινητό.



Σχήμα 6.12: Σελίδα ταξινόμησης νέου βιβλίου από Desktop και Smartphone.

Όπως γίνεται κατανοητό από την παραπάνω εικόνα, ο χρήστης μπορεί να εισαγάγει στο πλαίσιο που εμφανίζεται στην οθόνη τον πίνακα περιεχομένων ενός νέου βιβλίου έτσι ώστε η εφαρμογή, με την βοήθεια του αντίστοιχου μοντέλου, να το ταξινομήσει σε μια από τις διαθέσιμες κατηγορίες.

Παρακάτω παρουσιάζεται αναλυτικά η διαδικασία εισαγωγής του πίνακα περιεχομένων και η ταξινόμησή του από υπολογιστή και από κινητό.



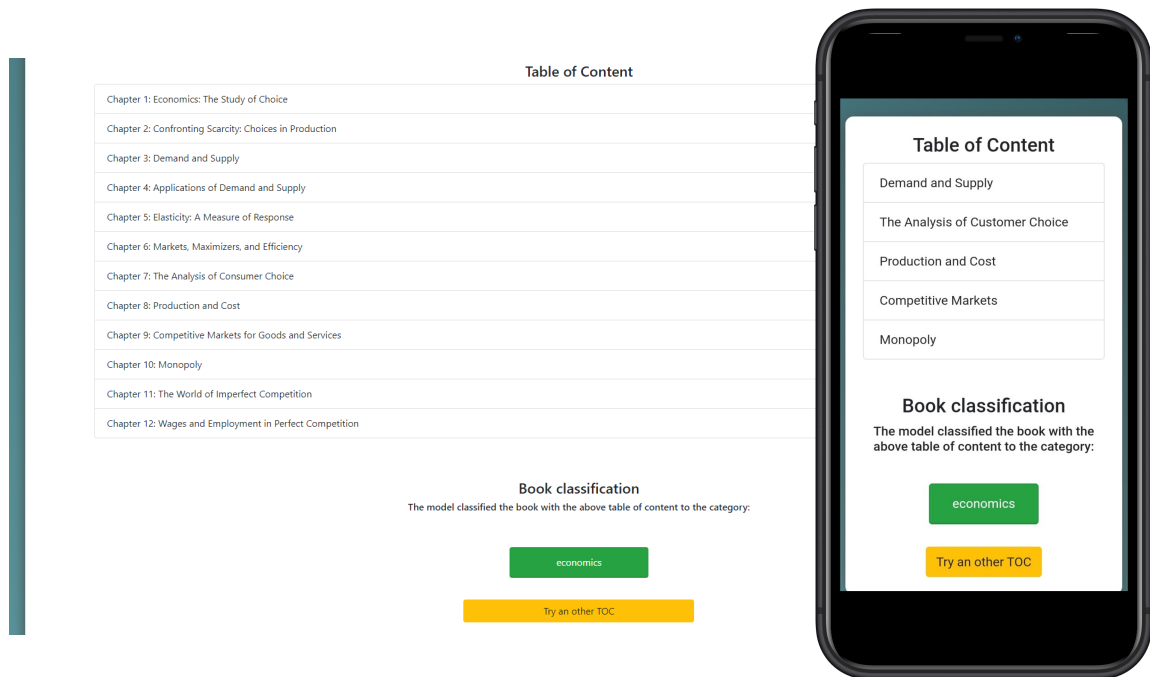
Σχήμα 6.13: Εισαγωγή πίνακα περιεχομένων νέου βιβλίου από Desktop και Smartphone.

Στο παράδειγμα αυτό, έγινε εισαγωγή του πίνακα περιεχομένων ενός βιβλίου οικονομικών. Στη συνέχεια πραγματοποιείται η διαδικασία ταξινόμησης από τον backend server της εφαρμογής. Το μοντέλο που επιλέχθηκε να ταξινομήσει τα βιβλία, είναι αυτό το οποίο εκπαιδεύτηκε στις εξής επτά κατηγορίες:

1. Επιστήμη Υπολογιστών
2. Φυσικές Επιστήμες
3. Ιατρική
4. Μηχανική
5. Μαθηματικά
6. Οικονομικά

## 7. Βιοεπιστήμες

Το μοντέλο των 26 κλάσεων δε διαθέτει την ίδια ακρίβεια με αυτό των 7, λόγω έλλειψης δεδομένων εκπαίδευσης. Κάτι τέτοιο ενδεχομένως θα οδηγούσε σε λανθασμένες προβλέψεις. Στην επόμενη εικόνα θα αποτυπωθούν τα αποτελέσματα του ταξινομητή.



Σχήμα 6.14: Αποτέλεσμα ταξινόμησης νέου βιβλίου από Desktop και Smartphone.

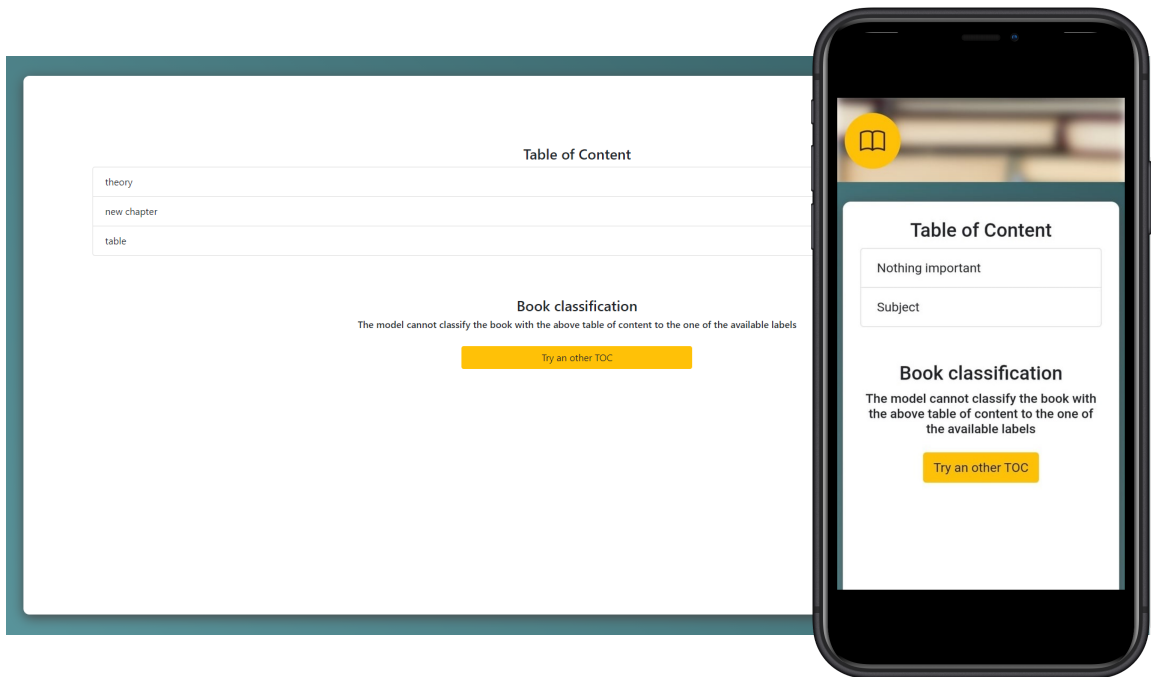
Σε αυτό το σημείο επιβεβαιώνεται η ευστοχία του μοντέλου να διακρίνει τα χαρακτηριστικά και να ταξινομεί σωστά τα βιβλία με βάση τον πίνακα περιεχομένων τους. Όσον αφορά το συγκεκριμένο βιβλίο, ανίχνευσε ότι ανήκει στην κατηγορία “Οικονομικά”, κάτι το οποίο ισχύει.

Σε κάποιες περιπτώσεις όμως τα δεδομένα εισόδου είναι λανθασμένα, ελλιπή ή δεν είναι ικανά να προσδιορίσουν την κατηγορία στην οποία ανήκει το αντίστοιχο βιβλίο. Χρησιμοποιώντας λοιπόν τα αποτελέσματα της ταξινόμησης και συγκεκριμένα αυτά που έδωσε η συνάρτηση softmax, θα τεθεί ένα κάτω όριο της κλάσης που υπερισχύει. Πιο συγκεκριμένα, όταν το ποσοστό πρόβλεψης της softmax για την υπερισχύουσα



κατηγορία είναι κάτω από 50%, τότε η πρόβλεψη του μοντέλου θα απορρίπτεται και θα εμφανίζεται στον χρήστη ότι το μοντέλο δεν μπόρεσε να ταξινομήσει το βιβλίο.

Παρακάτω απεικονίζεται η περίπτωση που μόλις αναλύθηκε:



Σχήμα 6.15: Αποτέλεσμα ταξινόμησης στην περίπτωση ελλιπούς εισόδου νέου βιβλίου από Desktop και Smartphone.

## Συμπεράσματα

Η ανάλυση και η επεξεργασία του πίνακα περιεχομένων των συγγραμμάτων αποδείχθηκε ότι μπορεί να συμβάλει θετικά στην ανάπτυξη μοντέλων βαθιάς μάθησης και να αποδώσει ουσιώδη αποτελέσματα. Στη συλλογή 56.401 ψηφιακών βιβλίων του Εκδοτικού Οίκου Springer, με την οποία εργαστήκαμε, τα τελικά μοντέλα ταξινόμησης 26 και 7 κλάσεων αποδεικνύονται αποτελεσματικά και ικανά να προβλέψουν σε ποια κατηγορία κατατάσσεται ένα βιβλίο αντλώντας δεδομένα αποκλειστικά από τον πίνακα περιεχομένων του. Τα ήδη εκπαιδευμένα μοντέλα, τα οποία διαθέτουν κάποια βάση γνώσης, βελτιώνουν αρκετά την απόδοση του συστήματος, όπως φάνηκε και από την εισαγωγή του μοντέλου word2vec ως πρώτου επιπέδου των μοντέλων. Η δημιουργία νέων χαρακτηριστικών για κάθε βιβλίο παρήγαγε νέα αποτελέσματα τα οποία επιβεβαιώνουν την ορθότητα των τελικών μοντέλων ταξινόμησης, καθώς προέκυψαν από τις ενεργοποιήσεις των εσωτερικών τους επιπέδων. Αξιοποιώντας τα νέα χαρακτηριστικά κατασκευάστηκαν μοντέλα clustering τα οποία τροφοδοτούνται από τα δεδομένα αυτά. Το τελικό μοντέλο που επιλέχθηκε, προβαίνει σε προεπεξεργασία των δεδομένων, δημιουργεί πέντε νέες κατηγορίες (clusters) και εντάσσει όλα τα βιβλία σε αυτές. Το σύστημα εύρεσης συναφών συγγραμμάτων σχεδιάστηκε με γνώμονα την αρχή ότι τα βιβλία που ο αλγόριθμος ανιχνεύει ως εγγύτερα με βάση τα χαρακτηριστικά τους, εκλαμβάνονται ως συναφή ως προς το περιεχόμενό τους. Από τα αποτελέσματα συνάγεται ότι το σύστημα λειτουργεί αποτελεσματικά αφού είναι σε θέση να προτείνει μια λίστα στοχευμένων βιβλίων για κάθε επιλεγμένο. Όλα τα παραπάνω μοντέλα και συστήματα, καθώς και όλες οι διαθέσιμες πληροφορίες για κάθε ηλεκτρονικό βιβλίο πρέπει να είναι διαθέσιμες για εξερεύνηση, αναζήτηση και δοκιμή. Η web εφαρμογή που υλοποιήθηκε, λειτουργεί άρτια σε υπολογιστή, tablet και κινητό και εξυπηρετεί κάθε ανάγκη του χρήστη για διάφορα σενάρια χρήσης.

Το σύστημα, τα μοντέλα και τα αποτελέσματα μπορούν να βελτιωθούν αξιοποιώντας δεδομένα και πληροφορίες, όπως η περίληψη, οι λέξεις-κλειδιά του κάθε βιβλίου και

λοιπά χαρακτηριστικά, τα οποία δεν χρησιμοποιήθηκαν στην παρούσα εργασία. Το έργο που παρουσιάστηκε στη διπλωματική εργασία δύναται να εξελιχθεί με την ενσωμάτωση των νέων πληροφοριών, παρέχοντας νέες δυνατότητες για την αύξηση της απόδοσης του συστήματος αλλά και προσθήκης νέων features στην web εφαρμογή.

# Βιβλιογραφία

- [1] Βλαχάβας Ιωάννης Κεφαλάς Πέτρος Βασιλειάδης Νικόλαος Κόκκορας Φώτης Σαακελαρίου Ηλίας. *Τεχνητή Νοημοσύνη*. Εκδόσεις Πανεπιστημίου Μακεδονίας, Εγνατία 156, 540 06 Θεσσαλονίκη, Γ Έκδοση 2011. σελίδες 3, 373 (Σχήμα 19.1), 379.
- [2] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*. σελίδα 1930 Σχήμα 1.
- [3] Bharath Raj. Data Augmentation: How to use Deep Learning when you have Limited Data. *KDnuggets* <https://www.kdnuggets.com>, 2018. Εικόνα από τη δημοσίευση.
- [4] Sumeet Pathania. A Brief Introduction To Self-Driving Cars. *Medium* <https://medium.com>, 2019. Εικόνα από την ενότητα 1. Computer Vision.
- [5] Wikipedia. [https://en.wikipedia.org/wiki/File:Conv\\_layers.png](https://en.wikipedia.org/wiki/File:Conv_layers.png). Αποτέλεσμα αναζήτησης για τα CNN neural networks, εικόνα απεικόνισης επιπέδων ενός συνελκτικού νευρωνικού δικτύου με όνομα "File:Conv layers.png".
- [6] Αθανάσιος Τάγαρης. *Ερμηνεύσιμα Μοντέλα Βαθιάς Μάθησης για την Ανάλυση Οπτικών Δεδομένων*. Διδακτορική Διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο, Ιούνιος 2020, σχήμα 2.15, σχήμα 2.16.
- [7] missinglink.ai. <https://missinglink.ai>. Keras Conv1D: Working with 1D Convolutional Neural Networks in Keras.
- [8] Wikipedia. [https://en.wikipedia.org/wiki/File:Recurrent\\_neural\\_network\\_unfold.svg](https://en.wikipedia.org/wiki/File:Recurrent_neural_network_unfold.svg). Αποτέλεσμα αναζήτησης για τα RNN neural networks,

εικόνα απεικόνισης ενός κόμβου σε ένα Αναδρομικό Νευρωνικό Δίκτυο με όνομα "File:Recurrent neural network unfold.svg".

- [9] Springer. <https://www.springer.com/gp>. Αρχική σελίδα του Springer.
- [10] SKlearn. <https://scikit-learn.org>. SKlearn Documentation - 2.3.1. Overview of clustering methods.
- [11] Η Πύλη για την Ελληνική Γλώσσα. <http://www.greek-language.gr>. Αποτέλεσμα αναζήτησης της λέξης νοημοσύνη.
- [12] Wikipedia. <https://en.wikipedia.org>. Αποτέλεσμα αναζήτησης για τον Clive Humby.
- [13] Κατερίνα Γεωργούλη. *Τεχνητή Νοημοσύνη - Μια εισαγωγική προσέγγιση*. Αποθετήριο Κάλλιπος <https://repository.kallipos.gr>, 2015. Κεφάλαιο 4 - Μηχανική Μάθηση, σελίδες 1.
- [14] Απόδοση Ελένη Γκαγκατσιού Simon Haykin. *Νευρωνικά Δίκτυα και Μηχανική Μάθηση*. Εκδόσεις Παπασωτηρίου, Στουρνάρη 49Α, 106 82 Αθήνα, Τρίτη Έκδοση 2010. σελίδες 1, 13.
- [15] Θεωρήσης Π. Δασκαλόπουλος. Παράρτημα 1: Νευρώνας και Νευροδιαβιβαστής. Γενικά στοιχεία της νευρωνικής λειτουργίας. [http://psi-gr.tripod.com/choc\\_20\\_app\\_neuro.html](http://psi-gr.tripod.com/choc_20_app_neuro.html), 2006. Κείμενο του άρθρου για νευρώνες εγκεφάλου και Σχήμα 1.4.
- [16] Lavanya Shukla. *Designing Your Neural Networks. Towards Data Science* <https://towardsdatascience.com>, 2019. Εικόνα από την ενότητα 1. Basic Neural Network Structure.
- [17] James Martin Daniel Jurafsky. *Speech and Language Processing*. Prentice Hall, 2008.
- [18] Γεωργούλη Αικατερίνη. *Επεξεργασία και Κατανόηση Φυσικής Γλώσσας*. Αποθετήριο Κάλλιπος <https://repository.kallipos.gr>, 2015. Κεφάλαιο 7 - Επεξεργασία και Κατανόηση Φυσικής Γλώσσας, σελίδες 1-2.

- [19] Ian Goodfellow, Yoshua Bengio, Aaron Courville. *Deep Learning (Adaptive Computation and Machine Learning series)*. The MIT Press, Illustrated edition 2016. σελίδα 226.
- [20] Diederik P. Kingma, Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2015.
- [21] Rachel Draelos. The History of Convolutional Neural Networks. *GLASS BOX* <https://glassboxmedicine.com>, Απρίλιος 13, 2019.
- [22] Wikipedia. <https://el.wikipedia.org/wiki>. Αποτέλεσμα αναζήτησης για την πράξη της συνέλιξης.
- [23] Μήτσιος Γεώργιος. *Αναδρομικά Νευρωνικά Δίκτυα και αυτόματη παραγωγή Hash-tag από Tweet του ενΤωιττερ*. Διπλωματική Εργασία, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Θεσσαλονίκη 2017, σελίδα 31 Εικόνα 6.
- [24] towards data science. <https://towardsdatascience.com>. Illustrated Guide to LSTM's and GRU's: A step by step explanation.
- [25] Giannopoulou E, Mitrou N. *An AI-based methodology for automatic classification of a multi-class corpus using information from the table of contents*.
- [26] Google wold2vec. <https://code.google.com/archive/p/word2vec>. εκπαιδευμένο μοντέλο word2vec της Google.
- [27] PCA. <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>. Η μέθοδος PCA της βιβλιοθήκης SKlearn.
- [28] Wikipedia. <https://el.wikipedia.org/wiki>. Αποτέλεσμα αναζήτησης για την Αλληλεπίδραση ανθρώπου-υπολογιστή.

# Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια

κτλ.	και τα λοιπά
π.χ.	παραδείγματος χάρη
βλ.	βλέπε
TN	Τεχνητή Νοημοσύνη
ML	Machine Learning
NN	Neural Network
ΕΦΓ	Επεξεργασία Φυσικής Γλώσσας
NLP	Natural Language Processing
NLU	Natural Language Understanding
ReLU	Rectified Linear Unit
MSE	Mean Squared Error
SGD	Stochastic Gradient Descent
TP	True Positive
FP	False Positive
FN	False Negative
TN	True Negative
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
HCI	Human-Computer Interaction
API	Application Programming Interface
PCA	Principal Component Analysis
SVD	Singular Value Decomposition
CFT	Clustering Feature Tree

# Απόδοση ξενόγλωσσων όρων

## Απόδοση

Τεχνητή Νοημοσύνη  
επιστήμη των υπολογιστών  
Μάθηση  
Μηχανική Μάθηση  
ερώτημα  
επιστήμη των δεδομένων  
Επιβλεπόμενη Μάθηση  
Μη Επιβλεπόμενη Μάθηση  
Ενισχυτική Μάθηση  
Ταξινόμηση  
Ομαδοποίηση  
Πρόβλεψη  
Ανάλυση Συσχετισμών  
δεδομένα εκπαίδευσης  
δεδομένα ελέγχου  
ανάκληση  
Επεξεργασία Φυσικής Γλώσσας  
αυτόνομο αυτοκίνητο  
Όραση Υπολογιστών  
νευρώνας  
Νευρωνικό Δίκτυο  
σώμα  
δενδρίτης

## Ξενόγλωσσος όρος

Artificial Intelligence  
computer science  
Learning  
Machine Learning  
query  
data science  
Supervised Learning  
Unsupervised Learning  
Reinforcement Learning  
Classification  
Clustering  
Prediction  
Association Analysis  
training dataset  
test dataset  
recall  
Natural Language Processing  
self-driving car  
Computer Vision  
neuron  
Neural Network  
body  
dendrite



**Απόδοση**

άξονας  
σύναψη  
Τεχνητό Νευρωνικό Δίκτυο  
τεχνητός νευρώνας  
βάρος  
στρώμα ή επίπεδο  
επίπεδο εισόδου  
επίπεδο εξόδου  
ενδιάμεσο ή κρυφό επίπεδο  
Επεξεργασία Φυσικής Γλώσσας  
Κατανόηση Φυσικής Γλώσσας  
Μηχανή Αναζήτησης  
εξυπηρετητής  
ταξινόμηση  
ταξινομητής  
ταξινόμηση κειμένου  
ετικέτα  
αθροιστής  
συνάρτηση ενεργοποίησης  
συνάρτηση περιορισμού  
συνάρτηση κατωφλίου  
συνάρτηση προσήμου  
γραμμική συνάρτηση  
μη γραμμική συνάρτηση  
σιγμοειδής συνάρτηση  
κατανομή πιθανότητας  
ανάκληση  
σφάλμα  
υποπροσαρμογή  
υπερπροσαρμογή

**Ξενόγλωσσος όρος**

axon  
synapse  
Artificial Neural Network  
artificial neuron  
weight  
layer  
input layer  
output layer  
hidden layer  
Natural Language Processing  
Natural Language Understanding  
Search Engine  
server  
classification  
classifier  
text classification  
label  
adder  
activation function  
squashing function  
threshold function  
sign function  
linear function  
non-linear function  
sigmoid function  
probability distribution  
recall  
error  
underfitting  
overfitting

**Απόδοση**

πλήρως συνδεδεμένοι  
μερικώς συνδεδεμένοι  
πρόσθια τροφοδότηση  
ανατροφοδότηση  
δέσμη  
μάθηση δέσμης  
επαυξητική μάθηση  
εποχή  
συνάρτηση κόστους  
συνάρτηση σφάλματος  
Μέση Τετραγωνική Απόκλιση  
σταυροειδή εντροπία  
κάθοδος μέγιστης κλίσης  
ρυθμός μάθησης  
Στοχαστική Κάθοδος Κλίσης  
επανάληψη  
προσαρμοστική εκτίμηση ροπής  
Απόσυρση  
Συνελικτικό Νευρωνικό Δίκτυο  
εικονοστοιχείο  
πλάτος  
ύψος  
βάθος  
ακρίβεια  
δυναδική κατηγοριοποίηση  
επίπεδο ομαδοποίησης  
Αναδρομικό Νευρωνικό Δίκτυο  
πίνακας  
πύλη επιλεκτικής συγκράτησης

**Ξενόγλωσσος όρος**

fully connected  
partially connected  
feedforward  
feedback  
batch  
batch learning  
incremental learning  
epoch  
cost function  
error function  
Mean Squared Error  
cross-entropy  
gradient descent  
learning rate  
Stochastic Gradient Descent  
iteration  
adaptive moment estimation  
Dropout  
Convolutional Neural Network  
pixel  
width  
height  
depth  
accuracy  
binary classification  
pooling layer  
Recurrent Neural Network  
vector  
forget gate

**Απόδοση**

Αλληλεπίδραση Ανθρώπου-Υπολογιστή  
διεπαφή χρήστη  
μη-σχεσιακός/ή  
βάση δεδομένων  
Ανθρωπολογία  
Τέχνη  
Κουλτούρα  
Οικονομικά  
Εκπαίδευση  
Μηχανική  
Περιβάλλον  
Φαγητό  
Ιστορία  
Ανθρωπιστικές Επιστήμες  
Νομικά  
Βιοεπιστήμες  
Γλωσσολογία  
Λογοτεχνία  
Διαχείριση  
Μαθηματικά  
Ιατρικά  
Μουσική  
Οργάνωση  
Φυσικές Επιστήμες  
Δημοφιλή Έργα  
Θρησκεία  
Κοινωνικές Επιστήμες  
Επιστήμη  
Μεταφορά  
πίνακας περιεχομένων

**Ξενόγλωσσος όρος**

Human-Computer Interaction  
user interface  
non-relational  
database  
Anthropology  
Art  
Culture  
Economics  
Education  
Engineering  
Environment  
Food  
History  
Humanities  
Law  
Life Sciences  
Linguistics  
Literature  
Management  
Mathematics  
Medicine  
Music  
Organisation  
Physical Sciences  
Popular Works  
Religion  
Social Sciences  
Science  
Transportation  
table of contents

## Απόδοση

ομοιότητα συνημιτόνων

Αποσύνθεση Ιδιαζουσών Τιμών

μέγεθος δέσμης

## Ξενόγλωσσος όρος

cosine similarity

Singular Value Decomposition

batch size