



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Πρόγραμμα Προπτυχιακών Σπουδών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Τεχνολογίες μηχανικής μάθησης και όρασης υπολογιστών για
την βελτιστοποίηση της απόδοσης στο πεδίο του αθλητισμού**

**Ιωάννης Καμπεράκης
Α.Μ. 71347254**

Εισηγητής: Δρ Παναγιώτης Καρκαζής, Επ. Καθηγητής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Τεχνολογίες μηχανικής μάθησης και όρασης υπολογιστών για την
βελτιστοποίηση της απόδοσης στον πεδίο του αθλητισμού**

**Ιωάννης Καμπεράκης
Α.Μ. 71347254**

Εισηγητής:

Δρ Παναγιώτης Καρκαζής, Επ. Καθηγητής

Εξεταστική Επιτροπή:

A/α	ΟΝΟΜ/ΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ	ΥΠΟΓΡΑΦΗ
1	Καρκαζής Παναγιώτης	Επ. Καθ. Τμ. Μηχ. Πληροφορικής και Υπολογιστών	
2	Νικόλαος Μυριδάκης	Επ. Καθ. Τμ. Μηχ. Πληροφορικής και Υπολογιστών	
3	Μάμαλης Βασίλειος	Καθ. Τμ. Μηχ. Πληροφορικής και Υπολογιστών	

Ημερομηνία εξέτασης 11/03/2022

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Ιωάννης Καμπεράκης του Φιλίππου, με αριθμό μητρώου 71347254, φοιτητής/τρια του Προγράμματος Προπτυχιακών Σπουδών του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών
Ιωάννης Καμπεράκης



ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της διπλωματικής μου εργασίας, θα ήθελα να ευχαριστήσω όλους τους καθηγητές τους οποίους συνάντησα κατά την διάρκεια των σπουδών μου. Με βοήθησαν να γίνω αυτό που είμαι σήμερα δίνοντας μου τις κατάλληλες ωθήσεις ώστε να εντοπίσω τα λάθη μου και να εξελιχθώ από αυτά.

Θέλω να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Παναγιώτη Καρκαζή, για τις γνώσεις που μου παρείχε τόσα χρόνια κατά την διάρκεια των διαλέξεων του και για την επίβλεψη και την υποστήριξη της εργασίας αυτής. Επίσης, θέλω να ευχαριστήσω θερμά και τον κ. Αποστόλη Αναγνωστόπουλο για την τόσο μεγάλη όρεξη που είχε να μεταλαμπαδεύσει τις γνώσεις που διαθέτει.

Τέλος θέλω να ευχαριστήσω την οικογένεια μου, η οποία μου παρείχε τα πάντα ώστε να βγάλω εις πέρας τις σπουδές μου. Με στήριζε και μου έδινε δύναμη στις δύσκολες στιγμές των σπουδών μου. Με βοηθούσε και με ενθάρρυνε κάθε στιγμή ώστε να τα βγάλω πέρα και με την βοήθεια τους τα κατάφερα. Χωρίς αυτούς δεν θα έφτανα έως εδώ.

ΠΕΡΙΛΗΨΗ

Ο όρος της Τεχνητής Νοημοσύνης έχει γίνει γνωστός σε όλο τον κόσμο. Ακούμε συνέχεια για τα έξυπνα κινητά, τα έξυπνα σπίτια, τις έξυπνες τηλεοράσεις και γενικά από συσκευές και υπολογιστικά συστήματα που παρέχουν ευφυείς υπηρεσίες. Δεν γίνεται πλήρως αντιληπτό όμως ότι ο τίτλος «έξυπνος» που έχει δοθεί πλέον στις συσκευές, οφείλεται κυρίως στην ανάπτυξη της Τεχνητής Νοημοσύνης και την ενσωμάτωσή της στην καθημερινότητα του ανθρώπου.

Οι δυνατότητες της Τεχνητής Νοημοσύνης και οι τομείς που μπορεί να χρησιμοποιηθεί αυξάνονται συνεχώς. Βρισκόμαστε στην εποχή όπου οι επιστήμονες συνειδητοποιούν τις προοπτικές και την διευκόλυνση που μπορούν να μας παρέχουν οι έξυπνες συσκευές, γι' αυτό και έχει ξεκινήσει η εντατική παραγωγή ευφύων υπολογιστικών συστημάτων και η δημιουργία λογισμικών που μαθαίνουν από τα δεδομένα που παράγονται καθημερινά.

Στο πλαίσιο αυτής της διπλωματικής, δημιουργείται ένα μοντέλο με την χρήση αλγορίθμων Μηχανικής Μάθησης και Όρασης Υπολογιστών, το οποίο είναι εκπαιδευμένο με βάση ένα σύνολο από βίντεο ελεύθερων βολών ενός ανθρώπου Α και ένας άνθρωπος Β τροφοδοτεί το βίντεο του στο μοντέλο για να δει κατά πόσο είναι σωστή η τεχνική της ελεύθερης βολής του σύμφωνα με την τεχνική του Α.

Λέξεις κλειδιά: Μηχανική Μάθηση (Machine Learning - ML), Όραση Υπολογιστών (Computer Vision - CV), Τεχνητή Νοημοσύνη (Artificial Intelligence - AI), Νευρωνικά Δίκτυα (Neural Networks - NN), Βαθιά Μάθηση (Deep Learning - DL), Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing - NLP), Γνωστική Υπολογιστική (Cognitive Computing - CC), Δεδομένα, Βελτιστοποίηση Αθλητικής Απόδοσης, Εκπαίδευση, Ευφυή, Μοντέλο, Έξυπνες Συσκευές, Ανάλυση Δεδομένων

ABSTRACT

The term Artificial Intelligence has become known throughout the world. We hear constantly about smart phones, smart homes, smart TVs and generally from devices and computer systems that provide intelligent services. However, it is not fully understood that the title "smart" that has now been given to devices, is mainly due to the development of Artificial Intelligence and its integration into everyday life.

The capabilities of Artificial Intelligence and the areas that can be used are constantly increasing. We are in a time where scientists are realizing the prospects and convenience that smart devices can provide us, which is why intensive production of intelligent computing systems and the creation of software that learns from the data produced daily has begun.

As part of this dissertation, a model is created using Machine Learning and Computer Vision algorithms, which is trained based on a set of free-shooting videos of a person A and a person B which feeds his video to the model to see if it is the technique of his free throw is correct according to the technique of A.

Keywords: Machine Learning (ML), Computer Vision (CV), Artificial Intelligence (AI), Neural Networks (NN), Deep Learning (DL), Natural Language Processing (NLP), Cognitive Computing (CC), Data, Athletic Performance Optimization, Training, Intelligent, Model, Smart Devices, Data Analysis

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ	8
ABSTRACT	9
ΕΙΣΑΓΩΓΗ	13
ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ	15
2.1 Ορισμός της Τεχνητής Νοημοσύνης	15
2.2 Διαχωρισμός της Τεχνητής Νοημοσύνης σε υποκατηγορίες	16
2.2.1 Νευρωνικά Δίκτυα	18
2.2.2 Βαθιά Μάθηση	23
2.2.3 Επεξεργασία Φυσικής Γλώσσας	29
2.2.4 Γνωστική Υπολογιστική	32
2.2.5 Όραση Υπολογιστών	34
ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ	37
3.1 Τι είναι η Μηχανική Μάθηση	37
3.2 Μέθοδοι Μηχανικής Μάθησης	42
ΕΦΑΡΜΟΓΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΑΘΛΗΤΙΚΩΝ ΕΠΙΔΩΣΕΩΝ	48
4.1 Παρουσίαση εφαρμογής	48
4.2 Βιβλιοθήκες και εξαρτήματα που χρησιμοποιήθηκαν	49
4.2.1 Βιβλιοθήκη MediaPipe	50
4.2.2 Βιβλιοθήκη scikit-learn	50
4.2.3 Βιβλιοθήκη pandas	51
4.2.4 Βιβλιοθήκη OpenCV	51
4.2.4 Βιβλιοθήκη matplotlib	52
4.2.5 Η Βιβλιοθήκη xlwt και το module os	52
4.2.6 Εξοπλισμός και μέσα που χρησιμοποιήθηκαν	53
4.3 Διαδικασία υλοποίησης της εφαρμογής	53
4.3.1 Συλλογή των δεδομένων	54
4.3.2 Προετοιμασία των δεδομένων	56
4.3.3 Επιλογή αλγορίθμου	57
4.3.4 Εκπαίδευση του μοντέλου	57
4.3.5 Αξιολόγηση του μοντέλου	58
4.3.6 Ρύθμιση απόδοσης	59
4.3.7 Παραγωγή αποτελεσμάτων	60
ΣΥΜΠΕΡΑΣΜΑΤΑ	70
ΒΙΒΛΙΟΓΡΑΦΙΑ	72

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Εικόνα 1: Βασικές υποκατηγορίες της Τεχνητής Νοημοσύνης	17
Εικόνα 2: Αναπαράσταση ενός βιολογικού νευρώνα	20
Εικόνα 3: Σχηματική αναπαράσταση του τεχνητού νευρώνα/perceptron	21
Εικόνα 4: Αναπαράσταση ενός Deep Neural Network	26
Εικόνα 5: CNN: Διαδικασία εξαγωγής χαρακτηριστικών	27
Εικόνα 6: Σχηματική αναπαράσταση ενός Ανατροφοδοτούμενου Νευρωνικού Δικτύου ..	28
Εικόνα 7: Η κίνηση του διαδικτύου τα τελευταία 30 χρόνια	39
Εικόνα 8: Εντοπισμός της στάσης του ανθρώπινου σώματος.....	49
Εικόνα 9: Λογότυπο του MediaPipe.....	50
Εικόνα 10: Λογότυπο του scikit-learn	51
Εικόνα 11: Λογότυπο του pandas	51
Εικόνα 12: Λογότυπο του OpenCV	52
Εικόνα 13: Λογότυπο της βιβλιοθήκης matplotlib	52
Εικόνα 14: Δημιουργία των Excel όπου θα εκχωρηθούν οι συντεταγμένες.....	55
Εικόνα 15: Μεταβλητές που περιέχουν το path των βίντεο	55
Εικόνα 16: Συναρτήσεις για διαχείριση ενός και πολλών βίντεο	55
Εικόνα 17: Αποθήκευση των συντεταγμένων στα Excel	56
Εικόνα 18: Έτοιμες συναρτήσεις της βιβλιοθήκης MediaPipe για το pose detection	56
Εικόνα 19: Συνάρτηση που εφαρμόζει τον k-means	58
Εικόνα 20: Συνάρτηση εκτύπωσης των συντεταγμένων.....	58
Εικόνα 21: Υπολογισμός του Silhouette score	59
Εικόνα 22: Παράδειγμα αποτελεσμάτων της μεθόδου Elbow.....	60
Εικόνα 23: Υλοποίηση της μεθόδου Elbow.....	60
Εικόνα 24: Συντεταγμένες του ώμου.....	61
Εικόνα 25: Συντεταγμένες του αγκώνα.....	61
Εικόνα 26: Συντεταγμένες του καρπού.....	62
Εικόνα 27: Διάγραμμα με τα αποτελέσματα του Silhouette score για τον ώμο.....	63
Εικόνα 28: Αποτελέσματα του Silhouette score για τον ώμο.....	63
Εικόνα 29: Αποτελέσματα της μεθόδου Elbow για τον ώμο.....	64
Εικόνα 30: Αποτελέσματα του Silhouette score σε διάγραμμα για τον αγκώνα	64
Εικόνα 31: Αποτελέσματα του Silhouette score για τον αγκώνα.....	65
Εικόνα 32: Αποτελέσματα της μεθόδου Elbow για τον αγκώνα.....	65
Εικόνα 33: Αποτελέσματα του Silhouette score σε διάγραμμα για τον καρπό.....	66
Εικόνα 34: Αποτελέσματα του Silhouette score για τον καρπό	66
Εικόνα 35: Αποτελέσματα της μεθόδου Elbow για τον καρπό.....	67
Εικόνα 36: Αποτελέσματα του k-means για τον ώμο	68
Εικόνα 37: Αποτελέσματα του k-means για τον αγκώνα	68
Εικόνα 38: Αποτελέσματα του k-means για τον καρπό	69

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

AI	Artificial Intelligence
AN	Artificial Neuron
ANN	Artificial Neural Network
BPNN	Back-Propagation Neural Network
CC	Cognitive Computing
CNN	Convolutional Neural Network
CV	Computer Vision
DL	Deep Learning
DNN	Deep Neural Network
DS	Dialog Systems
FFNN	Feed-Forward Neural Network
GM	Generative Models
IE	Information Extraction
IR	Information Retrieval
LSTM	Long Short-Term Memory
ML	Machine Learning
MT	Machine Translation
NLP	Natural Language Processing
NN	Neural Networks
PCA	Principal Component Analysis
RNN	Recurrent Neural Network
SVM	Support Vector Machine
TSVM	Transductive Support Vector Machine
WCSS	Within-Cluster Sum of Square

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

Στις μέρες μας η τεχνολογία εξελίσσεται με ταχύτατους ρυθμούς δίνοντας την ευκαιρία στην αγορά να παράγει νέα προϊόντα, εξελιγμένα, με νέες λειτουργίες και δυνατότητες. Πολλές νέες συσκευές, με καινούργιες και έξυπνες λειτουργίες, παράγονται σε σύντομο χρονικό διάστημα. Υπάρχει μεγάλος ανταγωνισμός στην αγορά, καθώς όλες οι εταιρείες που παράγουν τέτοιες συσκευές προσπαθούν καθημερινά να ανακαλύψουν νέους τρόπους για να γίνουν τα προϊόντα τους πιο χρήσιμα. Οι εφαρμογές των κινητών είχαν φτάσει το όριο σε θέμα λειτουργιών, το μόνο που μπορούσε να αλλάξει ήταν η πολυπλοκότητα των λειτουργιών και ο σχεδιασμός της εφαρμογής. Για να υπάρξει σημαντική αλλαγή, έπρεπε να δημιουργηθεί κάτι καινούργιο που θα άλλαζε τα δεδομένα.

Με την χρήση της Τεχνητής Νοημοσύνης καταφέραμε να προσθέσουμε νέες λειτουργίες, οι οποίες αξιοποιούν το μεγάλο και χαοτικό πλήθος δεδομένων, που παράγονται καθημερινά από διάφορες πηγές δίνοντας τους νόημα. Αυτή η δυνατότητα, όπως θα δούμε και στην συνέχεια είναι εξαιρετικά χρήσιμη καθώς ο άνθρωπος, δεν μπορεί πλέον να διαχειριστεί το πλήθος των δεδομένων που παράγονται. Οι ρυθμοί δημιουργίας των δεδομένων έχει αυξηθεί τόσο, όπου απαιτείται η χρήση της Τεχνητής Νοημοσύνης. Ανάλογα την χρήση, τα δεδομένα αναλύονται και παράγονται μοντέλα που προβλέπουν καταστάσεις και μας δείχνουν συσχετίσεις μεταξύ των δεδομένων. Για να επιτευχθούν αυτά τα αποτελέσματα χρησιμοποιούνται μερικές από τις υποκατηγορίες της Τεχνητής Νοημοσύνης όπως η Μηχανική Μάθηση, η Όραση Υπολογιστών, η Επεξεργασία Φυσικής Γλώσσας, τα Νευρωνικά Δίκτυα, η Γνωστική Υπολογιστική και η Βαθιά Μάθηση. Ανάλογα την περίπτωση χρησιμοποιείται και η κατάλληλη κατηγορία ή συνδυασμός αυτών.

Σκοπός της συγκεκριμένης διπλωματικής είναι η παρουσίαση των παραπάνω υποκατηγοριών και η επίδειξη των δυνατοτήτων της Τεχνητής Νοημοσύνης μέσω της δημιουργίας μίας εφαρμογής. Πρόκειται για μία εφαρμογή βελτιστοποίησης αθλητικής απόδοσης όπου με την χρήση πραγματικών αθλητών θα μπορούσε να χρησιμοποιηθεί για να βελτιώσει τις επιδόσεις των αθλητών. Στην ουσία η

εφαρμογή θα χρησιμοποιεί ένα μοντέλο Μηχανικής Μάθησης το οποίο θα έχει εκπαιδευτεί με την χρήση βίντεο. Τα βίντεο θα περιέχουν την εκτέλεση ελεύθερης βολής μπάσκετ, η ελεύθερη βολή θα εκτελείται από τον συγγραφέα αυτής της διπλωματικής. Αφού εκπαιδευθεί το μοντέλο, θα δέχεται βίντεο άλλων ανθρώπων που θα ρίχνουν και αυτοί ελεύθερη βολή, ώστε να παραχθεί ένα ποσοστό ομοιότητας. Αυτό το ποσοστό ομοιότητας σημαίνει κατά πόσο είναι «σωστή» η στάση σώματος του ανθρώπου που ρίχνει στο βίντεο την βολή σύμφωνα πάντα με την στάση σώματος του συγγραφέα της διπλωματικής. Έπειτα θα παρουσιαστούν τα αποτελέσματα, οι παρατηρήσεις και πως η Τεχνητή Νοημοσύνη θα διαμορφώσει το μέλλον της ανθρωπότητας.

Στα κεφάλαια που ακολουθούν θα παρουσιαστούν τα ακόλουθα:

- Ο ορισμός, τα χαρακτηριστικά, οι δυνατότητες, η χρησιμότητα και οι υποκατηγορίες στις οποίες χωρίζεται η Τεχνητή Νοημοσύνη (Κεφάλαιο 2, Τεχνητή Νοημοσύνη)
- Η έννοια και η χρησιμότητα της Μηχανικής Μάθησης, τα βήματα που απαιτούνται για την εκπαίδευση ενός μοντέλου και οι τρόποι με τους οποίους μπορεί να εκπαιδευθεί ένα μοντέλο (Κεφάλαιο 3, Μηχανική Μάθηση)
- Η ανάπτυξη, ο σχεδιασμός, οι τεχνολογίες που χρησιμοποιήθηκαν και τα αποτελέσματα που παράχθηκαν από το μοντέλο που δημιουργήθηκε (Κεφάλαιο 4, Υλοποίηση Λογισμικού)
- Τα συμπεράσματα, οι παρατηρήσεις, οι προσδοκίες, οι αλλαγές και οι πιθανότητες για ένα πολλά υποσχόμενο μέλλον που μπορεί να μας προσφέρει η Τεχνητή Νοημοσύνη (Κεφάλαιο 5, Συμπεράσματα)

ΚΕΦΑΛΑΙΟ 2

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

2.1 Ορισμός της Τεχνητής Νοημοσύνης

Ο όρος Τεχνητή Νοημοσύνη αρχικά συστήθηκε από τους φιλόσοφους οι οποίοι προσπάθησαν να σκεφτούν, να οραματιστούν και να περιγράψουν την διαδικασία της σκέψης του ανθρώπου ενσωματωμένη σε μία μηχανή [1]. Η συγκεκριμένη σκέψη κέντρισε το ενδιαφέρον των επιστημόνων του 20^{ου} αιώνα. Αυτό είχε ως αποτέλεσμα να συνεργαστούν επιστήμονες από διάφορες κατευθύνσεις (μαθηματικοί, ψυχολόγοι, μηχανικοί κλπ.) ώστε να διερευνήσουν και να προσπαθήσουν να υλοποιήσουν αυτή την ιδέα [1] [2]. Κάπως έτσι ξεκίνησε η ανάπτυξη της Τεχνητής Νοημοσύνης.

Η Τεχνητή Νοημοσύνη ασχολείται με την αναπαραγωγή των νοητικών ικανοτήτων του ανθρώπου στις μηχανές. Ανάλογα με το πως αυτή θα προσεγγιστεί, δηλαδή είτε με βάση την τεχνική που θα χρησιμοποιηθεί, είτε με βάση τον σκοπό που έχουμε, είτε με βάση την εφαρμογή ή είτε με βάση τις λειτουργίες που θέλουμε, ο όρος της Τεχνητής Νοημοσύνης μπορεί να αλλάξει και να γενικευθεί [3] [4]. Στην ουσία όμως είναι ένα σύνολο τεχνολογιών και τεχνικών που συμπληρώνουν τις νοητικές λειτουργίες των ανθρώπων. Μερικές από τις βασικές και σημαντικές νοητικές λειτουργίες του ανθρώπου είναι η δυνατότητα αναλυτικής σκέψης, η επίλυση προβλημάτων, η αυτοδιόρθωση, η ικανότητα λογικής και δημιουργικής σκέψης και η αντίληψη [5] [6] [7]. Γενικά η υλοποίηση της αντίληψης παίζει ένα μεγάλο ρόλο για την ανάπτυξη ενός προηγμένου ευφυούς συστήματος και για την εφαρμογή μερικών δυνατοτήτων της Τεχνητής Νοημοσύνης όπως η όραση υπολογιστών και η ομιλία. Ο άνθρωπος αντιλαμβάνεται και ερμηνεύει μέσω των αισθήσεων του τα διάφορα ερεθίσματα που δέχεται από το περιβάλλον. Αυτό πυροδοτεί την διαδικασία της σκέψης, φτάνοντας κάποια στιγμή στην λήψη απόφασης και αντίδρασης. Με βάση τα παραπάνω, ένα ευφυές σύστημα χρειάζεται οπτικά και ακουστικά αισθητήρια μέσα, σε συνδυασμό με έναν νου ο οποίος θα δέχεται και θα επεξεργάζεται τα σήματα των αισθητήριων μέσων. Ο απώτερος σκοπός είναι μέσω της ανάλυσης των δεδομένων που δέχεται, (είτε από το

περιβάλλον είτε από έναν προγραμματιστή) μία μηχανή να μαθαίνει, να εξελίσσεται και να προσαρμόζεται σε διάφορες καταστάσεις [4].

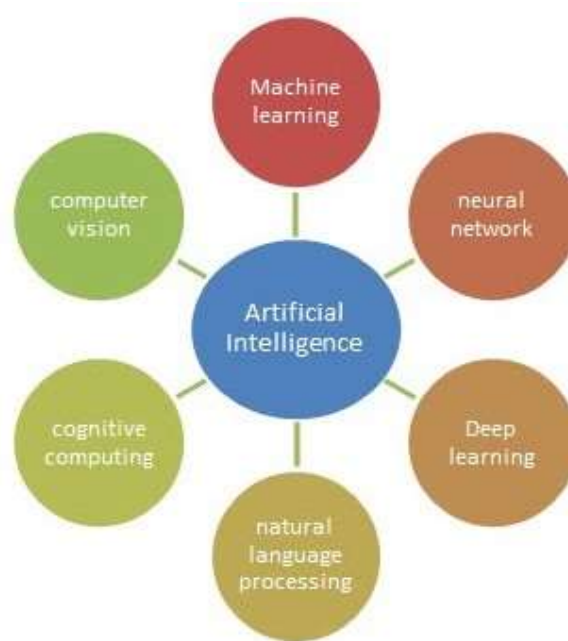
Σύμφωνα με τα παραπάνω και χρησιμοποιώντας την φαντασία μας, γίνονται αντιληπτές οι πιθανότητες εφαρμογής της Τεχνητής Νοημοσύνης. Αρχικά με την χρήση της, οι άνθρωποι θα έχουν περισσότερο ελεύθερο χρόνο, διότι οι δουλειές που χρειάζονται συγκεκριμένα βήματα, μπορούν να υλοποιηθούν εύκολα και γρήγορα με την χρήση της Τεχνητής Νοημοσύνης. Έτσι ο άνθρωπος απελευθερώνεται από αρκετές δουλειές, οι οποίες χρειάζονται επαναλαμβανόμενα και συγκεκριμένα βήματα, καθώς εξοικονομείται πολύτιμος χρόνος και ενέργεια [5]. Εκτός από την εξοικονόμηση χρόνου και ενέργειας προσφέρει επίσης και ασφάλεια, ιδιαίτερα στα εργοστάσια και τις βιομηχανίες. Αυτό συμβαίνει διότι δεν υπάρχει η ανθρώπινη επέμβαση, οπότε δεν υπάρχει περίπτωση τραυματισμού και λάθους. Ένα εύστοχο παράδειγμα της τεχνητής νοημοσύνης στην καθημερινότητα μας είναι ο φωνητικός βοηθός (voice assistant). Οι πιο διαδεδομένοι φωνητικοί βοηθοί είναι η Cortana της Microsoft, η Siri της Apple και η Alexa της Amazon [5]. Πρόκειται για μία έξυπνη λειτουργία που έχει προστεθεί σε πολλές συσκευές και παρέχει ακόμα ένα τρόπο αλληλεπίδρασης των ανθρώπων με τις συσκευές. Μερικές ακόμη από τις δυνατότητες της Τεχνητής Νοημοσύνης είναι η κατανόηση της ανθρώπινης ομιλίας, η κατανόηση διαφόρων γλωσσών, η δυνατότητα ομιλίας, η ρομποτική, η αναγνώριση ανθρώπων και άλλων αντικειμένων μέσω φωτογραφιών, η εξαγωγή συμπερασμάτων από ένα μεγάλο σύνολο δεδομένων κλπ [3]. Η λίστα των δυνατοτήτων και των εφαρμογών της Τεχνητής Νοημοσύνης συνεχίζει να αυξάνεται και να εμπλουτίζεται όσο περνάει ο καιρός καθώς προκύπτουν νέες ανάγκες και ιδέες.

2.2 Διαχωρισμός της Τεχνητής Νοημοσύνης σε υποκατηγορίες

Στο προηγούμενο υποκεφάλαιο έγινε μία σύντομη παρουσίαση της Τεχνητής Νοημοσύνης. Δόθηκε ο ορισμός της, εξετάστηκαν ορισμένες σύγχρονες εφαρμογές της και συζητήθηκαν μερικά από τα οφέλη της. Έχοντας κατά νου τα παραπάνω προκύπτουν νέες απορίες για αυτό τον τομέα της Επιστήμης των Υπολογιστών, όπως «Όλες αυτές οι εφαρμογές ανήκουν στην Τεχνητή Νοημοσύνη;», «Όλες αυτές οι λειτουργίες που είδαμε στο ορισμό της Τεχνητής Νοημοσύνης πως και πότε υλοποιούνται;», «Τι σχέση έχει η Μηχανική Μάθηση με την Τεχνητή Νοημοσύνη;».

Σε αυτό το υποκεφάλαιο, θα απαντηθούν οι παραπάνω ερωτήσεις καθώς και πολλές άλλες που πιθανόν υπάρχουν ή έχουν προκύψει.

Έχοντας τις παραπάνω απορίες, κάποιος ίσως αναζητούσε αν υπάρχουν υποκατηγορίες της Τεχνητής Νοημοσύνης. Η τεχνητή νοημοσύνη αποτελείται από διάφορες προσεγγίσεις οι οποίες αποτυπώνονται στην παρακάτω σχήμα:



Εικόνα 1: Βασικές υποκατηγορίες της Τεχνητής Νοημοσύνης [8]

Παρατηρώντας την **Εικόνα 1** αντιλαμβανόμαστε ότι η Τεχνητή Νοημοσύνη βρίσκεται στο κέντρο του σχήματος και όλες οι υποκατηγορίες συνδέονται μαζί της. Όπως είδαμε η Τεχνητή Νοημοσύνη είναι ένας τομέας ο οποίος έχει ως στόχο την παραγωγή έξυπνων συστημάτων. Μία από τις λειτουργίες που μπορεί να έχει ένα έξυπνο σύστημα είναι η κατανόηση της γλώσσας των ανθρώπων, όπου ένα σύστημα μπορεί να προγραμματιστεί να κατανοεί πολλές γλώσσες. Αυτό μπορεί να γίνει με την χρήση μίας από τις παραπάνω υποκατηγορίες και συγκεκριμένα με την Επεξεργασία Φυσικής Γλώσσας. Με βάση αυτό το παράδειγμα καταλαβαίνουμε ότι η Τεχνητή Νοημοσύνη βρίσκεται στο κέντρο γιατί παρέχει τον σκοπό ενώ οι υποκατηγορίες είναι τα μέσα με τα οποία θα επιτευχθεί αυτός ο σκοπός. Όλοι οι τομείς εξειδικεύονται και υλοποιούν διαφορετικές λειτουργίες, όμως κάθε τομέας έχει ως σκοπό να κάνει τα υπολογιστικά συστήματά ευφυή με τον δικό του τρόπο. Με τον συνδυασμό όλων των τομέων, δημιουργείται ένα υπολογιστικό σύστημα

πραγματικά ευφυές το οποίο μπορεί να μιλήσει, να καταλάβει, να σκεφτεί, να βλέπει, να αντιλαμβάνεται και πολλά άλλα που θα παρουσιαστούν στην συνέχεια.

Ουσιαστικά η Τεχνητή Νοημοσύνη αποτελείται από αυτές τις έξι βασικές υποκατηγορίες οι οποίες είναι οι εξής [8]:

- Μηχανική Μάθηση (Machine Learning)
- Νευρωνικά Δίκτυα (Neural Networks)
- Βαθιά Μάθηση (Deep Learning)
- Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing - NLP)
- Γνωστική Υπολογιστική (Cognitive Computing)
- Όραση Υπολογιστών (Computer Vision)

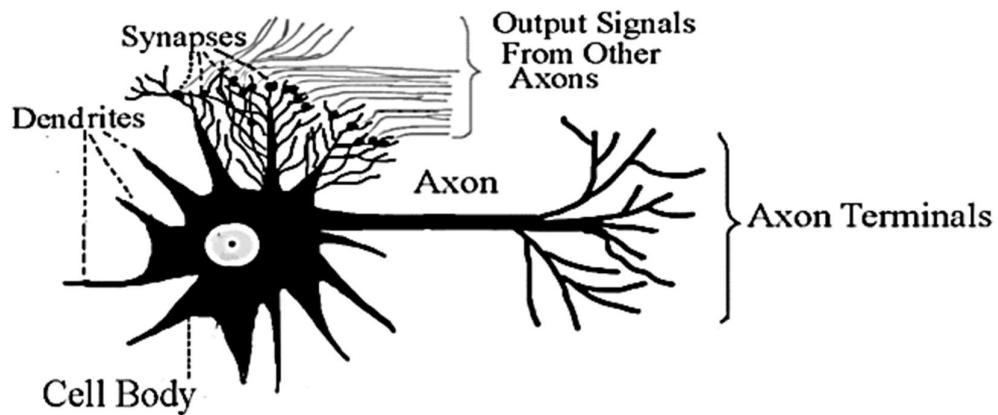
Κάθε μία από αυτές τις κατηγορίες βοηθά με διαφορετικό τρόπο στην μίμηση ορισμένων διανοητικών ικανοτήτων του ανθρώπινου εγκεφάλου. Κάθε μία έχει τα δικά της χαρακτηριστικά καθώς και τις δικές της τεχνικές και λειτουργίες. Αξίζει να παρουσιαστούν και να σχολιαστούν όλες οι υποκατηγορίες για να κατανοηθούν καλύτερα οι δυνατότητες της Τεχνητής Νοημοσύνης στο σύνολο της. Παρακάτω θα οριστούν οι κατηγορίες κατανοώντας τον σκοπό της κάθε μίας, θα δούμε τις δυνατότητες κάθε υποκατηγορίας, με ποιες τεχνικές επιτυγχάνει τον σκοπό της η κάθε κατηγορία καθώς και μερικούς τομείς που εφαρμόζονται. Θα εξαιρεθεί η Μηχανική Μάθηση καθώς το επόμενο κεφάλαιο είναι αφιερωμένο στην συγκεκριμένη υποκατηγορία.

2.2.1 Νευρωνικά Δίκτυα

Η πρώτη υποκατηγορία που πρόκειται να αναλυθεί φέρει το όνομα Νευρωνικά Δίκτυα ή αλλιώς Neural Networks. Κατά την ανάπτυξη της Τεχνητής Νοημοσύνης αρκετοί επιστήμονες αναρωτήθηκαν και ασχολήθηκαν με το εξής ερώτημα, «Πώς θα μπορεί να σκέφτεται ένα υπολογιστικό σύστημα;» [1]. Αυτό το ερώτημα ώθησε τους επιστήμονες στην ανακάλυψη των Νευρωνικών Δικτύων. Τα Νευρωνικά Δίκτυα αντιγράφουν τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου [9]. Αυτός ο τομέας της Τεχνητής Νοημοσύνης μπορεί να χαρακτηριστεί και ως «Το μυαλό της Τεχνητής Νοημοσύνης» αφού χρησιμοποιεί την νευροεπιστήμη, ένας κλάδος της βιολογίας όπου ασχολείται με την επιστημονική μελέτη του νευρικού συστήματος του ανθρώπου [8] [10] [11]. Συγκεκριμένα, στα Νευρωνικά Δίκτυα είναι πιο σημαντική η λειτουργία του εγκεφάλου, γι' αυτό και ασχολούνται πιο πολύ με τα νεύρα και την δομή του εγκεφάλου. Ο σκοπός των Νευρωνικών Δικτύων είναι ο

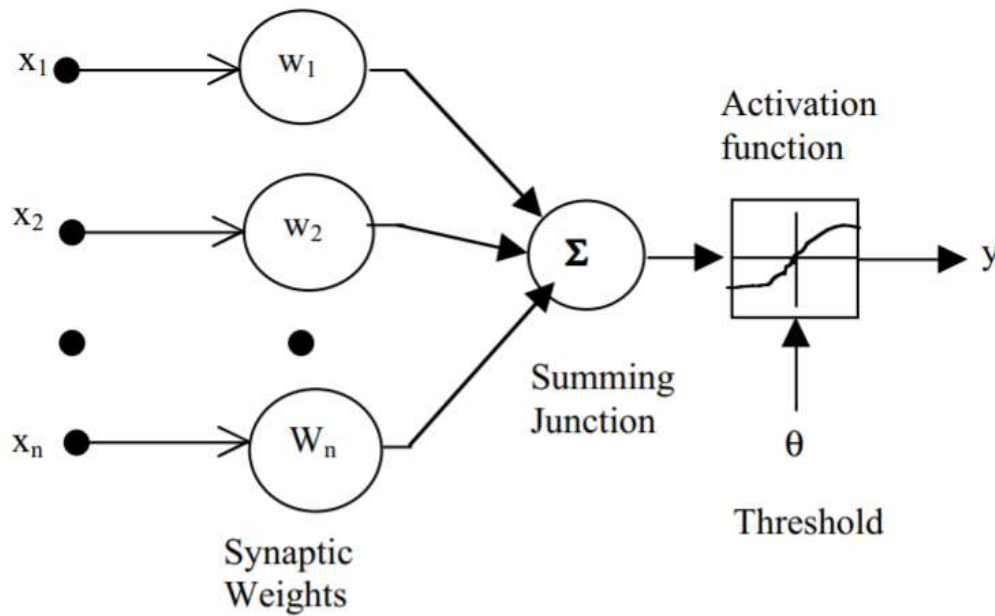
προγραμματισμός τεχνητών νευρώνων (artificial neurons) στα υπολογιστικά συστήματα ώστε ένα σύστημα να μπορεί να λύσει προβλήματα όπως ο άνθρωπος [12]. Δίνοντας του ένα πλήθος δεδομένων, το Νευρωνικό Δίκτυο εντοπίζει τις συσχετίσεις που μπορεί να υπάρχουν μεταξύ των δεδομένων [13]. Με την παραγωγή των τεχνητών νευρώνων και των νευρικών συνάψεων, καταφέρνουμε να μιμηθούμε σε έναν βαθμό μερικές από τις δυνατότητες του ανθρώπινου εγκεφάλου.

Για να αναπτυχθεί ο πρώτος τεχνητός νευρώνας έπρεπε να βασιστούν οι επιστήμονες στην δομή και τις λειτουργίες των βιολογικών νευρώνων [14] [15]. Ένας βιολογικός νευρώνας αποτελείται από το σώμα (Soma), τους δενδρίτες και τον άξονα (axon). Το σώμα είναι το κύριο μέρος του νευρώνα στον οποίο βρίσκεται και ο πυρήνας του. Ο άξονας έχει σχήμα κυλίνδρου το οποίο επεκτείνεται από το σώμα του νευρώνα και οι δενδρίτες, οι οποίοι έχουν δεντρική μορφή, είναι υπεύθυνοι για την λήψη των σημάτων από τους άλλους νευρώνες μέσω της σύνδεσης που έχουν με ένα τερματικό κουμπί το οποίο είναι το τερματικό σημείο ενός άξονα. Ο άξονας χωρίζεται σε αρκετά κλαδιά όπου το καθένα καταλήγει σε ένα τερματικό κουμπί που ουσιαστικά είναι το σημείο όπου πραγματοποιείται η σύναψη [14]. Το μικρό κενό που μεσολαβεί μεταξύ του τερματικού κουμπιού και του δενδρίτη ονομάζεται σύναψη και είναι το μέσο με το οποίο οι νευρώνες μπορούν να στείλουν σήματα στους υπόλοιπους νευρώνες, δηλαδή είναι ο τρόπος επικοινωνίας μεταξύ τους. Ένας νευρώνας μπορεί να έχει πολλές συναπτικές συνδέσεις. Τα εισερχόμενα σήματα τα οποία λαμβάνει ένας νευρώνας, αθροίζονται στο σώμα του νευρώνα και αν έχουν ληφθεί αρκετά σήματα δηλαδή αν ξεπερνούν ένα συγκεκριμένο κατώφλι, τότε διεγείρεται ο νευρώνας. [14] [15]



Εικόνα 2: Αναπαράσταση ενός βιολογικού νευρώνα [15]

Ο τεχνητός νευρώνας ή αλλιώς perceptron [8] [16], στην πραγματικότητα είναι το απλούστερο Νευρωνικό Δίκτυο που υπάρχει περιέχοντας μόνο τρία επίπεδα, το επίπεδο των εισόδων στο οποίο τα δεδομένα εισέρχονται στο Νευρωνικό Δίκτυο, το κρυφό επίπεδο (hidden layer) όπου εκεί πραγματοποιείται η επεξεργασία των δεδομένων και το επίπεδο των εξόδων όπου αποφασίζεται από το σύστημα η έξοδος δεδομένου των εισόδων [17]. Τα βασικά στοιχεία από τα οποία αποτελείται ένας τεχνητός νευρώνας είναι τα εξής: οι είσοδοι, τα βάρη, το βεβαρημένο άθροισμα, το κατώφλι και την συνάρτηση ενεργοποίησης [8] [16] [18].



Εικόνα 3: Σχηματική αναπαράσταση του τεχνητού νευρώνα/perceptron [15]

Αρχικά στον τεχνητό νευρώνα εισέρχονται μέσω των εισόδων τα δεδομένα τα οποία θέλουμε ο τεχνητός νευρώνας να επεξεργαστεί. Στην συνέχεια τα δεδομένα πολλαπλασιάζονται με τα βάρη των αντίστοιχων συνδέσεων. Αυτά τα βάρη στην αρχή ορίζονται αυθαίρετα. Οι συνδέσεις και τα βάρη είναι οι πιο σημαντικές παράμετροι σε ένα μοντέλο γιατί με αυτά αλλάζει η συμπεριφορά του μοντέλου, δηλαδή τα αποτελέσματα που παράγονται. Ένα μεγάλο συναπτικό βάρος μπορεί να επηρεάσει σημαντικά το αποτέλεσμα, γι' αυτό πρέπει να δίνεται μεγάλη προσοχή στην επιλογή των βαρών, διότι ανάλογα με την περίπτωση θα χρειαστούν διαφορετικά βάρη [19]. Έπειτα, υπολογίζεται το άθροισμα των βαρών και των εισόδων ή αλλιώς το βεβαρημένο άθροισμα και το αποτέλεσμα του αθροίσματος τροφοδοτείται σε μία συνάρτηση η οποία παράγει την έξοδο. Υπάρχει όμως και το κατώφλι (threshold) το οποίο είναι μία τιμή που πρέπει να ξεπεραστεί ώστε να ενεργοποιηθεί η έξοδος του τεχνητός νευρώνας. Υπάρχουν πολλών ειδών συναρτήσεις οι οποίες χρησιμοποιούνται γι' αυτό το σκοπό. Μερικές από αυτές είναι οι εξής [15]:

- Γραμμική συνάρτηση (Linear function)
- Σιγμοειδής συνάρτηση (Sigmoid function)
- Συνάρτηση κατωφλίου (Threshold Function)

Μόλις παρουσιάστηκε το απλούστερο Νευρωνικό Δίκτυο το οποίο αποκαλείται και Νευρωνικό Δίκτυο ενός επιπέδου (single layer neural network). Λόγω της απλής αρχιτεκτονικής του, το Νευρωνικό Δίκτυο ενός επιπέδου δεν είναι φτιαγμένο για να επιλύει πολύ σύνθετα προβλήματα, συνήθως χρησιμοποιείται για την υλοποίηση λογικών πυλών [20] [21]. Λόγω της απλότητας του, μπορεί να διαχειριστεί και να μάθει μόνο από γραμμικά διαχωρίσιμα δεδομένα και αυτός είναι και ένας από τους περιορισμούς του [20] [22]. Όμως χρησιμεύουν στην υλοποίηση των Πολυστρωματικών Νευρωνικών Δικτύων (multilayer neural networks).

Τα Πολυστρωματικά Νευρωνικά Δίκτυα (Artificial Neural Networks) είναι μία βελτίωση των νευρικών δικτύων ενός επιπέδου. Η πρώτη διαφορά που έχουν είναι το πλήθος των επιπέδων, όπου τα Πολυστρωματικά Νευρωνικά Δίκτυα μπορούν να περιέχουν αρκετά κρυφά επίπεδα. Έχοντας πολλά επίπεδα τα δεδομένα επεξεργάζονται καλύτερα και πλέον μπορούν λυθούν πιο σύνθετα προβλήματα [22]. Τα Πολυστρωματικά Νευρωνικά Δίκτυα μπορούν να επιλύσουν μη γραμμικά διαχωρίσιμα προβλήματα με την χρήση των κρυφών επιπέδων και αυτό τα καθιστά πολύ χρήσιμα [20] [22]. Κάθε τεχνητός νευρώνας είναι συνδεδεμένος με όλους τους νευρώνες του επόμενου επιπέδου και αυτό αυξάνει την επεξεργαστική ισχύ του δικτύου όμως αυτό αυξάνει την πολυπλοκότητα τους. Εφαρμόζονται συνήθως για την αναγνώριση ομιλίας, για εφαρμογές όρασης υπολογιστών και σε συστήματα που παράγονται προβλέψεις [20] [21]. Μερικοί ακόμη βασικοί τύποι Νευρωνικών Δικτύων είναι το Kohonen Νευρωνικό Δίκτυο και το Νευρωνικό Δίκτυο Οπισθοδιάδοσης (Back-Propagation Neural Network - BPNN).

Το Kohonen Νευρωνικό Δίκτυο είναι ένα δίκτυο το οποίο δέχεται δεδομένα πολλών διαστάσεων. Έπειτα αυτά τα δεδομένα ομαδοποιούνται και με αυτό τον τρόπο γίνεται μία συμπίεση των δεδομένων σε μικρότερες διαστάσεις όμως διατηρείται το αρχικό περιεχόμενό τους. Τα Kohonen Νευρωνικά Δίκτυα χρησιμοποιούνται κυρίως στην ιατρική για την αναγνώριση προτύπων και την κατηγοριοποίηση των δεδομένων [22] [23].

Το Νευρωνικό Δίκτυο Οπισθοδιάδοσης κατά την εκπαίδευση, προσπαθεί να ελαχιστοποιήσει το σφάλμα βάση των επιθυμητών αποτελεσμάτων που θέλουμε να έχουμε. Αν η έξοδος που παράγεται αποκλίνει από την επιθυμητή έξοδο, τότε μεταδίδεται προς τα πίσω ώστε το μοντέλο να μάθει από τα λάθη του και να γίνει βελτιστοποίηση των παραμέτρων του. Ουσιαστικά μεταδίδεται πίσω στους νευρώνες το σφάλμα ώστε κάθε νευρώνες να δει κατά πόσο ευθύνεται γι' αυτό.

Αυτή η διαδικασία επαναλαμβάνεται έως ότου οι τιμές των βαρών ελαχιστοποιήσουν το σφάλμα σε ανεκτό σημείο, όπου θα μπορούν να παράγονται οι επιθυμητές έξοδοι [24] [25].

Όπως συμπεραίνουμε, τα Νευρωνικά Δίκτυα έχουν πολλές δυνατότητες και με την ανακάλυψη και την εξέλιξη τους οι άνθρωποι επωφελούνται αρκετά. Χρησιμοποιούνται αρκετά στην καθημερινότητα μας καθώς είναι έμπιστα και ικανά να επιλύουν προβλήματα πραγματικού κόσμου. Για παράδειγμα στον αγροτικό τομέα, η χρήση Νευρωνικών Δικτύων είναι χρήσιμη καθώς η φύση αρκετών μηχανημάτων αφήνει τα περιθώρια στα Νευρωνικά Δίκτυα να ενσωματωθούν σε αυτά ώστε να τα ελέγχουν. Η επεκτασιμότητα και η προσαρμοστικότητα τους σε διάφορες καταστάσεις, είναι τα χαρακτηριστικά τα οποία χρίζουν τα Νευρωνικά Δίκτυα χρήσιμα για τέτοιου είδους ανάγκες και προβλήματα [26].

Μπορούν να χρησιμοποιηθούν επίσης και σε περιπτώσεις όπου είναι αναγκαία η επίβλεψη ορισμένων προϊόντων. Είναι εφικτό τα Νευρωνικά Δίκτυα να έχουν εκπαιδευθεί έτσι ώστε να αναγνωρίζουν χρώματα εικόνων και να καταλαβαίνουν εάν το χρώμα που έχει ένα φρούτο είναι φυσιολογικό ή όχι [26]. Μπορεί να αναγνωρίσει τι φρούτο είναι, αν το χρώμα που έχει είναι σύνηθες δεδομένου της εποχής που βρισκόμαστε και αν είναι υγιές. Αυτό είναι εξαιρετικά σημαντικό γιατί ένα πρόβλημα εντοπίζεται και αντιμετωπίζεται πριν κλιμακωθεί. Επίσης μπορεί να επιτευχθεί και η αναγνώριση προσώπων μέσω καμερών. Οι κάμερες μίας εταιρείας, χρησιμοποιώντας ένα κατάλληλα εκπαιδευμένο νευρωνικό δίκτυο, μπορούν να παρακολουθούν τις εισόδους της εταιρείας και ένα άτομο έχει πρόσβαση μόνο σε περίπτωση που δουλεύει στην εταιρεία. Σε αντίθετη περίπτωση, θα ενημερώνεται το Προσωπικό Ασφαλείας για την ύπαρξη αυτού του άγνωστου προσώπου στον χώρο της εταιρείας.

2.2.2 Βαθιά Μάθηση

Εφόσον αναλύθηκαν μέχρι ένα σημείο τα Νευρωνικά Δίκτυα, είναι εφικτή πλέον η ανάλυση της επόμενης υποκατηγορίας που είναι η Βαθιά Μάθηση ή αλλιώς, Deep Learning. Η Βαθιά Μάθηση είναι μία από τις βασικές υποκατηγορίες της Τεχνητής Νοημοσύνης, όμως μπορεί να θεωρηθεί και ως ένα υποσύνολο της Μηχανικής Μάθησης [27]. Η Μηχανική Μάθηση ουσιαστικά εκπαιδεύει τα υπολογιστικά συστήματα και τους δείχνει πως να μαθαίνουν από τα δεδομένα και πως να τα επεξεργάζονται, όμως με την Βαθιά Μάθηση τα υπολογιστικά συστήματα μπορούν

να είναι αυτοδίδακτα [8] [17] [28] [29]. Δηλαδή ένας υπολογιστής μπορεί πλέον να μάθει πως να μαθαίνει και πως να επεξεργάζεται τις διάφορες πληροφορίες βάση ενός μεγάλου συνόλου δεδομένων αναπαράγοντας τον τρόπο που λειτουργεί ο ανθρώπινος εγκέφαλος. Οι κύριες διαφορές τους βρίσκονται στον τρόπο με τον οποίο μαθαίνει ο κάθε αλγόριθμος και στο πλήθος των δεδομένων που χρησιμοποιεί κάθε αλγόριθμος.

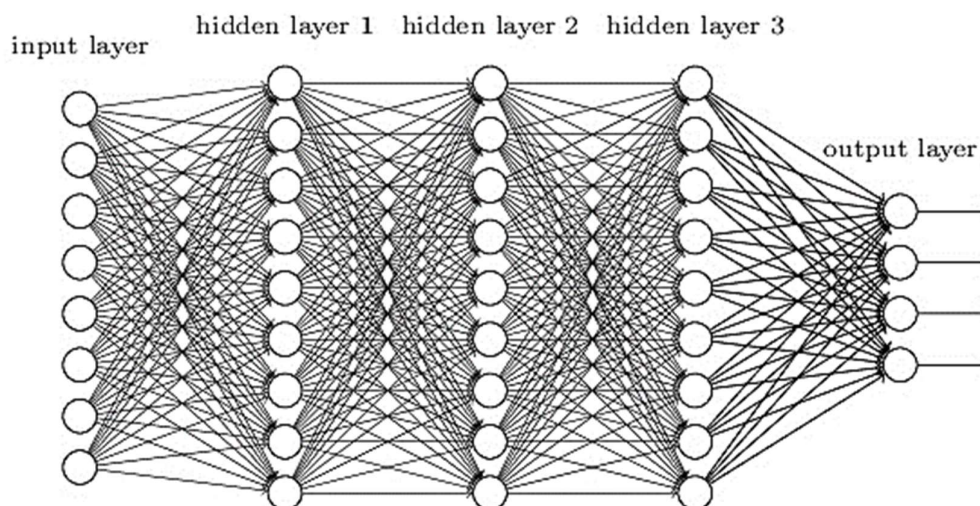
Αρχικά κατά την διάρκεια της εκμάθησης, τα μοντέλα Βαθιάς Μάθησης επιτρέπουν τη χρήση μεγάλων συνόλων δεδομένων [30]. Η εξαγωγή των χαρακτηριστικών πραγματοποιείται από το ίδιο το δίκτυο. Έτσι δεν υπάρχει η ανάγκη ανθρώπινης παρέμβασης κατά την διάρκεια εξαγωγής των χαρακτηριστικών [27] [31]. Αυτή η δυνατότητα είναι ιδιαίτερα σημαντική διότι το μεγαλύτερο πλήθος των δεδομένων πλέον είναι μη δομημένο (unstructured). Επίσης αυτά τα μοντέλα μπορούν να επεξεργαστούν και επισημασμένα σύνολα δεδομένων (labeled datasets) με την μέθοδο της επιβλεπόμενης μάθησης, αλλά γενικά αυτά τα μοντέλα δεν βασίζονται στα επισημασμένα σύνολα δεδομένων [27]. Με λίγα λόγια η επιβλεπόμενη μάθηση βασίζεται στα επισημασμένα δεδομένα (labeled data) τα οποία κατηγοριοποιούνται σε κατηγορίες. Με βάση αυτά ο αλγόριθμος μαθαίνει τα χαρακτηριστικά των δεδομένων κάθε κατηγορίας με σκοπό να τα αναγνωρίζει [32].

Σε αντίθετη περίπτωση, στα μοντέλα Μηχανικής Μάθησης είναι αναγκαία η ανθρώπινη παρέμβαση κατά την διάρκεια της εκπαίδευσης. Ο άνθρωπος ο οποίος θα παρέμβει κατά την διάρκεια της εκπαίδευσης, καθορίζει την ιεραρχία των χαρακτηριστικών ώστε να γίνουν κατανοητές οι διαφορές των δεδομένων εισόδου. Συνήθως χρησιμοποιούνται περισσότερα δομημένα δεδομένα (structured data) για την εκμάθηση ενός μοντέλου Μηχανικής Μάθησης [27]. Αν θέλαμε να εκπαιδεύσουμε ένα μοντέλο το οποίο αναγνωρίζει μήλα και κεράσια μέσω εικόνων, πρέπει ο άνθρωπος που θα παρέμβει να καθορίσει τα χαρακτηριστικά που διακρίνουν κάθε φρούτο. Μερικά από αυτά μπορεί να είναι το χρώμα και το μέγεθος. Ιεραρχικά το μέγεθος είναι ψηλότερα από το χρώμα διότι είναι ένα χαρακτηριστικό το οποίο διακρίνει το κεράσι από το άλλο φρούτο. Αλλιώς μπορούν να χρησιμοποιηθούν ετικέτες στα δεδομένα όπως «μήλο», «κεράσι» ώστε μέσω της επιβλεπόμενης μάθησης να βελτιωθεί η διαδικασία της εκπαίδευσης [27].

Τα μοντέλα Βαθιάς Μάθησης δέχονται ως εισόδους διάφορες πηγές δεδομένων όπως κείμενα, εικόνες, ήχο και βίντεο. Τα δεδομένα που δέχονται, μπορούν να είναι

και μη δομημένα και μπορούν από αυτά να προσδιοριστούν τα χαρακτηριστικά που τα διακρίνουν, αυτόματα. Έχοντας αυτά κατά νου καταλαβαίνουμε ότι τα μοντέλα Βαθιάς Μάθησης, αναλύοντας τα δεδομένα εισόδου εξάγουν μοτίβα και ομαδοποιούν τα δεδομένα κατάλληλα. Οπότε με βάση το προηγούμενο παράδειγμα, ένα τέτοιο δίκτυο θα μπορούσε να δεχθεί τις εικόνες των μήλων και των κερασιών και να τις αντιστοιχίσει στις κατάλληλες κατηγορίες βάση των ομοιοτήτων και των διαφορών τους. Τέλος, όσων αφορά την δεύτερη διαφορά τους, τα μοντέλα Βαθιάς Μάθησης χρειάζονται περισσότερα δεδομένα κατά την εκπαίδευση τους ώστε να βελτιστοποιηθεί η ακρίβεια τους, ενώ ένα μοντέλο Μηχανικής Μάθησης δεν χρειάζεται πολλά δεδομένα ώστε να εκπαιδευθεί λόγω της διαδικασίας από την οποία περνάνε τα δεδομένα εισόδου [27].

Η Βαθιά Μάθηση δεν μπορεί να υπάρξει χωρίς τα Νευρωνικά Δίκτυα καθώς βασίζεται στην λογική και την δομή τους [17] [27]. Η διαφοροποίηση που διακρίνεται μεταξύ τους είναι στο πλήθος των επιπέδων, δηλαδή ένας αλγόριθμος Βαθιάς Μάθησης χρειάζεται παραπάνω από τρία επίπεδα για να μπορέσει να λειτουργήσει και να παράξει τα επιθυμητά αποτελέσματα, άρα είναι αναγκαία η χρήση πολλαπλών κρυφών επιπέδων [8] [17] [30] [33] [34]. Σε αυτή την περίπτωση έχουμε ένα Deep Neural Network (DNN), δηλαδή ένα δίκτυο διασυνδεδεμένων τεχνητών νευρώνων χωρισμένων σε πολλαπλά κρυφά επίπεδα. Μόλις τα δεδομένα τροφοδοτηθούν στο δίκτυο, οι τεχνητοί νευρώνες αναλύουν και πραγματοποιούν μαθηματικές πράξεις στα δεδομένα έως ότου έχουμε ανεκτό ποσοστό ακρίβειας [35]. Χρειάζεται ένα μεγάλο πλήθος δεδομένων για να γίνει σωστή εκπαίδευση καθώς και αρκετή υπολογιστική ισχύς [29] [30].



Εικόνα 4: Αναπαράσταση ενός Deep Neural Network [36]

Στον χώρο της Βαθιάς Μάθησης υπάρχουν διάφοροι τύποι Νευρωνικών Δικτύων, με διαφορετικές λειτουργίες και με διαφορετικούς σκοπούς. Οι πιο βασικοί τύποι Νευρωνικών Δικτύων που χρησιμοποιούνται στην Βαθιά Μάθηση είναι τα εξής:

- Τεχνητό Νευρωνικό Δίκτυο (Artificial Neural Network - ANN)
- Συνελικτικό Νευρωνικό Δίκτυο (Convolutional Neural Networks - CNN)
- Ανατροφοδοτούμενο Νευρωνικό Δίκτυο (Recursive/Recurrent Neural Network - RNN)

Ο πρώτος τύπος Νευρωνικών Δικτύων είναι τα Τεχνητά Νευρωνικά Δίκτυα (ANN). Αυτά τα Νευρωνικά Δίκτυα είναι ουσιαστικά τα Πολυστρωματικά Νευρωνικά Δίκτυα που παρουσιάστηκαν στο προηγούμενο υποκεφάλαιο. Ονομάζονται και Feed-Forward Neural Networks (FFNN) επειδή οι συνδέσεις μεταξύ των τεχνητών νευρώνων έχουν μία συγκεκριμένη φορά (forward). Άρα τα δεδομένα μεταβιβάζονται στον επόμενο κόμβο, δεν γίνεται να μεταβιβαστούν προς τα πίσω [31].

Το επόμενο Νευρωνικό Δίκτυο Βαθιάς Μάθησης που θα αναλυθεί είναι το Συνελικτικό Νευρωνικό Δίκτυο (CNN). Η έμπνευση για την δημιουργία των Συνελικτικών Νευρωνικών Δικτύων προέρχεται από τον οπτικό φλοιό ο οποίος αποτελείται από διάφορες περιοχές οι οποίες βοηθάνε τον άνθρωπο να αντιλαμβάνεται τις κινήσεις, τα χρώματα, τα σχήματα, τις αποστάσεις και γενικά όλες τις λειτουργίες που αφορούν την όραση [37]. Τα Συνελικτικά Νευρωνικά Δίκτυα χρησιμοποιούνται κυρίως σε εφαρμογές που ασχολούνται με την ανίχνευση αντικειμένων, ανίχνευση γωνιών, αναγνώριση αντικειμένων και πολλά άλλα, καθώς

σε τέτοιες εφαρμογές επιτυγχάνει συναρπαστικά αποτελέσματα λόγω των συνελκτικών επιπέδων [38]. Ένα από τα σημαντικά χαρακτηριστικά των CNNs είναι ότι επιτυγχάνουν την μείωση των παραμέτρων και αυτό βοήθησε πάρα πολύ τους ερευνητές και τους προγραμματιστές να δημιουργήσουν μεγαλύτερα μοντέλα τα οποία θα επιλύουν πιο πολύπλοκες διαδικασίες. Επίσης, καθώς οι είσοδοι προχωράνε στα επόμενα επίπεδα συνδυάζονται με τα χαρακτηριστικά που έχουν βρεθεί από πριν και παράγονται χαρακτηριστικά υψηλότερου επιπέδου. Δηλαδή, στο πρώτο επίπεδο θα αναγνωριστούν οι γωνίες, στο επόμενο επίπεδο θα αναγνωριστούν πιο απλά σχήματα βάση των χαρακτηριστικών που βρέθηκαν από το πρώτο επίπεδο και όσο μεταβιβάζονται σε επόμενα επίπεδα αναγνωρίζονται χαρακτηριστικά υψηλότερου επιπέδου όπως είναι ένα πρόσωπο [38].

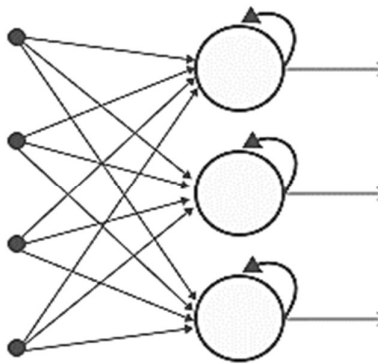


Εικόνα 5: CNN: Διαδικασία εξαγωγής χαρακτηριστικών [39]

Η αναγνώριση των χαρακτηριστικών και των προτύπων σε κάθε επίπεδο γίνεται με την χρήση των φίλτρων. Τα φίλτρα είναι σαν ένας πίνακας, του οποίου τις διαστάσεις τις ορίζει ο προγραμματιστής, όπου διανύει όλη την εικόνα και εντοπίζει τα σημεία ενδιαφέροντος [21] [38] [40]. Χρησιμοποιώντας τα παραπάνω εργαλεία και τεχνικές, τα Συνελκτικά Νευρωνικά Δίκτυα είναι πιο έμπιστα για εφαρμογές που ασχολούνται με την ανάλυση εικόνας, την αναγνώριση ομιλίας και κυρίως στην όραση υπολογιστών. Μερικά από τα αρνητικά του είναι η πολυπλοκότητα του κατά την σχεδίαση και η ταχύτητα του ειδικά όταν υπάρχουν πολλά κρυφά επίπεδα [20].

Ο τελευταίος τύπος Νευρωνικών Δικτύων Βαθιάς Μάθησης που θα αναλυθεί ονομάζεται Ανατροφοδοτούμενο Νευρωνικό Δίκτυο. Όταν διαβάζει ο άνθρωπος ένα βιβλίο, είναι ανέφικτο να κατανοηθεί το νόημα του βιβλίου μόνο από την τρέχουσα σελίδα που βρίσκεται. Για να επιτευχθεί αυτό, πρέπει να συνδυαστούν οι

πληροφορίες και των προηγούμενων σελίδων. Μόνο έτσι ένας άνθρωπος μπορεί να κατανοήσει πραγματικά το θέμα και το νόημα ενός βιβλίου και γενικότερα ενός κειμένου. Αυτή η διαδικασία υλοποιείται από τα Ανατροφοδοτούμενα Νευρωνικά Δίκτυα. [21] [41] [42]. Τα RNNs δέχονται ως εισόδους δεδομένα χρονοσειρών και διαδοχικά δεδομένα. Αφού τροφοδοτηθούν τα δεδομένα στο δίκτυο, στην συνέχεια περνούν από τα επίπεδα του. Καθώς διανύουν τα επίπεδα, παράγονται πληροφορίες και συσχετίσεις για τα δεδομένα. Η λειτουργία που κάνει ένα δίκτυο RNN να ξεχωρίζει, συγκριτικά με τα υπόλοιπα Νευρωνικά Δίκτυα που έχουν εξηγηθεί, είναι ότι επαναχρησιμοποιεί τις πληροφορίες που παράχθηκαν από προηγούμενες εισόδους. Σε όλα τα δίκτυα που παρουσιάστηκαν παραπάνω, οι εισοδοί δεν έχουν καμία σχέση με τις εξόδους, όμως οι εξοδοί των RNNs εξαρτώνται από τα προηγούμενα στοιχεία της ακολουθίας [41] [42] [43]. Για να πραγματοποιήσουν αυτή την λειτουργία τα RNNs, έπρεπε να τροποποιηθούν σχηματικά. Εκτός από τις συνδέσεις οι οποίες συνδέουν κάθε στοιχείο με κάποιο άλλο στοιχείο επόμενου επιπέδου, υπάρχουν συνδέσεις οι οποίες καταλήγουν στο ίδιο επίπεδο ή σε προηγούμενο [20] [21] [40] [41] [42] [43].



Εικόνα 6: Σχηματική αναπαράσταση ενός Ανατροφοδοτούμενου Νευρωνικού Δικτύου [31]

Αυτή είναι η βασική ιδέα των Ανατροφοδοτούμενων Νευρωνικών Δικτύων. Υπάρχουν διάφοροι τύποι Ανατροφοδοτούμενων Νευρωνικών Δικτύων, με διαφορετικές λειτουργίες και με διαφορετικούς σκοπούς το καθένα. Ο πιο βασικός τύπος Ανατροφοδοτούμενων Νευρωνικών Δικτύων είναι τα Δίκτυα Νευρώνων Μακράς-Βραχείας Μνήμης (Long Short-Term Memory - LSTM).

Στα Ανατροφοδοτούμενα Νευρωνικά Δίκτυα, μία πληροφορία επηρεάζει την τρέχουσα κατάσταση όμως μπορεί αυτή η πληροφορία να μην προέρχεται από το κοντινό παρελθόν. Σε αυτή την περίπτωση, το RNN δεν θα κάνει ακριβή πρόβλεψη γιατί δεν θα μπορεί να συσχετίσει τις κατάλληλες πληροφορίες ώστε να παραχθεί σωστή πρόβλεψη. Γι' αυτό τον λόγο δημιουργήθηκαν τα Δίκτυα Νευρώνων Μακράς-Βραχείας Μνήμης. Τα LSTM για να το πετύχουν αυτό, περιέχουν κύτταρα μνήμης (memory cell) στα κρυφά επίπεδα, τα οποία απαρτίζονται από την πύλη εισόδου (input gate) η οποία ελέγχει το πότε οι νέες πληροφορίες θα εισέλθουν στο κύτταρο μνήμης, από την πύλη λήθης (forget gate) η οποία ελέγχει πόση πληροφορία από προηγούμενες καταστάσεις διατηρείται και από την πύλη εξόδου (output gate) η οποία ελέγχει τις πληροφορίες που θα χρησιμοποιηθούν εκτός του κυττάρου μνήμης [21] [41] [43].

Γενικά, τα Ανατροφοδοτούμενα Νευρωνικά Δίκτυα, είναι απαραίτητα για την υλοποίηση αρκετών σύγχρονων λειτουργιών και εφαρμογών. Για παράδειγμα, η Siri, ο προσωπικός βοηθός που είναι ενσωματωμένος στις συσκευές της εταιρείας Apple, χρησιμοποιεί Ανατροφοδοτούμενα Νευρωνικά Δίκτυα για να μπορεί να κατανοεί διάφορες γλώσσες και να εκτελεί τις εντολές που της δίνουν οι κάτοχοι των συσκευών αναλύοντας τις προτάσεις που ακούει (εντολές). Ο προσωπικός βοηθός γνωρίζει τις ενέργειες που πρέπει να πραγματοποιήσει ώστε να εκτελεστεί η κατάλληλη εντολή. Μερικές ακόμα γνωστές εφαρμογές των Ανατροφοδοτούμενων Νευρωνικών Δικτύων είναι η ανάλυση συναισθήματος στα κείμενα, η μετάφραση λέξεων ή φράσεων από μία γλώσσα σε μία άλλη, εντοπισμός ορθογραφικών και γραμματικών λαθών [20] [43].

2.2.3 Επεξεργασία Φυσικής Γλώσσας

Ακολουθεί η περιληπτική αναφορά της υποκατηγορίας της Τεχνητής Νοημοσύνης που ονομάζεται Επεξεργασία Φυσικής Γλώσσας ή εν συντομία NLP (Natural Language Processing). Αν παρατηρήσουμε τα κινητά του 21^{ου} αιώνα, τα περισσότερα αν όχι όλα, διαθέτουν την λειτουργία της υπαγόρευσης κειμένου. Με αυτή την λειτουργία, ο κάτοχος του κινητού μπορεί να υπαγορεύει προτάσεις και αυτές να γράφονται ως κείμενο στο κινητό. Αρκετοί θα γνωρίζουν πλέον πως η συγκεκριμένη λειτουργία έχει άμεση σχέση με την Τεχνητή Νοημοσύνη, όμως λίγοι αντιλαμβάνονται ότι για να δημιουργηθεί η λειτουργία αυτή ήταν αναγκαία η χρήση του NLP. Το NLP είναι μία υποκατηγορία της Τεχνητής Νοημοσύνης η οποία

ασχολείται με την ανάλυση και την επεξεργασία των γλωσσών που χρησιμοποιεί ο άνθρωπος για να επικοινωνήσει, κάνοντας χρήση διάφορων υπολογιστικών τεχνικών [44] [45] [46]. Η επικοινωνία με τους υπολογιστές είτε μπορεί να γίνεται με την χρήση της φωνής, όπως είδαμε και με την υπαγόρευση κειμένου, μέσω γραπτού κειμένου ή μέσω εισαγωγής δεδομένων με την χρήση ενός πληκτρολογίου [29] [45]. Με την χρήση του NLP, τα υπολογιστικά συστήματα μπορούν να καταλαβαίνουν την γλώσσα που μιλάμε και να πραγματοποιούν διάφορες λειτουργίες όπως μετάφραση γλωσσών [8] [35] [44] [45]. Απώτερος σκοπός του NLP είναι η επίτευξη επεξεργασίας γλωσσών από τους υπολογιστές, παρόμοια με αυτή των ανθρώπων. Έτσι θα μπορέσει να υπάρξει επικοινωνία μεταξύ ανθρώπων και υπολογιστών [28] [45].

Προκειμένου ο άνθρωπος να κατανοήσει ή να παράξει κάποια πρόταση, χρειάζεται να ενεργοποιηθούν διάφορα σημεία στον εγκέφαλό του ώστε να γίνει η κατάλληλη γλωσσική ανάλυση [44]. Υπάρχουν διάφορα επίπεδα γλωσσικής ανάλυσης τα οποία χρησιμοποιεί ο ανθρώπινος εγκέφαλος. Σε ένα σύστημα NLP μπορεί να επιλεχθεί ποιο ή ποια επίπεδα γλωσσικής ανάλυσης θα χρησιμοποιηθούν ανάλογα με τις ανάγκες της εφαρμογής [44] [45]. Στην συνέχεια θα παρουσιαστούν τα επίπεδα γλωσσικής ανάλυσης που υπάρχουν.

Όταν μία εφαρμογή απαιτεί την χρήση ενός συστήματος NLP το οποίο θα δέχεται ως είσοδο την ανθρώπινη φωνή, τότε είναι απαραίτητη η φωνολογική ανάλυση. Το συγκεκριμένο επίπεδο διαχειρίζεται και ερμηνεύει την φωνή, κωδικοποιώντας τα ηχητικά κύματα. Έτσι παράγει ένα ψηφιοποιημένο σήμα, το οποίο στη συνέχεια, με βάση διάφορους κανόνες, το ερμηνεύει και παράγει αποτελέσματα [44].

Έπειτα ακολουθεί η μορφολογική ανάλυση. Όταν ο άνθρωπος συναντάει μια άγνωστη λέξη, τότε την αναλύει και βρίσκει τα συστατικά της μορφώματα (πρόθεμα, επίθεμα, ριζικό κλπ.). Ένα σύστημα NLP, εκμεταλλεύεται αυτή την διαδικασία και την χρησιμοποιεί προκειμένου να πραγματοποιήσει την μορφολογική ανάλυση. Αναλύοντας μία λέξη και βρίσκοντας τα συστατικά της μορφώματα, ένα σύστημα μπορεί να προσδίδει νόημα στις λέξεις που δέχεται [44] [46].

Στην συνέχεια περιγράφεται η λεκτική ανάλυση. Στην λεκτική ανάλυση, το κείμενο διαιρείται σε παραγράφους, προτάσεις και λέξεις, εντοπίζεται και αναλύεται δηλαδή η δομή του λόγου. Χρησιμοποιείται ένα λεξικό της εκάστοτε γλώσσας το οποίο είναι γεμάτο με λεξήματα. Μπορεί να εφαρμοστεί για την ανίχνευση ανεπιθύμητων

μηνυμάτων, όμως για πιο σύνθετες εφαρμογές χρειάζεται ένα πιο προηγμένο σύστημα [44] [47].

Στις περισσότερες γλώσσες που χρησιμοποιούν οι άνθρωποι, η σύνταξη των προτάσεων είναι πολύ σημαντική γι' αυτό και χρειάζεται προσοχή καθώς η παραμικρή αλλαγή στην σύνταξη μίας πρότασης μπορεί να αλλάξει τελείως το νόημα της. Γι' αυτό χρησιμοποιείται το επίπεδο της συντακτικής ανάλυσης. Το επίπεδο της συντακτικής ανάλυσης ασχολείται με την ανάλυση των λέξεων σε μία πρόταση, ψάχνοντας τις εξαρτήσεις που έχουν οι λέξεις μεταξύ τους [44] [46].

Μία λέξη μπορεί να έχει διαφορετική σημασία ανάλογα με τον τρόπο που θα διατυπωθεί μέσα σε μία πρόταση. Προκειμένου ένα σύστημα NLP να καταλαβαίνει κάθε φορά την σωστή σημασία που έχει μία λέξη στην πρόταση, χρησιμοποιείται η σημασιολογική ανάλυση. Στο συγκεκριμένο επίπεδο καθορίζονται οι πιθανές έννοιες των λέξεων σε μία πρόταση [44] [45].

Το επίπεδο της συνδιάλεξης (discourse) ασχολείται με την ανάλυση της δομής και του νοήματος σε κείμενα μεγάλων εκτάσεων. Δεν αναλύει απλά την σύνδεση των λέξεων σε μία πρόταση, αντίθετα, προσπαθεί να βρει τις συσχετίσεις που υπάρχουν μεταξύ των λέξεων και των προτάσεων [44] [47].

Το τελευταίο επίπεδο που θα αναλυθεί ονομάζεται ρεαλιστική/πρακτική ανάλυση. Σε αυτό το επίπεδο βρίσκεται η πραγματική σημασία των λέξεων μέσα σε μία πρόταση, χωρίς αυτή να υπάρχει ήδη μέσα στην πρόταση. Το σύστημα πρέπει να αντιλαμβάνεται τις προθέσεις και τους στόχους που κρύβονται σε μία πρόταση. Επίσης πρέπει να χρησιμοποιήσει αρκετή γνώση η οποία προέρχεται από τον πραγματικό κόσμο, ώστε το σύστημα να καταλάβει το πραγματικό νόημα που μπορεί να κρύβεται πίσω από μία λέξη ή πρόταση. Για παράδειγμα, σε ένα σύστημα εκχωρείται η φράση «Μου έσπασες τα νεύρα», σημασιολογικά το ρήμα «έσπασες» σημαίνει χαλάω ή καταστρέφω όμως πρακτικά καταλαβαίνουμε ότι το νόημα είναι διαφορετικό αφού το ρήμα «έσπασες» χρησιμοποιείται μεταφορικά. Η δουλειά αυτού του επιπέδου είναι να βρίσκει και να κατανοεί αυτές τις περιπτώσεις [44] [47].

Γνωρίζοντας αυτά, γίνεται αντιληπτό το εύρος χρήσης αυτής της υποκατηγορίας. Έχουν φτιαχτεί πολλές εφαρμογές οι οποίες αξιοποιούν το NLP και χρησιμοποιούνται αρκετά σήμερα. Μερικές από αυτές είναι οι εξής [44] [45]:

- Εξαγωγή Πληροφοριών (Information Extraction)
- Ανάκτηση Πληροφορίας (Information Retrieval)
- Συστήματα Διαλόγου (Dialog Systems)

- Σύνοψη (Summarization)
- Αυτόματη Μετάφραση(Machine Translation)

Ένα κλασσικό παράδειγμα εφαρμογής του NLP είναι η Siri, ο προσωπικός βοηθός που βρίσκουμε στις συσκευές της εταιρείας Apple. Η Siri για να μπορεί να πραγματοποιεί όλες αυτές τις λειτουργίες, κάνει χρήση αρκετών μεθόδων γλωσσικής ανάλυσης. Με βάση τα παραπάνω, μπορούν να δημιουργηθούν παρόμοιες εφαρμογές οι οποίες θα κάνουν πιο εύκολη και ευχάριστη την αλληλεπίδραση με τις συσκευές.

2.2.4 Γνωστική Υπολογιστική

Η επόμενη υποκατηγορία που πρόκειται να αναλυθεί είναι η Γνωστική Υπολογιστική (Cognitive Computing). Τα ευφυή υπολογιστικά συστήματα, στοχεύουν στην διευκόλυνση και την καλυτέρευση της ποιότητας ζωής του ανθρώπου. Επεξεργάζονται μεγάλους όγκους δομημένων και μη δομημένων δεδομένων πολύ πιο γρήγορα και αποτελεσματικά από τον άνθρωπο. Έτσι, τα συγκεκριμένα συστήματα μπορούν να φανούν πολύ χρήσιμα ως βοηθοί και σύμβουλοι του ανθρώπου καθώς με βάση τα αποτελέσματα ενός μοντέλου, ο άνθρωπος θα μπορεί εύκολα να παράγει χρήσιμα συμπεράσματα. Όλες αυτές οι λειτουργίες που αναφέρθηκαν φαίνονται αρκετές, όμως με την Γνωστική Υπολογιστική δημιουργούνται νέοι ορίζοντες και παύουν να ισχύουν τα όρια που έως σήμερα ήταν γνωστά.

Τα τελευταία χρόνια έχουν δοθεί αρκετοί ορισμοί για την Γνωστική Υπολογιστική, οι οποίοι δεν απέχουν αρκετά μεταξύ τους. Συνδυάζοντας τους αποτυπώνεται ένας εξαιρετικά περιεκτικός ορισμός από τον οποίο γίνεται εύκολα αντιληπτός ο σκοπός και η χρήση της συγκεκριμένης υποκατηγορίας. Η Γνωστική Υπολογιστική είναι ένα διεπιστημονικό πεδίο έρευνας και εφαρμογών, που έχει ως στόχο την δημιουργία μοντέλων και μηχανισμών λήψης απόφασης χρησιμοποιώντας τον τομέα της ψυχολογίας, της νευροεπιστήμης, της φυσικής, της στατιστικής ανάλυσης, της επεξεργασίας σήματος και της γνωστικής επιστήμης [48] [49]. Πρακτικά, ένα μοντέλο Γνωστικής Υπολογιστικής είναι ικανό να παράγει συλλογισμούς παρόμοιους με αυτούς του ανθρώπου. Επίσης, ενσωματώνεται η δυνατότητα παραγωγής σκέψεων, συναισθημάτων και πλέον τα συστήματα διαθέτουν και την ικανότητα της γνώσης [48] [49].

Με την χρήση αυτών των μοντέλων, ένα υπολογιστικό σύστημα θα μπορεί να επιλύει σύνθετα προβλήματα ακολουθώντας μια διαδικασία παρόμοια με αυτή της σκέψης [50] [51] [58]. Η γνωστική επιστήμη, η οποία αναφέρθηκε προηγουμένως, είναι μία επιστήμη η οποία ασχολείται με τον τρόπο που ο ανθρώπινος εγκέφαλος διαχειρίζεται την πληροφορία παρατηρώντας διάφορες πτυχές όπως είναι η λογική, το συναίσθημα, η γλώσσα επικοινωνίας, η προσοχή και η αντίληψη [49] [58]. Η γνωστική διαδικασία του ανθρώπου απαρτίζεται από συγκεκριμένα στάδια. Αρχικά ο άνθρωπος δέχεται πληροφορίες από το περιβάλλον μέσω των αισθητήριων οργάνων του όπως είναι τα μάτια του και το δέρμα του. Την πληροφορία του περιβάλλοντος την δέχεται ο οργανισμός ως είσοδο και στην συνέχεια αυτή μεταφέρεται στον εγκέφαλο μέσω των νεύρων. Εκεί γίνονται πολύπλοκες επεξεργασίες προκειμένου η είσοδος να επεξεργαστεί και να αποθηκευτεί στην μνήμη. Αφού γίνει η επεξεργασία, τα αποτελέσματα της μεταβιβάζονται σε διάφορα μέρη του σώματος μέσω του νευρικού συστήματος και αυτά με την σειρά τους αντιδρούν αναλόγως. Αυτός είναι ο τρόπος με τον οποίο κάθε άνθρωπος γνωρίζει και αντιλαμβάνεται το περιβάλλον που βρίσκεται σε όλη την διάρκεια της ζωής του. Άρα είναι απαραίτητη η καλύτερη κατανόηση του γνωστικού συστήματος καθώς αυτό θα βοηθήσει στην παραγωγή ακριβέστερων και πιο προηγμένων μοντέλων [49].

Τα έξυπνα υπολογιστικά συστήματα παρουσιάζουν παρόμοια χαρακτηριστικά. Πλέον μπορούν να:

- να ακούν την φωνή του ανθρώπου και να την επεξεργάζονται
- να αναλύουν τις γλώσσες επικοινωνίας που χρησιμοποιεί ο άνθρωπος
- να βλέπουν το περιβάλλον και να παρατηρούν
- να αισθάνονται διάφορα περιβαλλοντικά δεδομένα (θερμοκρασία, υγρασία κλπ.) και να τα επεξεργάζονται
- να κινούνται και να ελέγχουν τις κινήσεις τους

Μέσω όλων αυτών των δυνατοτήτων που έχουν αποκτήσει, τα ευφυή σύστημα μπορούν να μαθαίνουν από το περιβάλλον τους με την χρήση συσκευών (αισθητήρες, κάμερες, μικρόφωνα κλπ.). Εκτός από την αλληλεπίδραση με το περιβάλλον, τα έξυπνα υπολογιστικά συστήματα μπορούν να μαθαίνουν και μέσω της αλληλεπίδρασης με τους ανθρώπους [50]. Η Γνωστική Υπολογιστική έχει ένα ευρύ φάσμα εφαρμογής. Την συναντάμε σε τομείς όπως η ρομποτική, τα έμπειρα συστήματα, η όραση υπολογιστών, τα αυτόνομα οχήματα και πολλά άλλα. Έχουν

δημιουργηθεί ήδη εφαρμογές που αξιοποιούν τον τομέα της Γνωστικής Υπολογιστικής, μερικές από τις πιο γνωστές εφαρμογές είναι ο Watson της IBM, η Cortana της Microsoft και η Siri της Apple. Επίσης υπάρχουν εφαρμογές που έχουν φτιαχτεί για τα πανεπιστήμια και τα βιντεοπαιχνίδια. Έχει αναπτυχθεί ένας ευφυής ακαδημαϊκός σύμβουλος ο οποίος απαντάει ερωτήσεις σχετικές με το πανεπιστήμιο. Αυτό το σύστημα χρησιμοποιεί τους αλγορίθμους που χρησιμοποιεί το σύστημα Watson προκειμένου να καταλάβει την ερώτηση και να απαντήσει [50]. Όσο για τα βιντεοπαιχνίδια, αποτελεί μεγάλη πρόκληση η ενσωμάτωση της Γνωστικής Υπολογιστικής διότι είναι αναγκαία η ανάπτυξη σύνθετων μαθηματικών μοντέλων και η ύπαρξη ενός μεγάλου συνόλου δεδομένων για την εκπαίδευση των αλγορίθμων [50].

Τα μοντέλα Γνωστικής Υπολογιστικής, αναλύοντας μεγάλες ποσότητες δεδομένων και εκτελώντας διάφορες διεργασίες, βελτιστοποιούν την απόδοση των εφαρμογών μειώνοντας το κόστος και αυξάνοντας την αποτελεσματικότητα τους [49] [50]. Επίσης προσφέρουν χρήσιμες προτάσεις οι οποίες βοηθούν στην λήψη απόφασης [49]. Για να βοηθήσουν τους ανθρώπους στην λήψη απόφασης, παράγουν ένα σύνολο πιθανών λύσεων και τις προτείνουν [52]. Έπειτα ο άνθρωπος επιλέγει κάποια από αυτές τις λύσεις. Οι λύσεις που παράγονται, βασίζονται σε ένα μεγάλο πλήθος δεδομένων με το οποίο τροφοδοτούνται τα υπολογιστικά συστήματα. Για παράδειγμα, η Γνωστική Υπολογιστική βοηθάει τους γιατρούς και τους νοσοκόμους με την διάγνωση των ασθενών και με βάση την διάγνωση που θα γίνει, προτείνονται διάφορες θεραπείες και επιλέγεται η καταλληλότερη από τους ειδικούς [52]. Με την μίμηση της διαδικασίας της ανθρώπινης σκέψης, ο άνθρωπος μπορεί να δημιουργήσει συστήματα τα οποία θα τον βοηθούν να πάρει σωστές αποφάσεις σε διάφορες καταστάσεις [53] [57]. Ωστόσο, τα συγκεκριμένα συστήματα χρειάζονται αρκετές βελτιώσεις προκειμένου να μιμηθούν απόλυτα την διαδικασία δημιουργίας συναισθημάτων, την παραγωγή γνώσης, τον τρόπο επεξεργασίας των πληροφοριών και την λήψη αποφάσεων του ανθρώπινου εγκεφάλου [48] [57] [59].

2.2.5 Όραση Υπολογιστών

Η τελευταία υποκατηγορία της Τεχνητής Νοημοσύνης είναι η Όραση Υπολογιστών. Για πολλά χρόνια οι επιστήμονες προσπαθούσαν να εφεύρουν έναν τρόπο όπου οι υπολογιστές θα μπορούν να δουν και να αντιληφθούν τον χώρο

γύρω τους. Μετά από χρόνια έρευνας και προσπαθειών, ανακαλύφθηκαν αρκετές τεχνικές επεξεργασίας εικόνας και σε συνδυασμό με τα Νευρωνικά Δίκτυα και την Βαθιά Μάθηση δημιουργήθηκαν πολύ χρήσιμα μοντέλα Όρασης Υπολογιστών [54]. Η Όραση Υπολογιστών επιτρέπει στα υπολογιστικά συστήματα να επεξεργαστούν και να αναγνωρίσουν σε εικόνες και βίντεο, αντικείμενα (objects), ακριβώς όπως ένας άνθρωπος. Το υπολογιστικό σύστημα ερμηνεύει αυτό που βλέπει, παράγει συμπεράσματα και παίρνει αποφάσεις βάση των εικόνων που βλέπει [8] [35]. Προσπαθεί να μιμηθεί τον τρόπο λειτουργίας του ανθρώπινου οπτικού συστήματος. Κατανοώντας το καλύτερα, οι επιστήμονες μπορούν να φτιάξουν πιο ακριβή μοντέλα που θα διαθέτουν περισσότερες λειτουργίες [55]. Στοχεύει στην δημιουργία συστημάτων τα οποία εξάγουν πληροφορίες από τις εικόνες, προκειμένου να τις αναλύσουν, να κατανοήσουν το περιεχόμενό τους και να εξυπηρετήσουν τις ανάγκες μίας συγκεκριμένης εφαρμογής [56]. Για να βελτιωθούν τα αποτελέσματα που παράγουν τα μοντέλα Όρασης Υπολογιστών, είναι αναγκαία και η εξέλιξη των εξαρτημάτων που χρησιμοποιούν όπως για παράδειγμα οι κάμερες. Πλέον έχουν δημιουργηθεί κάμερες με εξαιρετική ανάλυση, εύκολη εγκατάσταση στο υπολογιστικό σύστημα και με πολλές λειτουργίες χωρίς να αυξάνεται πολύ το κόστος τους [60]. Μερικές από τις εφαρμογές που χρησιμοποιείται η Όραση Υπολογιστών είναι στα αυτόνομα αμάξια όπου τα βοηθάει να κατατοπιστούν στον χώρο και να πάρουν κατάλληλες αποφάσεις όσων αφορά την ταχύτητα του οχήματος, στην ρομποτική όπου το ρομπότ θα μπορεί να σχεδιάσει τις επόμενες κινήσεις του και στην παρακολούθηση και τον έλεγχο χώρων μέσω καμερών [56]. Ο υπολογιστής αντιλαμβάνεται και καταλαβαίνει αυτό που βλέπει και με βάση τους κανόνες με τους οποίους έχει προγραμματιστεί, λαμβάνει τις κατάλληλες αποφάσεις. Για παράδειγμα, σε ένα κατάστημα που είναι εγκατεστημένο ένα σύστημα καμερών, είναι εφικτό σε περίπτωση που ανιχνευθεί άνθρωπος μέσα στο κατάστημα μετά από κάποια συγκεκριμένη ώρα, να ειδοποιείται η αστυνομία και ο ιδιοκτήτης του καταστήματος για να γνωρίζουν ότι εντοπίστηκε ύποπτη κινητικότητα [55]. Έχοντας τις εικόνες και τα βίντεο ως εισόδους, ο αλγόριθμος Όρασης Υπολογιστών καταφέρνει να εξάγει από αυτά πληροφορία με την χρήση διάφορων τεχνικών. Μερικές από τις βασικότερες τεχνικές είναι οι εξής [55]:

- Ανίχνευση αντικειμένων (object detection)
- Ανίχνευση προσώπου (face detection)
- Ανίχνευση γωνιών (corner detection)

- Εντοπισμός αντικειμένων (object tracking)
- Ταίριασμα προτύπων (template matching).

Με τον συνδυασμό των παραπάνω τεχνικών, τα μοντέλα μπορούν να είναι πραγματικά έξυπνα και να παρέχουν λύσεις σε αρκετά προβλήματα που αντιμετωπίζουμε καθημερινά.

ΚΕΦΑΛΑΙΟ 3

ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

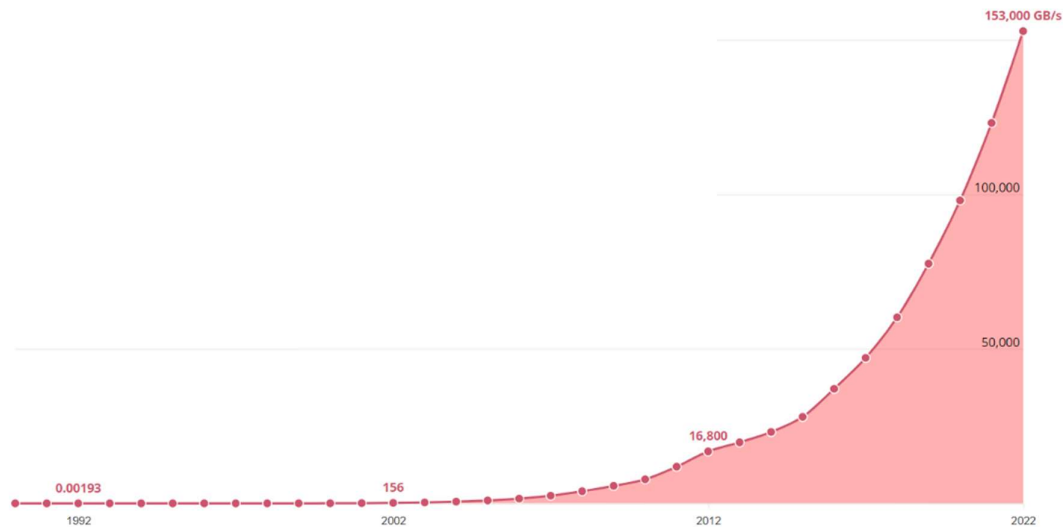
3.1 Τι είναι η Μηχανική Μάθηση

Στην εποχή που ζούμε με την εξέλιξη της τεχνολογίας, το πλήθος των δεδομένων έχει αυξηθεί δραματικά και συνεχίζει να αυξάνεται καθημερινά. Η πρόκληση για τους ερευνητές ήταν να ερμηνεύσουν αυτή τη μάζα δεδομένων και να παραχθούν συμπεράσματα από αυτά. Για τον σκοπό αυτό χρησιμοποιείται η Μηχανική Μάθηση (Machine Learning) [61]. Η Μηχανική Μάθηση είναι ένα υποπεδίο της επιστήμης των υπολογιστών με την χρήση του οποίου τα συστήματα υπολογιστών προσδίδουν νόημα στα δεδομένα. Ουσιαστικά πρόκειται για ένα υποεπίπεδο της Τεχνητής Νοημοσύνης που αναγνωρίζει πρότυπα στα δεδομένα χρησιμοποιώντας έναν αλγόριθμο [62]. Ο κύριος στόχος της Μηχανικής Μάθησης είναι η εκπαίδευση των υπολογιστικών συστημάτων με την χρήση ενός αλγορίθμου. Όσο εκπαιδεύεται ένα υπολογιστικό σύστημα αποκτά εμπειρία. Μέσω της εμπειρίας, τα υπολογιστικά συστήματα μαθαίνουν και αυτοβελτιώνονται οπότε δεν προγραμματίζονται εκ νέου, ούτε είναι απαραίτητη η ανθρώπινη παρέμβαση κατά την διάρκεια της εκπαίδευσης. Όλα εξαρτώνται από τον αλγόριθμο ο οποίος βελτιώνει την συμπεριφορά του όσο αυξάνεται η εμπειρία του [63] [64].

Ο άνθρωπος είναι το πιο ευφυές είδος που υπάρχει αυτή την στιγμή στην Γη, διότι μπορεί να λύσει σύνθετα προβλήματα, μπορεί να παράγει σύνθετες σκέψεις, να κατανοήσει σύνθετες έννοιες και να μάθει από την καθημερινότητα του. Αφού ο άνθρωπος είναι το εξυπνότερο είδος, γιατί είναι αναγκαία η εκπαίδευση των υπολογιστικών συστημάτων; Γιατί ο άνθρωπος χρειάζεται πολύ χρόνο για να επεξεργαστεί αυτό το πλήθος δεδομένων που παράγονται καθημερινά [61]. Με την Μηχανική Μάθηση ο προγραμματιστής μπορεί να εισάγει στον αλγόριθμο ένα μεγάλο πλήθος δεδομένων. Έπειτα ο αλγόριθμος αναλύει τα δεδομένα, παράγει προβλέψεις και λαμβάνει αποφάσεις με βάση τα δεδομένα εισόδου. Σε περίπτωση που ανιχνευθεί κάποιο σφάλμα τότε είναι εφικτό να διορθωθεί και να ενσωματωθεί στον αλγόριθμο αυτή η πληροφορία ώστε να βελτιωθεί η ικανότητα λήψης αποφάσεων του αλγορίθμου [63]. Άρα, στόχος της Μηχανικής Μάθησης είναι η

αποτελεσματική επίλυση προβλημάτων μεγάλης κλίμακας, η αυτοματοποίηση διάφορων καθημερινών εργασιών, η διευκόλυνση και η καλυτέρευση της ζωής του ανθρώπου.

Τα τελευταία χρόνια, με βάση έρευνες, ο όγκος των δεδομένων έχει αυξηθεί εκθετικά. Το 2020 εκτιμάται ότι κάθε άνθρωπος παρήγαγε 1.7 MB (MegaBytes) σε ένα δευτερόλεπτο και εκτιμάται ότι το πλήθος των δεδομένων που παράγονται στο διαδίκτυο ημερησίως αγγίζει τα 2.5 Quintillion Bytes [65] [66] [67]! Από αυτά γίνεται κατανοητό ότι έχει αυξηθεί δραματικά το πλήθος των χρηστών που χρησιμοποιούν το διαδίκτυο, καθώς και οι υπηρεσίες και οι εφαρμογές που βρίσκονται σε αυτό. Για την ακρίβεια, 4.71 δισεκατομμύρια είναι το πλήθος των χρηστών που είναι συνδεδεμένοι καθημερινά στο διαδίκτυο. Επίσης έχει αυξηθεί ο χρόνος που καταναλώνουν οι χρήστες στο διαδίκτυο, όπου πλέον ο μέσος χρήστης καταναλώνει 7 ώρες την ημέρα πλοηγώντας στο διαδίκτυο και συγκεκριμένα οι περισσότεροι από τους χρήστες χρησιμοποιούν το κινητό τηλέφωνο τους [65] [66] [67]. Φαίνεται ότι οι άνθρωποι την τωρινή εποχή αναζητούν συνέχεια και θέλουν να μάθουν πράγματα, χρειάζονται πληροφορίες. Αυτό συμπεραίνεται από το πλήθος των αναζητήσεων που γίνονται καθημερινά όπου σύμφωνα με την Google παράγονται 3.5 δισεκατομμύρια αναζητήσεις την ημέρα [65] [68]. Επιπρόσθετα, την ίδια χρονολογία εκτιμάται ότι το μέγεθος της κίνησης του διαδικτύου ήταν μεγαλύτερη από 3 Zettabytes δηλαδή 3 τρισεκατομμύρια GB (GigaBytes) και μέχρι το 2022 η κίνηση του διαδικτύου προβλέπεται ότι θα έχει αυξηθεί κατά 50% συγκριτικά με την κίνηση του διαδικτύου το 2020 [69]. Στην επομένη εικόνα φαίνεται η άνοδος της παγκόσμιας κίνησης (traffic) του διαδικτύου τα τελευταία 30 χρόνια.



Εικόνα 7: Η κίνηση του διαδικτύου τα τελευταία 30 χρόνια [69]

Όπως είδαμε, τα τελευταία χρόνια το πλήθος των δεδομένων και η κίνηση του διαδικτύου έχουν αυξηθεί δραματικά και θα συνεχίσουν να αυξάνονται. Αυτό το μεγάλο πλήθος των δεδομένων αποθηκεύεται και το μεγαλύτερο ποσοστό τους μένει ανεκμετάλλευτο ενώ θα μπορούσε να χρησιμοποιηθεί προς όφελος μας. Είναι αδύνατο όλα αυτά τα δεδομένα να μπορέσει να τα διαχειριστεί ο άνθρωπος, ειδικά με την ταχύτητα που παράγονται πλέον. Αντιμετωπίζοντας αυτό το πρόβλημα, έγινε αναζήτηση για έναν τρόπο που θα βοηθήσει τον άνθρωπο σε αυτό το έργο. Η λύση βρέθηκε και ήταν τα έξυπνα υπολογιστικά συστήματα.

Προκειμένου τα υπολογιστικά συστήματα να αποκαλούνται έξυπνα, χρειάζονται ένα μοντέλο. Ένα μοντέλο δημιουργείται μέσω της εκπαίδευσης, δηλαδή με βάση ορισμένα δεδομένα που έχουν συλλεχθεί και σε συνδυασμό με την επιλογή ενός αλγορίθμου Μηχανικής Μάθησης, παράγεται ένα μοντέλο. Έπειτα με την χρήση αυτού του μοντέλου, είναι εφικτή η πρόβλεψη των αποτελεσμάτων και η παραγωγή συμπερασμάτων που έχουν ως σκοπό την βελτιστοποίηση. Όμως, για να είναι εφικτή η έγκυρη πρόβλεψη των αποτελεσμάτων πρέπει πρώτα να ακολουθηθούν ορισμένα βήματα τα οποία θα διασφαλίσουν την εγκυρότητα και την αξιοπιστία των αποτελεσμάτων [70] [71]. Τα βήματα εκπαίδευσης ενός μοντέλου είναι τα εξής:

- Συλλογή των δεδομένων
- Προετοιμασία των δεδομένων
- Επιλογή αλγορίθμου
- Εκπαίδευση του μοντέλου

- Αξιολόγηση του μοντέλου
- Ρύθμιση απόδοσης
- Παραγωγή προβλέψεων

Το πρώτο βήμα είναι η συλλογή των δεδομένων η οποία είναι εξαιρετικά σημαντική. Η ακρίβεια και η αξιοπιστία του μοντέλου που πρόκειται να παραχθεί εξαρτάται από τα δεδομένα από τα οποία εκπαιδεύεται. Επίσης όσα περισσότερα δεδομένα έχουμε, τόσο καλύτερες προβλέψεις θα παράγονται. Επομένως, πρέπει να δοθεί μεγάλη έμφαση στην ποσότητα και την ποιότητα των δεδομένων [70] [71].

Έπειτα ακολουθεί η προετοιμασία των δεδομένων όπου τα δεδομένα τοποθετούνται σε ένα σημείο όπου μπορεί να τα αντλήσει ο αλγόριθμος, για παράδειγμα μία βάση δεδομένων ή σε ένα αρχείο Excel. Αφού τοποθετηθούν τα δεδομένα στο σημείο αυτό, αλλάζουμε την σειρά τους ώστε να αποφύγουμε την επανάληψη κάποιου μοτίβου. Επίσης πρέπει να συλλεχθεί ίδιο πλήθος δεδομένων για κάθε κατηγορία. Για παράδειγμα, για να δημιουργηθεί ένα μοντέλο το οποίο αναγνωρίζει το γιασεμί και το νυχτολούλουδο πρέπει να εισάγουμε στο μοντέλο ίδιο πλήθος φωτογραφιών και για τα δύο είδη λουλουδιών ώστε να μην αναγνωρίζει την μία κατηγορία περισσότερο από την άλλη. Τέλος, τα δεδομένα χωρίζονται σε δυο κατηγορίες, δεδομένα για εκπαίδευση και δεδομένα για αξιολόγηση. Αν χρησιμοποιηθούν όλα τα δεδομένα για την εκπαίδευση του μοντέλου τότε δεν θα είναι εφικτή η αντικειμενική αξιολόγηση του μοντέλου, γι' αυτό τα δεδομένα χωρίζονται στις παραπάνω κατηγορίες με την χρήση μίας αναλογίας (π.χ. 80/20) [70] [71].

Αφού συλλεχθούν και προετοιμαστούν τα δεδομένα, το επόμενο βήμα είναι η επιλογή του αλγορίθμου. Ανάλογα με τον τύπο των δεδομένων που έχουμε επιλέγουμε και τον αλγόριθμο που θα χρησιμοποιηθεί κατά την εκπαίδευση, για παράδειγμα υπάρχουν αλγόριθμοι που παράγουν καλύτερα μοντέλα για αριθμητικά δεδομένα, υπάρχουν αλγόριθμοι που παράγουν καλύτερα μοντέλα όταν έχουμε ως δεδομένα εικόνες κ.ο.κ. Επίσης παίζει σημαντικό ρόλο το πλήθος των χαρακτηριστικών που θα χρησιμοποιηθούν κατά την εκπαίδευση, για παράδειγμα τα πέταλα του γιασεμιού έχουν άσπρο χρώμα ενώ του νυχτολούλουδου έχουν φούξια χρώμα [70] [71].

Κατά την εκπαίδευση του μοντέλου αφού τα δεδομένα έχουν επεξεργαστεί και έχουν διαχωριστεί σε κατηγορίες, δεδομένα για εκπαίδευση και δεδομένα για αξιολόγηση, πλέον ο αλγόριθμος παίρνει τα δεδομένα για εκπαίδευση και μαθαίνει

με βάση τα χαρακτηριστικά του κάθε είδους. Δηλαδή, έχοντας τα χαρακτηριστικά του γιασεμιού και του νυχτολούλουδου ο αλγόριθμος πλέον καταλαβαίνει και μαθαίνει ποια είναι τα χαρακτηριστικά του γιασεμιού και ποια του νυχτολούλουδου οπότε πλέον μπορεί να κάνει τον διαχωρισμό μεταξύ αυτών των δύο ειδών [70] [71].

Μετά την εκπαίδευση του μοντέλου πρέπει να μάθουμε την ακρίβεια του, γι' αυτό και είναι σημαντική η αξιολόγηση του μοντέλου. Για να μάθουμε την ακρίβεια του πρέπει να γίνει χρήση της δεύτερης κατηγορίας δεδομένων, δηλαδή των δεδομένων για αξιολόγηση. Έτσι γίνεται αντιληπτό πόσο καλά έχει εκπαιδευθεί το μοντέλο βάση των δεδομένων για εκπαίδευση και τις τιμές που έχουν δοθεί στις παραμέτρους του αλγορίθμου [70] [71].

Σε περίπτωση που ο αλγόριθμος δεν παράγει τα επιθυμητά αποτελέσματα πρέπει να γίνουν αλλαγές ώστε να αυξηθεί η απόδοση του. Αυτό το βήμα ονομάζεται ρύθμιση απόδοσης. Ουσιαστικά, πρέπει είτε να ελεγχθούν τα δεδομένα εκπαίδευσης ξανά ή να αλλαχθούν οι τιμές των παραμέτρων του αλγορίθμου. Όσο για την δεύτερη περίπτωση, υπάρχουν μεταβλητές οι οποίες ελέγχουν την διαδικασία μάθησης του μοντέλου, για παράδειγμα σε έναν αλγόριθμο συσταδοποίησης επιλέγεται το πλήθος των συστάδων που θα χωριστούν τα δεδομένα. Γενικά αυτές οι παράμετροι επηρεάζουν την διαδικασία μάθησης και η ρύθμιση τους αποτελεί πειραματική διαδικασία καθώς κάθε σύνολο δεδομένων έχει διαφορετικές ανάγκες, οπότε η προσαρμογή είναι ένα πολύ βασικό χαρακτηριστικό για την δημιουργία ακριβών και αξιόπιστων μοντέλων [70] [71] [72] [73].

Τέλος, αφού η απόδοση του μοντέλου είναι υψηλή μένει η παραγωγή προβλέψεων. Σε αυτό το βήμα πλέον ο αλγόριθμος θεωρείται αξιόπιστος οπότε του δίνονται νέα δεδομένα και παράγονται προβλέψεις, δηλαδή πλέον αν δοθεί στο μοντέλο μία φωτογραφία γιασεμιού ως είσοδος τότε το μοντέλο θα αναγνωρίσει ότι είναι γιασεμί, αντίστοιχα θα γίνει το ίδιο και στην περίπτωση που δοθούν ως είσοδοι φωτογραφίες του νυχτολούλουδου [70] [71].

Εν κατακλείδι, παρατηρώντας την διαδικασία εκμάθησης και παραγωγής ενός μοντέλου Μηχανικής Μάθησης, γίνεται αντιληπτό ότι πρέπει να δοθεί ιδιαίτερη προσοχή στην ποιότητα των δεδομένων. Αν δεν έχουμε τα επιθυμητά δεδομένα τότε είναι δύσκολο να παραχθεί ένα αξιόπιστο μοντέλο.

3.2 Μέθοδοι Μηχανικής Μάθησης

Στο συγκεκριμένο κεφάλαιο αναλύονται οι βασικές μέθοδοι μάθησης που μπορούν να ακολουθηθούν, προκειμένου να εκπαιδευθούν τα μοντέλα και να επιλύσουν διάφορα προβλήματα. Καθημερινά συναντώνται διαφόρων ειδών προβλήματα. Ανάλογα το αντικείμενο και τις ανάγκες του προβλήματος πρέπει να βρεθεί ο κατάλληλος τρόπος αντιμετώπισης. Για να λυθεί το πρόβλημα που αντιμετωπίζεται και να παραχθούν τα καλύτερα δυνατά αποτελέσματα, πρέπει το πρόβλημα να κατηγοριοποιηθεί κατάλληλα προκειμένου να επιλεγεί και να εφαρμοστεί ο καταλληλότερος αλγόριθμος Μηχανικής Μάθησης. Υπάρχουν μερικές κατηγορίες στις οποίες μπορούν να τοποθετηθούν όλα τα προβλήματα που συναντώνται. Αφού το πρόβλημα τοποθετηθεί στην κατάλληλη κατηγορία, τότε είναι εύκολο να επιλεγεί και ο βέλτιστος αλγόριθμος Μηχανικής Μάθησης [74]. Τα προβλήματα μπορούν να κατηγοριοποιηθούν στις εξής κατηγορίες:

- Προβλήματα ταξινόμησης (Classification problem)
- Προβλήματα ομαδοποίησης (Clustering problem)
- Προβλήματα ανίχνευσης ανωμαλιών (Anomaly detection problem)
- Προβλήματα παλινδρόμησης (Regression problem)
- Προβλήματα ενίσχυσης (Reinforcement problem)

Αρχικά γίνεται αναφορά στα προβλήματα ταξινόμησης. Σε αυτού του είδους προβλήματα, οι έξοδοι που μπορούν να προκύψουν είναι συγκεκριμένες και γνωστές στο μοντέλο αφού τα δεδομένα είναι κατηγοριοποιημένα. Γι' αυτό και οι έξοδοι που μπορούν να παραχθούν θα είναι σχετικές με κάποια από αυτές τις κατηγορίες. Ένα πρόβλημα ταξινόμησης μπορεί να είναι δυαδικών ή πολλαπλών κατηγοριών [74].

Τα προβλήματα ομαδοποίησης συναντώνται κυρίως στους αλγόριθμους μάθησης χωρίς επίβλεψη. Αυτό συμβαίνει γιατί αρχικά ο αλγόριθμος δέχεται μη επισημασμένα σύνολα δεδομένων, οπότε προσπαθεί να βρει συσχετίσεις μεταξύ των δεδομένων με βάση τις ομοιότητές τους. Έπειτα γίνεται η επισήμανση των διάφορων κατηγοριών που εντάσσονται οι έξοδοι του μοντέλου [74].

Υπάρχουν μοντέλα τα οποία έχουν μάθει και αναλύουν ένα συγκεκριμένο μοτίβο. Σε περίπτωση που υπάρξει αλλαγή ή κάποιου είδους ανωμαλία στο μοτίβο, τότε ο αλγόριθμος εντοπίζει την αλλαγή και την καταγράφει ή ενημερώνει τον διαχειριστή. Αυτή η κατηγορία προβλημάτων ονομάζεται ανίχνευση ανωμαλιών. Για

παράδειγμα, οι εταιρείες πιστωτικών καρτών χρησιμοποιούν αλγόριθμους ανίχνευσης ανωμαλιών, ώστε να ειδοποιούν τον πελάτη τους σε περίπτωση που εντοπιστεί διαφορετική συμπεριφορά συναλλαγών από τη συνήθη [74].

Επίσης υπάρχουν και τα προβλήματα παλινδρόμησης. Οι αλγόριθμοι παλινδρόμησης χρησιμοποιούνται κυρίως για την αντιμετώπιση προβλημάτων που έχουν ως εξόδους αριθμητικές τιμές [74].

Η τελευταία κατηγορία προβλημάτων είναι τα προβλήματα ενίσχυσης. Οι συγκεκριμένοι αλγόριθμοι χρησιμοποιούνται στην περίπτωση όπου μία απόφαση πρέπει να ληφθεί βάση προηγούμενων εμπειριών. Το μοντέλο υιοθετεί την κατάλληλη συμπεριφορά που πρέπει να έχει μέσω της μεθόδου δοκιμής και σφάλματος σε ένα συνεχώς μεταβαλλόμενο περιβάλλον. Ο τρόπος που προγραμματίζεται το μοντέλο είναι μέσω των ανταμοιβών και των κυρώσεων χωρίς όμως να είναι προσδιορισμένος ο τελικός στόχος [74].

Αφού κατανοηθεί σε ποια κατηγορία από τις παραπάνω ανήκει το πρόβλημα, έπειτα πρέπει να επιλεγεί η κατάλληλη μέθοδος εκπαίδευσης Μηχανικής Μάθησης προκειμένου το μοντέλο να εκπαιδευθεί και να επιλύσει το πρόβλημα που αντιμετωπίζεται. Οι βασικότερες μέθοδοι εκπαίδευσης Μηχανικής Μάθησης είναι οι ακόλουθοι:

- Επιβλεπόμενη μάθηση (Supervised learning)
- Μη-επιβλεπόμενη μάθηση (Unsupervised learning)
- Ημι-επιβλεπόμενη μάθηση (Semi supervised learning)
- Ενισχυτική μάθηση (Reinforcement learning)

Παρακάτω αναλύονται και εξηγούνται οι μέθοδοι προκειμένου να κατανοηθούν οι λειτουργίες τους.

Όπως έχει γίνει γνωστό, στόχος της Μηχανικής Μάθησης είναι η μίμηση του τρόπου με τον οποίο οι άνθρωποι επεξεργάζονται τις πληροφορίες τις οποίες δέχονται ώστε να επιτύχουν τον εκάστοτε σκοπό τους. Ένας στόχος που μπορεί να έχει ένα μοντέλο είναι η αναγνώριση προτύπων, όπου σε αυτή την περίπτωση το μοντέλο θα μπορεί να διακρίνει αντικείμενα με βάση τα χαρακτηριστικά τους. Το μοντέλο μαθαίνει πως να τα ξεχωρίζει μέσω της επαναλαμβανόμενης τροφοδότησης εικόνων ή χαρακτηριστικών αυτών των αντικειμένων. Στα δεδομένα τα οποία δέχεται ο αλγόριθμος, έχει τοποθετηθεί μία ετικέτα, δηλαδή είναι κατηγοριοποιημένα ώστε να μπορεί ο αλγόριθμος να καταλάβει τα χαρακτηριστικά κάθε αντικειμένου [75]. Στους αλγόριθμους επιβλεπόμενης μάθησης είναι

απαραίτητη η ανθρώπινη παρέμβαση. Τα δεδομένα εισόδου είναι απαραίτητο να είναι χωρισμένα σε δεδομένα εκπαίδευσης και δεδομένα αξιολόγησης, διότι ένας αλγόριθμος θα εκπαιδευθεί με την χρήση των δεδομένων εκπαίδευσης και έπειτα με βάση τα πρότυπα τα οποία θα έχει μάθει από αυτά, δέχεται ως είσοδο τα δεδομένα αξιολόγησης ώστε να παραχθούν προβλέψεις ή να πραγματοποιηθεί ταξινόμηση [76]. Στην επιβλεπόμενη μάθηση τα μοντέλα μαθαίνουν μέσω παραδειγμάτων. Ένα σύνολο παραδειγμάτων ή αλλιώς ένα σύνολο δεδομένων εκπαίδευσης, τροφοδοτείται στο μοντέλο με τα σωστά αποτελέσματα και με βάση αυτά ο αλγόριθμος μαθαίνει από αυτά. Η ακρίβεια του αυξάνεται επειδή ο αλγόριθμος γνωρίζει τα σωστά αποτελέσματα, γι' αυτό και όταν παράγει αποτελέσματα τα συγκρίνει με τα σωστά και έτσι βελτιστοποιούνται οι προβλέψεις του. Χρησιμοποιείται σε περιπτώσεις όπου έχουμε ένα σύνολο δεδομένων και με βάση αυτά θέλουμε να κάνουμε προβλέψεις. Για παράδειγμα, έχοντας ένα e-shop, είναι εφικτό να τροφοδοτηθεί στο μοντέλο το ιστορικό πλοήγησης ενός χρήστη και με βάση αυτό, να του προταθούν σχετικά προϊόντα με αυτά που είχε ψάξει [74]. Η επιβλεπόμενη μάθηση χρησιμοποιεί συγκεκριμένες τεχνικές κατά την διαδικασία της εκπαίδευσης. Οι πιο βασικές από αυτές είναι οι ακόλουθες:

- Δέντρα απόφασης (Decision tree)
- Μηχανή Διανυσμάτων Υποστήριξης (Support Vector Machine - SVM)
- Naïve Bayes

Τα Δέντρα απόφασης είναι γραφήματα στα οποία οι επιλογές και τα αποτελέσματα αναπαρίστανται σε δένδροειδή μορφή. Κάθε κόμβος του δέντρου αντιπροσωπεύει ένα γεγονός ή μία επιλογή και οι άκρες του αντιπροσωπεύουν τους κανόνες επιλογής ή τις συνθήκες επιλογής. Τα Δέντρα απόφασης απαρτίζονται από κόμβους οι οποίοι αντιπροσωπεύουν χαρακτηριστικά μίας κατηγορίας που πρόκειται να ταξινομηθεί και από κλάδους όπου ο καθένας αντιπροσωπεύει μία τιμή την οποία μπορεί να λάβει ένας κόμβος. [76]. Στατιστικά μέτρα όπως ο δείκτης Gini και η εντροπία υπολογίζουν την αξία κάθε κόμβου [74].

Οι Μηχανές Διανυσματικής Υποστήριξης είναι ένας αλγόριθμος επιβλεπόμενης μάθησης, που μπορεί να χρησιμοποιηθεί για την επίλυση προβλημάτων ταξινόμησης και παλινδρόμησης χρησιμοποιώντας την έννοια του υπολογισμού περιθωρίου (margin calculation). Κάθε στοιχείο των δεδομένων απεικονίζεται γραφικά ως ένα σημείο σε ένα χώρο n -διαστάσεων. Το n είναι το πλήθος των χαρακτηριστικών που έχουμε στην συλλογή των δεδομένων. Κάθε συντεταγμένη

που απεικονίζεται γραφικά, αντιστοιχίζεται σε κάποια τιμή χαρακτηριστικού των δεδομένων της συλλογής. Βρίσκοντας μία γραμμή που ονομάζεται υπερεπίπεδο (hyper plane) ο αλγόριθμος ταξινομεί τα δεδομένα σε διαφορετικές κλάσεις, μεγιστοποιώντας τις αποστάσεις των κοντινότερων σημείων των κλάσεων [74] [76].

Η τεχνική Naïve Bayes πραγματοποιεί ταξινόμηση χρησιμοποιώντας το θεώρημα πιθανοτήτων του Bayes. Το θεώρημα του Bayes ουσιαστικά σχετίζει την τρέχουσα πιθανότητα με την αρχική πιθανότητα, δηλαδή υπολογίζει την μεταγενέστερη πιθανότητα κάποιου γεγονότος A έχοντας δεδομένη μία προηγούμενη πιθανότητα ενός γεγονότος B που αντιπροσωπεύεται από το $P(A/B)$ ως εξής [74]:

$$P(A/B) = \frac{P(B/A) \times P(A)}{P(B)}$$

Ο παραπάνω τύπος χρησιμοποιεί την εξής λογική, η πιθανότητα να συμβεί το ενδεχόμενο A δεδομένου του B, ισούται με την πιθανότητα να συμβεί το B δεδομένου του A επί την πιθανότητα να συμβεί το A, προς την πιθανότητα να συμβεί το B [74, 77]. Η συγκεκριμένη τεχνική χρησιμοποιείται κυρίως για ταξινόμηση κειμένων [76].

Η δεύτερη μέθοδος εκπαίδευσης είναι η μη-εποπτευόμενη μάθηση. Σε αυτή την μέθοδο δεν υπάρχει κάποια καθοδήγηση κατά την διάρκεια της εκπαίδευσης. Χρησιμοποιείται σε περιπτώσεις όπου οι κατηγορίες των δεδομένων είναι άγνωστες. Τα δεδομένα εκπαίδευσης δεν είναι επισημασμένα γι' αυτό χρησιμοποιείται η μέθοδος της μη-εποπτευόμενης μάθησης, για να βρεθούν οι συσχετίσεις μεταξύ των μη επισημασμένων δεδομένων [74] [75] [76]. Όταν εισάγονται νέα δεδομένα, ο αλγόριθμος προσπαθεί με βάση την εμπειρία του από προηγούμενα δεδομένα να κατηγοριοποιήσει τις νέες εισόδους [75] [76]. Οι πιο βασικές τεχνικές που χρησιμοποιούνται στην μη-επιβλεπόμενη μάθηση είναι οι εξής:

- Ομαδοποίηση K-means
- Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis - PCA)

Η τεχνική K-means είναι ένας δημοφιλής τρόπος με τον οποίο μπορεί να λυθεί το πρόβλημα της ομαδοποίησης χρησιμοποιώντας τις συστάδες. Τα δεδομένα ταξινομούνται με την χρήση των κέντρων. Με την έναρξη εκτέλεσης του αλγορίθμου ορίζονται “k” κέντρα. Η τοποθέτηση των κέντρων είναι μία σημαντική διαδικασία καθώς μπορούν να επηρεάσουν τα αποτελέσματα δραματικά. Ενδείκνυται τα κέντρα να τοποθετούνται μακριά το ένα από το άλλο. Αφού τοποθετηθούν τα

κέντρα, τότε τα δεδομένα συσχετίζονται με το πλησιέστερο κέντρο και έτσι δημιουργούνται οι συστάδες. Τα κέντρα αλλάζουν θέση σε κάθε γύρο καθώς υπολογίζεται ο μέσος όρος των δεδομένων που βρίσκονται σε κάθε συστάδα. [74] [76].

Στην Ανάλυση Κύριων Συνιστωσών, χρησιμοποιείται ο ορθογώνιος μετασχηματισμός για να μετατρέψει τις παρατηρήσεις των πιθανά συσχετίσιμων μεταβλητών σε ένα σύνολο τιμών οι οποίες θα είναι γραμμικά μη συσχετίσιμες. Οι συγκεκριμένες μεταβλητές ονομάζονται κύριες συνιστώσες. Δουλειά του αλγόριθμου είναι να εξηγήσει την διακύμανση και την συν-διακύμανση ενός συνόλου μεταβλητών με την χρήση γραμμικών συνδυασμών. Με αυτό τον τρόπο, μειώνονται οι διαστάσεις των δεδομένων. Έτσι απλοποιούνται οι υπολογισμοί και γίνονται πιο γρήγορα [76].

Στην ημι-επιβλεπόμενη μάθηση, συνδυάζονται τα χαρακτηριστικά της επιβλεπόμενης και της μη-επιβλεπόμενης μάθησης. Τα δεδομένα στις προηγούμενες μεθόδους είτε είναι επισημασμένα είτε δεν είναι. Στην συγκεκριμένη μέθοδο, ένα μέρος των δεδομένων είναι επισημασμένο ενώ τα υπόλοιπα δεδομένα δεν είναι επισημασμένα λόγω του υψηλού κόστους που χρειάζεται για να πραγματοποιηθεί η συγκεκριμένη δουλειά καθώς και για την έλλειψη έμπειρων ατόμων [74] [76] [78]. Σε αυτή την περίπτωση, τα επισημασμένα δεδομένα χρησιμοποιούνται για να βοηθήσουν στην διαδικασία εκπαίδευσης και ταξινόμησης των μη επισημασμένων δεδομένων [75] [78]. Παρακάτω, παρουσιάζονται μερικές τεχνικές ημι-επιβλεπόμενης μάθησης. Μερικές από αυτές είναι οι εξής:

- Μεταγωγική Μηχανή Διανυσμάτων Υποστήριξης (Transductive Support Vector Machine - TSVM)
- Παραγωγικά Μοντέλα (Generative Models - GM)
- Αυτό-εκπαίδευση (Self-Training)

Οι Μεταγωγικές Μηχανές Διανυσμάτων Υποστήριξης χρησιμοποιούνται όταν υπάρχουν μη επισημασμένα δεδομένα στα σετ δεδομένων. Με την χρήση αυτής της τεχνικής τα δεδομένα επισημαίνονται έτσι ώστε να υπάρχει διαφορά συγκριτικά με τα μη επισημασμένα δεδομένα [76].

Ακολουθούν τα Παραγωγικά Μοντέλα, όπου σε αυτού τους αλγόριθμους είναι εφικτή η παραγωγή δεδομένων. Μοντελοποιώντας το $P(x, y)$ μπορεί να χρησιμοποιηθεί αυτή η κατανομή πιθανοτήτων για να δημιουργηθούν σημεία δεδομένων [76].

Τέλος, στην Αυτό-εκπαίδευση χρησιμοποιείται ένας ταξινομητής ο οποίος εκπαιδεύεται με την χρήση ενός μικρού πλήθους επισημασμένων δεδομένων. Αφού εκπαιδευθεί με τα επισημασμένα δεδομένα, τροφοδοτείται με τα μη επισημασμένα δεδομένα. Έπειτα, οι ετικέτες που προβλέφθηκαν καθώς και τα μη επισημασμένα σημεία, προστίθενται στο σύνολο των δεδομένων εκπαίδευσης και η παραπάνω διαδικασία επαναλαμβάνεται έχοντας ως στόχο την βελτιστοποίηση [76].

Η τελευταία μέθοδος μάθησης είναι η Ενισχυτική Μάθηση. Θεωρείται μία ενδιάμεση μέθοδος καθώς στον αλγόριθμο παρέχονται μόνο απαντήσεις που του υποδεικνύουν αν η έξοδος είναι σωστή ή όχι. Η κύρια δουλειά του αλγορίθμου είναι η απόκλιση και έγκριση διάφορων πιθανοτήτων προκειμένου να παραχθεί το επιθυμητό αποτέλεσμα [74] [76]. Ο αλγόριθμος αντιστοιχίζει διάφορες πιθανές περιπτώσεις με συγκεκριμένες ενέργειες. Δεν του υποδεικνύονται ποιες ενέργειες πρέπει να κάνει, ανακαλύπτει ποιες ενέργειες πρέπει να κάνει παρατηρώντας ποιες από αυτές έχουν την μεγαλύτερη αποδοχή σε μία συγκεκριμένη περίπτωση. Λειτουργεί με την χρήση ενός συστήματος ανταμοιβής [79]. Συναντώνται σημαντικές διαφορές ανάμεσα στην Ενισχυτική Μάθηση και στις υπόλοιπες μεθόδους που αναλύθηκαν προηγουμένως. Συγκρίνοντας την Ενισχυτική Μάθηση με την Επιβλεπόμενη Μάθηση, παρατηρούμε ότι στην Επιβλεπόμενη Μάθηση ένα μοντέλο δεν μπορεί να μάθει εύκολα σε προβλήματα που είναι αναγκαία η αλληλεπίδραση. Επίσης η Ενισχυτική Μάθηση δεν χρησιμοποιείται για ανακάλυψη προτύπων όπως η Μη-επιβλεπόμενη Μάθηση, αλλά προσπαθεί να μεγιστοποιήσει το σήμα επιβράβευσης.

Έχοντας αναλύσει τις βασικές μεθόδους Μηχανικής Μάθησης, πλέον είναι εφικτή η μετάβαση στο πρακτικό κομμάτι της διπλωματικής. Στο επόμενο κεφάλαιο παρουσιάζεται η διαδικασία υλοποίησης του μοντέλου Μηχανικής Μάθησης για την βελτιστοποίηση της ελεύθερης βολής.

ΚΕΦΑΛΑΙΟ 4

ΕΦΑΡΜΟΓΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΑΘΛΗΤΙΚΩΝ ΕΠΙΔΩΣΕΩΝ

4.1 Παρουσίαση εφαρμογής

Σε αυτό το κεφάλαιο θα αναλυθούν τα βήματα που ακολουθήθηκαν για την υλοποίηση της εφαρμογής βελτιστοποίησης αθλητικών επιδόσεων, καθώς και οι τεχνολογίες που χρησιμοποιήθηκαν. Στο τέλος, θα παρουσιαστούν τα αποτελέσματα και διάφορες παρατηρήσεις. Η εφαρμογή αποτελείται από την Όραση Υπολογιστών και την Μηχανική Μάθηση. Αρχικά θα παρουσιαστεί μία από τις δυνατότητες της Όρασης Υπολογιστών, ο εντοπισμός της στάσης σώματος του ανθρώπου, που βοηθά στην διαδικασία παραγωγής των δεδομένων.

Η Όραση Υπολογιστών παρέχει την λειτουργία του εντοπισμού της στάσης σώματος του ανθρώπου. Είναι μία σχετικά νέα τεχνική η οποία έχει ξεκινήσει να χρησιμοποιείται και να αναπτύσσεται τα τελευταία χρόνια. Ουσιαστικά, η συγκεκριμένη τεχνολογία είναι ικανή να εντοπίζει τα κύρια σημεία του ανθρώπινου σώματος όπως τους ώμους, τους αγκώνες, την λεκάνη, τα γόνατα κλπ. Έχοντας αυτή την τεχνολογία, μπορούν να δημιουργηθούν διάφορες χρήσιμες εφαρμογές που θα μας βοηθήσουν στην καθημερινότητα μας [89]. Για παράδειγμα έχουν δημιουργηθεί εφαρμογές γυμναστικής, οι οποίες είναι εκπαιδευμένες να αναγνωρίζουν διαφόρων ειδών ασκήσεις, ώστε όταν κάποια άσκηση δεν εκτελείται σωστά τότε η εφαρμογή κάνει υποδείξεις στον αθλητή, προκειμένου να τον βοηθήσει και να αποτρέψει τυχόν τραυματισμούς [89]. Επίσης έχουν δημιουργηθεί εφαρμογές, όπου εντοπίζεται η κίνηση του ανθρώπου κατά την διάρκεια που τρέχει, εντοπίζεται η μπάλα και το σερβίς στο τένις [70]. Όπως γίνεται αντιληπτό, με τον εντοπισμό της στάσης του ανθρώπινου σώματος μπορούν να δημιουργηθούν πολλές χρήσιμες εφαρμογές οι οποίες θα βοηθούν τον άνθρωπο σε καθημερινές δραστηριότητες και όχι μόνο.

Η δεύτερη τεχνολογία η οποία χρησιμοποιείται στην εφαρμογή της βελτιστοποίησης της αθλητικής απόδοσης, είναι η Μηχανική Μάθηση που αναλύθηκε σε προηγούμενο κεφάλαιο. Ουσιαστικά θα χρησιμοποιηθεί η μη-επιβλεπόμενη μάθηση για την ανάλυση των δεδομένων και την παραγωγή

αποτελεσμάτων και συμπερασμάτων. Με τον συνδυασμό και των δύο κατηγοριών, παράγεται μία εφαρμογή η οποία αρχικά εντοπίζει την στάση σώματος του ανθρώπου που ρίχνει βολή μπάσκετ και στην συνέχεια αναλύει τις τιμές των σημείων ενδιαφέροντος, που είναι ο δεξής ώμος, αγκώνας και καρπός, προκειμένου να παραχθούν συμπεράσματα όσων αφορά την ομοιότητα της βολής του παίκτη Α και του παίκτη Β. Επίσης, παράγεται και ένα ποσοστό ομοιότητας για τους δύο παίκτες, προκειμένου να δει ο αθλητής που θέλει να βελτιστοποιήσει την βολή του, σε τι επίπεδο βρίσκεται. Στην **Εικόνα 8** φαίνεται ένα στιγμιότυπο από τα βίντεο εκπαίδευσης, στο οποίο έχει γίνει εντοπισμός της ανθρώπινης στάσης.



Εικόνα 8: Εντοπισμός της στάσης του ανθρώπινου σώματος

4.2 Βιβλιοθήκες και εξαρτήματα που χρησιμοποιήθηκαν

Προκειμένου να υλοποιηθεί η εφαρμογή και να παραχθούν τα επιθυμητά αποτελέσματα, ήταν αναγκαία η χρήση διάφορων πακέτων (packages), βιβλιοθηκών (libraries) και εξαρτημάτων. Σε αυτό το υποκεφάλαιο, αναφέρονται και εξηγούνται οι βιβλιοθήκες, τα πακέτα και οποιοδήποτε μέσο χρησιμοποιήθηκε για την εκπόνηση αυτού του έργου.

4.2.1 Βιβλιοθήκη MediaPipe

Για να εκπαιδευθεί το μοντέλο, πρέπει να υπάρχει ένα σύνολο δεδομένων, με τις συντεταγμένες του δεξιού ώμου, αγκώνα και καρπού, κατά την ρίψη ελεύθερων βολών. Από τα βίντεο του αθλητή, πρέπει να συλλεχθούν τα δεδομένα και να δημιουργηθεί το σύνολο των δεδομένων. Για τον συγκεκριμένο σκοπό χρησιμοποιείται το MediaPipe.



Εικόνα 9: Λογότυπο του MediaPipe [80]

Το MediaPipe είναι ένα framework ανοιχτού κώδικα, που έχει δημιουργηθεί από την Google. Παρέχει έτοιμες λύσεις για διάφορα ζητήματα Μηχανικής Μάθησης και είναι συμβατό με πολλαπλές πλατφόρμες (cross platform). Είναι ένα προσαρμόσιμο framework που στοχεύει στην εύκολη ενσωμάτωση του στις εφαρμογές. Παρέχει έτοιμο κώδικα για πολλά σύνθετα προβλήματα [81]. Μερικές από αυτές τις υπηρεσίες που παρέχει είναι οι ακόλουθες [82]:

- Εντοπισμός προσώπου (Face detection)
- Εντοπισμός παλάμης (Palm detection)
- Εντοπισμός του ανθρώπινου σώματος (Pose detection)
- Εντοπισμός αντικειμένων (Object detection)
- Εντοπισμός της ίριδας (Iris detection)

Η υπηρεσία που παρέχει το MediaPipe και χρησιμοποιείται για την υλοποίηση της παρούσας διπλωματικής, είναι αυτή του εντοπισμού του ανθρώπινου σώματος. Το MediaPipe παρέχει το Pose Landmark Model το οποίο εντοπίζει την θέση των κυριότερων σημείων του ανθρώπινου σώματος [83].

4.2.2 Βιβλιοθήκη scikit-learn

Η scikit-learn είναι μία βιβλιοθήκη ανοιχτού κώδικα η οποία παρέχει συναρτήσεις για την υλοποίηση επιβλεπόμενης και μη επιβλεπόμενης εκπαίδευσης. Είναι από τις πιο γνωστές και ίσως πιο χρήσιμες βιβλιοθήκες της Python, για την επίλυση προβλημάτων Μηχανική Μάθησης [84].



Εικόνα 10: Λογότυπο του *scikit-learn* [84]

Παρέχει πολλές λειτουργίες για την επεξεργασία των δεδομένων, αρκετές έτοιμες συναρτήσεις για την εκπαίδευση μοντέλων και πολλά άλλα tools [84].

4.2.3 Βιβλιοθήκη pandas

Το pandas είναι ένα πακέτο της Python που χρησιμοποιείται για γρήγορη και ευέλικτη επεξεργασία επισημασμένων και μη επισημασμένων δεδομένων. Είναι ένα από τα πιο δυνατά πακέτα της Python για ανάλυση δεδομένων, καθώς η επεξεργασία των δεδομένων γίνεται εύκολα με την χρήση του pandas.



Εικόνα 11: Λογότυπο του *pandas* [85]

Παρέχει πολλές λειτουργίες για την επεξεργασία των δεδομένων όπως καθορισμός των δεδομένων, γέμισμα των δεδομένων, κανονικοποίηση των δεδομένων, συγχώνευση των δεδομένων και πολλά άλλα.

4.2.4 Βιβλιοθήκη OpenCV

Η OpenCV (Open Source Computer Vision Library) είναι μία βιβλιοθήκη ανοιχτού κώδικα για μηχανική μάθηση και όραση υπολογιστών. Η συγκεκριμένη βιβλιοθήκη περιέχει μία μεγάλη γκάμα κλασικών και υπερσύγχρονων αλγορίθμων μηχανικής μάθησης και όρασης υπολογιστών.



Εικόνα 12: Λογότυπο του OpenCV [86]

Μερικές από τις εφαρμογές που μπορούν να χρησιμοποιηθούν οι αλγόριθμοι είναι:

- Αναγνώριση αντικειμένων
- Παρακολούθηση κινήσεων έχοντας μία κάμερα
- Εύρεση όμοιων εικόνων από ένα σύνολο δεδομένων
- Εξαγωγή τρισδιάστατων μοντέλων (3D) για διάφορα αντικείμενα

4.2.4 Βιβλιοθήκη matplotlib

Η βιβλιοθήκη matplotlib χρησιμοποιείται για την δημιουργία στατικών, δυναμικών και αλληλεπιδραστικών διαγραμμάτων. Παρέχει μία μεγάλη γκάμα διαγραμμάτων στους χρήστες, τα οποία μπορούν να τα επεξεργαστούν ώστε να έχουν τα χρώματα και τις ετικέτες που επιθυμούν. Θα μπορούσε να πει κάποιος ότι είναι μία εναλλακτική του MATLAB.



Εικόνα 13: Λογότυπο της βιβλιοθήκης matplotlib

4.2.5 Η Βιβλιοθήκη xlwt και το module os

Η βιβλιοθήκη xlwt χρησιμοποιείται για την δημιουργία υπολογιστικών φύλλων (spreadsheets). Με την βοήθεια της είναι εφικτή η εγγραφή και η ανάγνωση δεδομένων σε Excel παλιάς μορφής (.xls) έχοντας ως σκοπό την αποθήκευση των δεδομένων για περαιτέρω επεξεργασία [87]. Επίσης με την χρήση του module os, γίνεται εφικτή η διαχείριση των φακέλων. Ένας προγραμματιστής μπορεί με την χρήση αυτού του module, να δημιουργεί και να αφαιρεί φακέλους, να αλλάζει τον

κατάλογο που βρίσκεται, να βρίσκει άμεσα σε ποιο φάκελο είναι και να μπορεί να προσπελάσει τα περιεχόμενα κάθε φακέλου [88].

4.2.6 Εξοπλισμός και μέσα που χρησιμοποιήθηκαν

Δεδομένου ότι θέλουμε να κάνουμε εγγραφή βίντεο για αθλητές του μπάσκετ, χρειάζονται τα εξής:

- Μπάλα μπάσκετ σωστών διαστάσεων
- Κάμερα ή κινητό τηλέφωνο για την εγγραφή των βίντεο

Τα μεγέθη της μπάλας μπάσκετ αλλάζουν, ανάλογα το φύλο του αθλητή. Υπάρχουν οι παιδικές μπάλες μεγέθους 5, οι μπάλες για τις αθλήτριες μεγέθους 6 και οι μπάλες για τους αθλητές μεγέθους 7. Προκειμένου να ληφθούν καλύτερα αποτελέσματα, ανάλογα το φύλο και την ηλικία του αθλητή, πρέπει να επιλεγεί το σωστό μέγεθος μπάλας. Αλλάζοντας το μέγεθος της μπάλας επηρεάζεται το αποτέλεσμα, διότι ο αθλητής πρέπει να χρησιμοποιήσει διαφορετική δύναμη συγκριτικά με αυτή που ίσως θα έχει συνηθίσει. Γι' αυτό προτείνεται να χρησιμοποιηθεί το σωστό μέγεθος.

Όσων αφορά το μέσο λήψης των βίντεο, χρησιμοποιήθηκε μία φωτογραφική μηχανή Nikon DSLR D3500 με φακό 18 - 55 mm. Χαρακτηρίζεται ως μία μέτρια φωτογραφική μηχανή για βασική χρήση. Επίσης και ο φακός είναι απλός για βασική χρήση και όχι ιδιαίτερα ακριβός. Γενικά χρησιμοποιήθηκαν εξαρτήματα τα οποία είναι πιθανό ή μπορούν να αγοραστούν από τους περισσότερους (mid-range).

4.3 Διαδικασία υλοποίησης της εφαρμογής

Προκειμένου να υλοποιηθεί η εφαρμογή, έπρεπε να ακολουθηθούν συγκεκριμένα βήματα, τα οποία είναι αλληλένδετα μεταξύ τους και γι' αυτό πρέπει να εκτελεστούν με κατάλληλη σειρά. Τα βήματα που ακολουθήθηκαν, είναι αυτά που παρουσιάστηκαν στο κεφάλαιο 3.1 *Τι είναι η Μηχανική Μάθηση*. Ουσιαστικά τα βήματα που παρουσιάστηκαν θεωρητικά, εφαρμόζονται πρακτικά, καθώς αποτελούν την βασική διαδικασία εκπαίδευσης και παραγωγής μοντέλων Μηχανικής Μάθησης. Στα επόμενα υποκεφάλαια εξηγείται για κάθε βήμα, οι ενέργειες που έγιναν, ο κώδικας που δημιουργήθηκε και τα αποτελέσματα που παράχθηκαν.

4.3.1 Συλλογή των δεδομένων

Κατά την συλλογή των δεδομένων, έπρεπε να δοθεί μεγάλη σημασία στον τρόπο λήψης των βίντεο. Αρχικά η απόσταση και η γωνία λήψης των βίντεο έπρεπε να προσεχθούν αρκετά, καθώς αν άλλαζαν, τότε τα δεδομένα δεν θα ήταν σχετικά μεταξύ τους διότι θα άλλαζε η κλίμακα. Άρα η φωτογραφική μηχανή ή γενικά το μέσο λήψης, πρέπει να είναι σταθερό και αμετάβλητο. Αυτό επιτυγχάνεται με την χρήση τριπόδου ή ο άνθρωπος που καταγράφει το βίντεο, να παραμένει σταθερός στην ίδια θέση καθ' όλη την διάρκεια της συλλογής των δεδομένων. Τα ίδια περίπου ισχύουν και για τον αθλητή ο οποίος ρίχνει τις βολές. Πρέπει να ρίχνει τις βολές από την ίδια θέση κάθε φορά προκειμένου η κλίμακα να μην αλλάξει.

Τα βίντεο λήφθηκαν πρωινή και ηλιόλουστη μέρα, προκειμένου το μοντέλο Όρασης Υπολογιστών MediaPipe να είναι πιο ακριβές. Αυτή η απόφαση πάρθηκε, συλλογιζόμενος την αδυναμία που αντιμετωπίζουν οι περισσότερες κάμερες με τις βραδινές λήψεις, οι οποίες χάνουν σημαντικές πληροφορίες λόγω της έλλειψης φωτός. Επίσης, είναι σημαντικό να σημειωθεί ότι και οι 3 αθλητές ρίχνουν βολή με το δεξί χέρι.

Στην συνέχεια παρουσιάζεται το αρχείο Python με όνομα `human_pose_estimation_MediaPipe.py`, που εξάγει τα δεδομένα από τα βίντεο και τα εκχωρεί σε ένα αρχείο Excel. Στην συνέχεια εξηγείται η λογική του και παρουσιάζονται τμήματα του.

Τα κύρια σημεία του κώδικα είναι η δημιουργία των Excel και οι συναρτήσεις που εξάγουν τις συντεταγμένες του δεξιού ώμου, αγκώνα και καρπού και τις εκχωρούν στα κατάλληλα Excel. Πρώτα δημιουργούνται τα Excel και δημιουργείται ένα φύλλο (Sheet 1) για κάθε ένα από αυτά. Στην συνέχεια ορίζονται τα ονόματα των στηλών ώστε να γνωρίζουμε αργότερα τι περιέχει κάθε στήλη των Excel.

```
style = xlwt.easyxf('font: bold 1') # Specifying style for the .csv files
dataset_xls = xlwt.Workbook() # Workbook for training set
real_world_set_A_xls = xlwt.Workbook() # Workbook for test set of athlete A
real_world_set_B_xls = xlwt.Workbook() # Workbook for test set of athlete B

sheet_dataset = dataset_xls.add_sheet('Sheet 1') # add_sheet is used to
create sheet in .csv
sheet_real_world_set_A = real_world_set_A_xls.add_sheet('Sheet 1')
sheet_real_world_set_B = real_world_set_B_xls.add_sheet('Sheet 1')

sheet_dataset.write(0, 0, 'Right shoulder x', style) # adding names to the
columns of training_set_csv
sheet_dataset.write(0, 1, 'Right shoulder y', style) # to recognize the data
that we have
sheet_dataset.write(0, 2, 'Right elbow x', style)
sheet_dataset.write(0, 3, 'Right elbow y', style)
sheet_dataset.write(0, 4, 'Right wrist x', style)
sheet_dataset.write(0, 5, 'Right wrist y', style)
```

Εικόνα 14: Δημιουργία των Excel όπου θα εκχωρηθούν οι συντεταγμένες

Αφού δημιουργηθούν τα Excel και δοθεί όνομα στις στήλες τους, έπειτα πρέπει να διευκρινιστούν οι τοποθεσίες των αρχείων (paths) που βρίσκονται τα βίντεο, για να μπορέσουν οι συναρτήσεις να τα επεξεργαστούν.

```
training_set_folder = "D:/Diploma-Thesis/thesis_ML-CV/dataset/"
test_set_A_folder = "D:/Diploma-Thesis/thesis_ML-CV/real_world_set_A/"
test_set_B_folder = "D:/Diploma-Thesis/thesis_ML-CV/real_world_set_B/"

xls_name_training = "dataset.xls"
xls_name_test_A = "real_world_set_A.xls"
xls_name_test_B = "real_world_set_B.xls"
```

Εικόνα 15: Μεταβλητές που περιέχουν το path των βίντεο

Αφού γίνουν όλα αυτά, έπειτα καλούνται οι ακόλουθες συναρτήσεις. Η συνάρτηση `multiple_videos_find_landmarks_human_pose_estimation`, είναι η συνάρτηση η οποία δέχεται πλήθος από βίντεο, τα αναλύει και εκχωρεί τις τιμές στα Excel. Ενώ η συνάρτηση `single_video_find_landmarks_human_pose_estimation`, δέχεται ένα βίντεο την φορά και δημιουργεί ένα Excel με τις τιμές που εντοπίζει.

```
multiple_videos_find_landmarks_human_pose_estimation(sheet_dataset,
dataset_xls, training_set_folder, xls_name_training)

single_video_find_landmarks_human_pose_estimation(sheet_real_world_set_A,
real_world_set_A_xls, test_set_A_folder, xls_name_test_A)
single_video_find_landmarks_human_pose_estimation(sheet_real_world_set_B,
real_world_set_B_xls, test_set_B_folder, xls_name_test_B)
```

Εικόνα 16: Συναρτήσεις για διαχείριση ενός και πολλών βίντεο

Το κύριο σώμα των συναρτήσεων έχει ενδιαφέρον. Αφού εντοπιστούν τα σημεία του σώματος τότε ξεχωρίζονται οι συντεταγμένες του δεξιού ώμου, αγκώνα και

καρπού. Έπειτα κάθε τιμή, ανάλογα ποιο μέρος του χεριού είναι, εκχωρείται στις κατάλληλες στήλες του Excel.

```
if results.pose_landmarks:
    mpDraw.draw_landmarks(img, results.pose_landmarks,
                           mpPose.POSE_CONNECTIONS)
    for id, lm in enumerate(results.pose_landmarks.landmark):
        h, w, c = img.shape # get image's width, height and channel
        cx, cy = int(lm.x * w), int(lm.y * h) # convert normalized values to
        real pixel values
        cv2.circle(img, (cx, cy), 6, (0, 0, 255), cv2.FILLED) # print a red
        filled circle on the detected joint

        if id == 12: # right shoulder -> insert it at the right column and
        row in the .csv
            sheet_set_csv.write(rows12 + 1, 0, cx)
            sheet_set_csv.write(rows12 + 1, 1, cy)
            rows12 += 1
        elif id == 14: # right elbow -> insert it at the right column and
        row in the .csv
            sheet_set_csv.write(rows14 + 1, 2, cx)
            sheet_set_csv.write(rows14 + 1, 3, cy)
            rows14 += 1
        elif id == 16: # right wrist -> insert it at the right column and
        row in the .csv
            sheet_set_csv.write(rows16 + 1, 4, cx)
            sheet_set_csv.write(rows16 + 1, 5, cy)
            rows16 += 1
    set_csv.save(name_csv)
```

Εικόνα 17: Αποθήκευση των συντεταγμένων στα Excel

Για να σχεδιαστούν οι γραμμές και οι κουκίδες στα μέρη του σώματος όπως βλέπουμε στα παραπάνω στιγμιότυπα και για να εντοπιστούν τα σημεία του σώματος του ανθρώπου, ήταν αναγκαία η χρήση των παρακάτω συναρτήσεων που παρέχονται από την βιβλιοθήκη MediaPipe για τους συγκεκριμένους σκοπούς.

```
mpPose = mp.solutions.pose
pose = mpPose.Pose()
mpDraw = mp.solutions.drawing_utils
```

Εικόνα 18: Έτοιμες συναρτήσεις της βιβλιοθήκης MediaPipe για το pose detection

4.3.2 Προετοιμασία των δεδομένων

Κατά την προετοιμασία των δεδομένων, τα βίντεο χωρίστηκαν στους φακέλους dataset, real_world_set_A και real_world_set_B. Ο φάκελος dataset περιέχει τα βίντεο μου, που ουσιαστικά με αυτά τα δεδομένα βασίζεται ο αλγόριθμος και με βάση αυτά τα δεδομένα θα ελεγχθεί κατά πόσο είναι σωστός ο τρόπος βολής του αθλητή. Άρα μπορούμε να πούμε ότι ο συγκεκριμένος φάκελος περιέχει το σύνολο εκπαίδευσης (training set) ενώ άλλοι δύο φάκελοι περιέχουν τα βίντεο των αθλητών που θέλουν να βελτιστοποιήσουν την ελεύθερη βολή τους, άρα αυτοί οι φάκελοι

είναι τα σύνολα δοκιμής (test set). Αφού εκχωρήθηκαν τα δεδομένα στα Excel, έπειτα ελέγχθηκε αν υπάρχουν διπλότυπα στις τιμές. Δεν θέλουμε να υπάρχουν ίδιες τιμές στο σύνολο εκπαίδευσης, ώστε το μοντέλο να μην είναι προκατειλημμένο (biased).

4.3.3 Επιλογή αλγορίθμου

Ο αλγόριθμος που θα χρησιμοποιηθεί στην εφαρμογή είναι ο K-means. Ο συγκεκριμένος αλγόριθμος αναλύθηκε και προηγουμένως στο κεφάλαιο 3.2 Μέθοδοι Μηχανικής Μάθησης, είναι ένας αλγόριθμος που ανάλογα το πλήθος των συστάδων, ορίζονται αυθαίρετα ορισμένα κέντρα. Με βάση αυτά τα κέντρα χρησιμοποιούνται μετρικές απόστασης, όπως είναι η Ευκλείδεια απόσταση, ώστε να δημιουργηθούν οι επιθυμητές συστάδες. Χρησιμοποιείται ο αλγόριθμος K-means διότι σχετικά με άλλους αλγορίθμους είναι προσαρμόσιμος σε μεγάλη μάζα δεδομένων και αυτό είναι ένα χαρακτηριστικό το οποίο απαιτεί η εφαρμογή καθώς πρέπει να διαχειριστούμε ένα αρκετά μεγάλο πλήθος συντεταγμένων x και y του δεξιού ώμου, αγκώνα και καρπού. Ένα ακόμη χαρακτηριστικό που με έκανε να επιλέξω τον συγκεκριμένο αλγόριθμο είναι η προσαρμοστικότητα του σε νέα δεδομένα, καθώς με την χρήση της μεθόδου Elbow και με το Silhouette score μπορούμε εύκολα κάθε φορά, να βρίσκουμε το πλήθος συστάδων που θα μας παράξουν τα καλύτερα δυνατά αποτελέσματα ανάλογα με το σύνολο δεδομένων που έχουμε κάθε φορά.

4.3.4 Εκπαίδευση του μοντέλου

Για να εκπαιδευθεί το μοντέλο, του τροφοδοτούμε τα δεδομένα που έχουν συλλεχθεί στα Excel. Συγκεκριμένα στην περίπτωση μας, τροφοδοτούνται τα δεδομένα των αθλητών, προκειμένου να αναλυθούν με την μέθοδο της συσταδοποίησης, η κίνηση που διανύουν τα σημεία ενδιαφέροντος δηλαδή ο δεξιός ώμος, αγκώνας και καρπός του αθλητή. Παράγονται κατάλληλα διαγράμματα, προκειμένου να μπορέσουμε να δούμε τα αποτελέσματα της συσταδοποίησης. Παρακάτω φαίνεται το σημείο του κώδικα που υλοποιεί την συσταδοποίηση.

```
def find_clusters(dataset):
    km = KMeans(n_clusters=3)
    y_predicted = km.fit_predict(datasetB)
    dataset['cluster'] = y_predicted

    df1 = dataset[dataset.cluster == 0]
    df2 = dataset[dataset.cluster == 1]
    df3 = dataset[dataset.cluster == 2]

    plt.scatter(df1["Right shoulder x"], df1["Right shoulder y"],
color='green')
    plt.scatter(df2["Right shoulder x"], df2["Right shoulder y"],
color='red')
    plt.scatter(df3["Right shoulder x"], df3["Right shoulder y"],
color='black')
    plt.xlabel('x')
    plt.ylabel('y')
    plt.legend()
    plt.show()
```

Εικόνα 19: Συνάρτηση που εφαρμόζει τον k-means

Ο κώδικας που χρησιμοποιείται για να εκτυπωθούν τα διαγράμματα των βολών είναι ο ακόλουθος.

```
def show_plots(dataset):
    plt.scatter(dataset["Right shoulder x"], dataset["Right shoulder y"]) #
plot the datasetA
    plt.xlabel('x')
    plt.ylabel('y')
    plt.show()

    plt.scatter(dataset["Right elbow x"], dataset["Right elbow y"]) # plot
the datasetA
    plt.xlabel('x')
    plt.ylabel('y')
    plt.show()

    plt.scatter(dataset["Right wrist x"], dataset["Right wrist y"]) # plot
the datasetA
    plt.xlabel('x')
    plt.ylabel('y')
    plt.show()
```

Εικόνα 20: Συνάρτηση εκτύπωσης των συντεταγμένων

4.3.5 Αξιολόγηση του μοντέλου

Στην μη επιβλεπόμενη μάθηση, δεν μπορεί να ελεγχθεί η ακρίβεια του μοντέλου με τον ίδιο τρόπο που ελέγχεται στους αλγορίθμους ταξινόμησης. Με τους αλγορίθμους μη επιβλεπόμενης μάθησης, πραγματοποιείται συσταδοποίηση, δηλαδή εφαρμόζονται μετρικές απόστασης προκειμένου να δημιουργηθούν συστάδες μεταξύ των δεδομένων. Στην συγκεκριμένη περίπτωση θα χρησιμοποιηθεί ο αλγόριθμος συσταδοποίησης K-means. Πρέπει να δοθεί μεγάλη

προσοχή στο πλήθος των συστάδων που θα επιλεγεί, καθώς είναι ένα από τα σημαντικότερα σημεία τα οποία επηρεάζουν πολύ τα αποτελέσματα του αλγορίθμου αυτού. Για να μπορέσουμε να αξιολογήσουμε αν το πλήθος των συστάδων που έχει επιλεγεί παράγει τα βέλτιστα αποτελέσματα, τότε θα χρησιμοποιηθεί ο συντελεστής Silhouette. Ο συντελεστής Silhouette χρησιμοποιείται για να υπολογιστεί, πόσο καλά έχουν συσταδοποιηθεί τα δεδομένα με την χρήση μίας τεχνικής συσταδοποίησης [90]. Τα αποτελέσματα του συντελεστή κυμαίνονται από 0 έως 1. Όσο το αποτέλεσμα πλησιάζει το 1, τότε ο αλγόριθμος συσταδοποίησης έχει χωρίσει τα δεδομένα επαρκώς αρκετά καλά. Αν το αποτέλεσμα του συντελεστή πλησιάζει το 0 τότε τα δεδομένα δεν έχουν χωριστεί καλά. Παρακάτω φαίνεται ο τρόπος με τον οποίο κωδικοποιείται και χρησιμοποιείται στην εφαρμογή ο συντελεστής Silhouette.

```
# calculate silhouette score of the dataset
for i in range(2, 10):
    kmeans = KMeans(n_clusters=i, max_iter=100)
    kmeans.fit(dataset)
    score = silhouette_score(dataset, kmeans.labels_)
    print("For cluster: {}, the silhouette score is: {}".format(i, score))

silhouette_coefficients = []
for i in range(2, 10):
    kmeans = KMeans(n_clusters=i, max_iter=100)
    kmeans.fit(dataset)
    score = silhouette_score(dataset, kmeans.labels_)
    silhouette_coefficients.append(score)

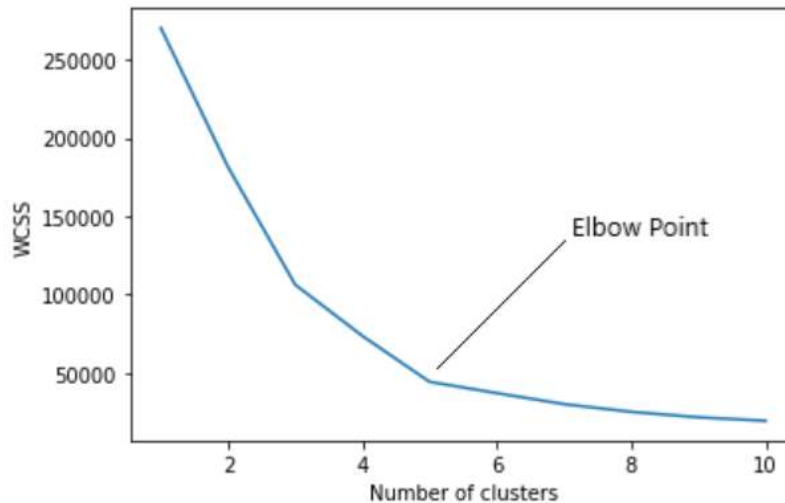
plt.plot(range(2, 10), silhouette_coefficients)
plt.xticks(range(2, 10))
plt.xlabel("number of clusters")
plt.ylabel("Silhouette coefficient")
plt.show()
```

Εικόνα 21: Υπολογισμός του Silhouette score

4.3.6 Ρύθμιση απόδοσης

Ο πρώτος τρόπος με τον οποίο μπορούμε να αξιολογήσουμε την επίδοση του μοντέλου, είναι με την χρήση του συντελεστή Silhouette. Μπορούμε όμως βελτιώσουμε την απόδοση του μοντέλου με την χρήση ακόμα μίας μεθόδου που ονομάζεται μέθοδος Elbow (αγκώνας). Με την μέθοδο Elbow ουσιαστικά προσπαθούμε αλλάζουμε το πλήθος των συστάδων μεταξύ 1 έως 10 και για κάθε πλήθος συστάδων που επιλέγεται, υπολογίζεται το WCSS (Within-Cluster Sum of Square) [91]. Το WCSS είναι το τετραγωνισμένο άθροισμα των αποστάσεων

μεταξύ κάθε σημείου με το κέντρο σε μία συστάδα. Όταν παρουσιάζουμε τα αποτελέσματα των αθροισμάτων, σχηματίζεται μια γραμμή η οποία κάποια στιγμή θα κάνει μία απότομη κλίση και θα σχηματίσει μία γωνία (εξού και αγκώνας) [91].



Εικόνα 22: Παράδειγμα αποτελεσμάτων της μεθόδου Elbow [91]

Στο σημείο που θα υπάρξει απότομη αλλαγή στα αποτελέσματα των αθροισμάτων, εκεί θα είναι και αριθμός των συστάδων που βελτιστοποιούν τα αποτελέσματα για το σύνολο δεδομένων μας. Παρακάτω φαίνεται ο τρόπος με τον οποίο κωδικοποιείται και χρησιμοποιείται η μέθοδος Elbow στην εφαρμογή.

```
# calculate the number of clusters with elbow function
wss = []
for i in range(1, 10):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=1)
    kmeans.fit(dataset)
    wss.append(kmeans.inertia_)

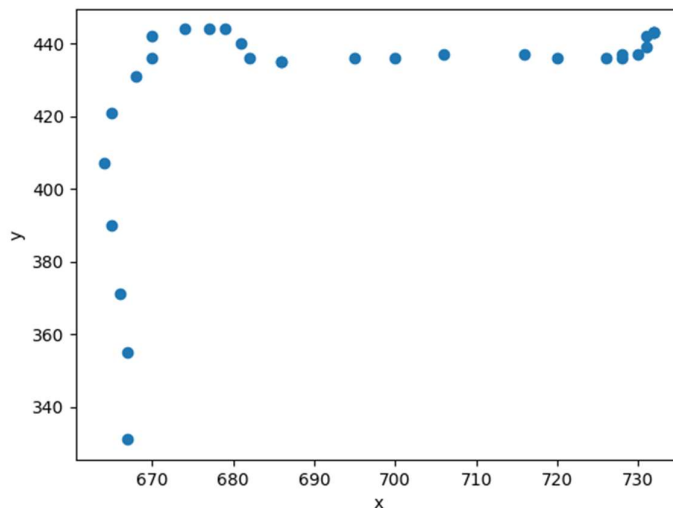
plt.plot(range(1, 10), wss)
plt.title('The Elbow method')
plt.xlabel('Number of clusters')
plt.ylabel('Sum of squared distances')
plt.show()
```

Εικόνα 23: Υλοποίηση της μεθόδου Elbow

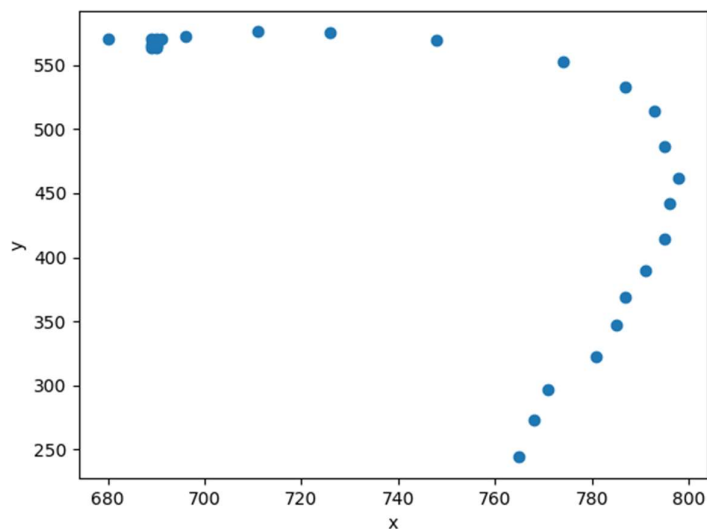
4.3.7 Παραγωγή αποτελεσμάτων

Το πρόγραμμα που χρησιμοποιεί τον αλγόριθμο μη επιβλεπόμενης μάθησης είναι το `machine_learning_algorithm.py`. Τρέχοντας την συνάρτηση `kmeans_clustering_visualization()` που έχει υλοποιηθεί μέσα στο

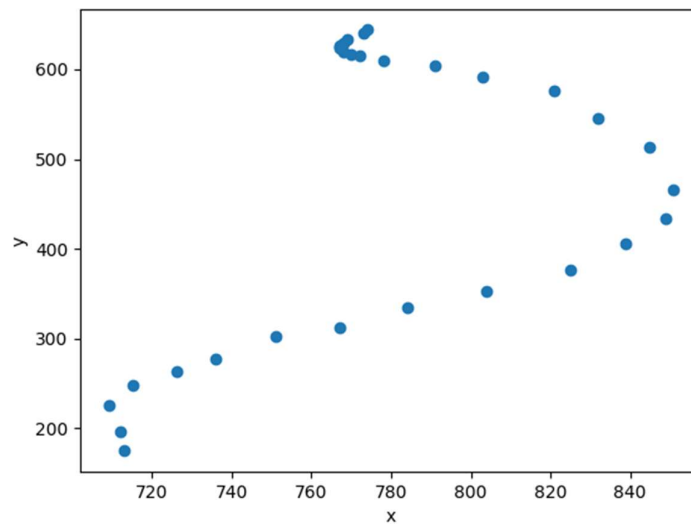
πρόγραμμα, παράγονται τα διαγράμματα και τα αποτελέσματα του αλγορίθμου Μηχανικής Μάθησης. Είναι η κύρια συνάρτηση του προγράμματος καθώς είναι υπεύθυνη για την παραγωγή όλων των αποτελεσμάτων. Όταν τροφοδοτούνται στο πρόγραμμα οι συντεταγμένες του αθλητή, πρώτα αναπαρίστανται σε διαγράμματα οι συντεταγμένες του δεξιού ώμου, αγκώνα και καρπού. Τα διαγράμματα που παράγονται είναι τα ακόλουθα.



Εικόνα 24: Συντεταγμένες του ώμου

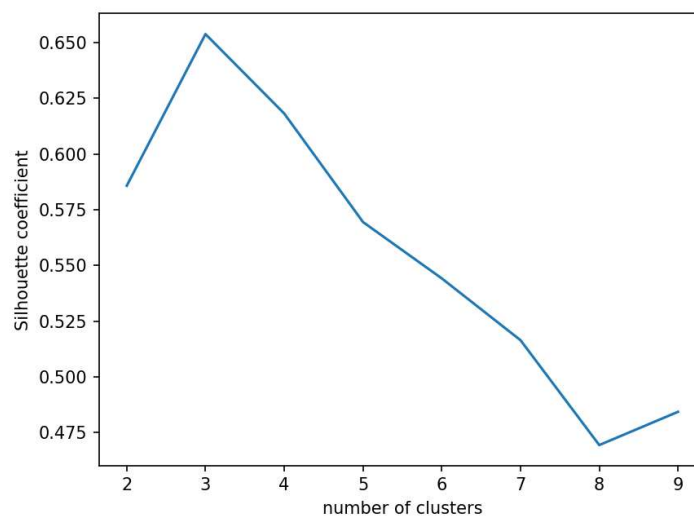


Εικόνα 25: Συντεταγμένες του αγκώνα



Εικόνα 26: Συντεταγμένες του καρπού

Παρατηρώντας τις συντεταγμένες στα διαγράμματα, μπορούμε να καταλάβουμε ότι κάθε διάγραμμα αποτυπώνει την κίνηση κάποιου σημείου του χεριού του αθλητή. Βλέποντας τα διαγράμματα είναι εύκολο να παρατηρηθεί η διαδρομή που ακολουθεί κάθε αθλητής κατά την διάρκεια της βολής. Με την χρήση συντεταγμένων έμπειρων βολών, θα μπορούσαμε να παρατηρήσουμε τις διαφορές που έχουν συγκριτικά με του αθλητή που θέλει να βελτιστοποιήσει την βολή του. Βλέπουμε ήδη τις δυνατότητες που προσφέρει η Όραση Υπολογιστών στην εφαρμογή μας. Όμως μπορούμε να βελτιστοποιήσουμε την διαδικασία της ανάλυσης και της βελτιστοποίησης με την χρήση της Μηχανικής Μάθησης. Παρακάτω παρουσιάζονται τα αποτελέσματα του αλγορίθμου K-means. Πρώτα όμως θα υπολογίσουμε και θα δούμε το βέλτιστο πλήθος συστάδων που πρέπει να χρησιμοποιηθούν για τις συντεταγμένες μας. Στην συνέχεια φαίνονται αρχικά τα αποτελέσματα του Silhouette score για τον ώμο.

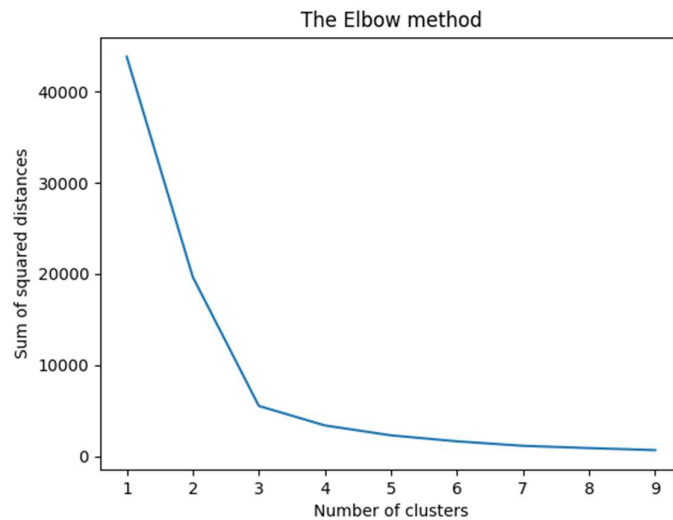


Εικόνα 27: Διάγραμμα με τα αποτελέσματα του *Silhouette score* για τον ώμο

```
For cluster: 2, the silhouette score is: 0.5857772182987869
For cluster: 3, the silhouette score is: 0.6538332570252036
For cluster: 4, the silhouette score is: 0.6181881318552553
For cluster: 5, the silhouette score is: 0.5541684172985899
For cluster: 6, the silhouette score is: 0.5342726238204583
For cluster: 7, the silhouette score is: 0.5082923949263085
For cluster: 8, the silhouette score is: 0.4683336298431778
For cluster: 9, the silhouette score is: 0.3792079571918586
```

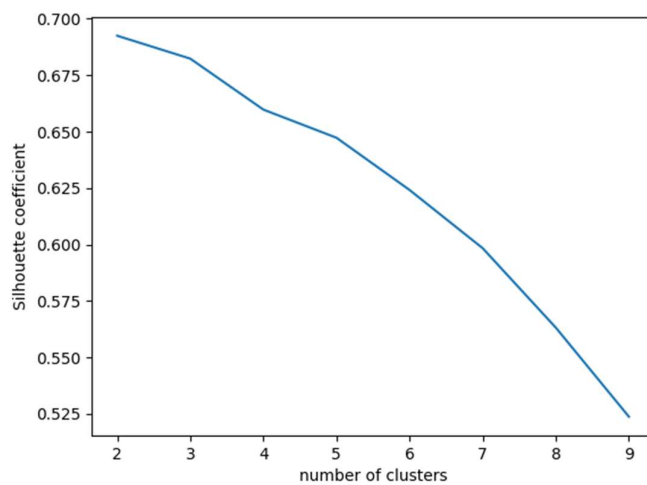
Εικόνα 28: Αποτελέσματα του *Silhouette score* για τον ώμο

Βλέπουμε ότι η βέλτιστη τιμή για το k με βάση το *Silhouette score*, έχοντας ως σύνολο δεδομένων εισόδου τις συντεταγμένες των ώμων είναι το 3. Αμέσως μετά βλέπουμε και τα αποτελέσματα της μεθόδου Elbow.



Εικόνα 29: Αποτελέσματα της μεθόδου Elbow για τον ώμο

Η μέθοδος Elbow μας επιβεβαιώνει ότι η βέλτιστη τιμή του k είναι το 3. Η ίδια διαδικασία συνεχίζεται και για τον αγκώνα.

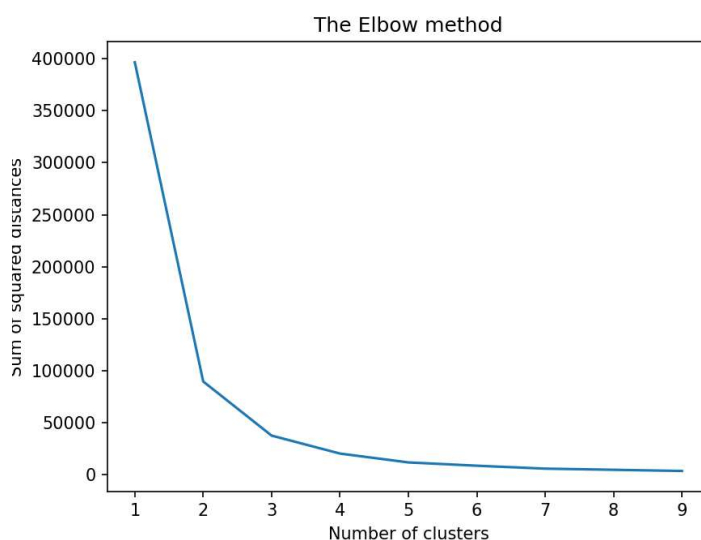


Εικόνα 30: Αποτελέσματα του Silhouette score σε διάγραμμα για τον αγκώνα

```
For cluster: 2, the silhouette score is: 0.6925073004034449
For cluster: 3, the silhouette score is: 0.6701260810532891
For cluster: 4, the silhouette score is: 0.6597692351578406
For cluster: 5, the silhouette score is: 0.6472921355768596
For cluster: 6, the silhouette score is: 0.6256934493175526
For cluster: 7, the silhouette score is: 0.5998991649850273
For cluster: 8, the silhouette score is: 0.5726022855386337
For cluster: 9, the silhouette score is: 0.5424583692616878
```

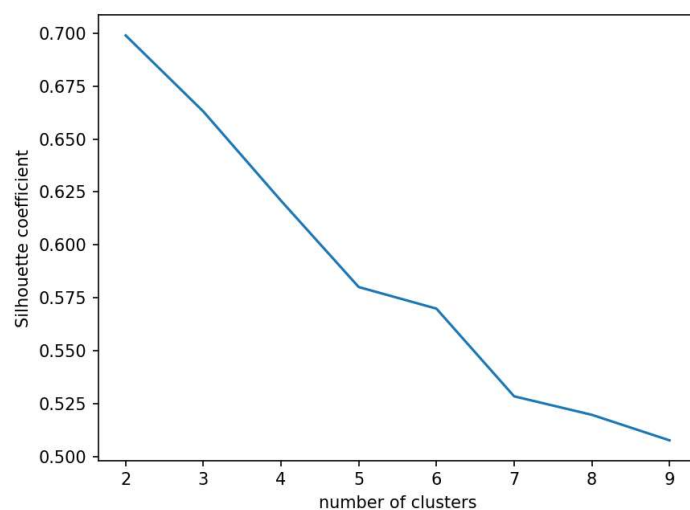
Εικόνα 31: Αποτελέσματα του Silhouette score για τον αγκώνα

Πάλι βλέπουμε ότι η βέλτιστη τιμή του k για τα δεδομένα της είναι το 3, από ότι φαίνεται. Στην συνέχεια ακολουθούν και τα αποτελέσματα της μεθόδου Elbow.



Εικόνα 32: Αποτελέσματα της μεθόδου Elbow για τον αγκώνα

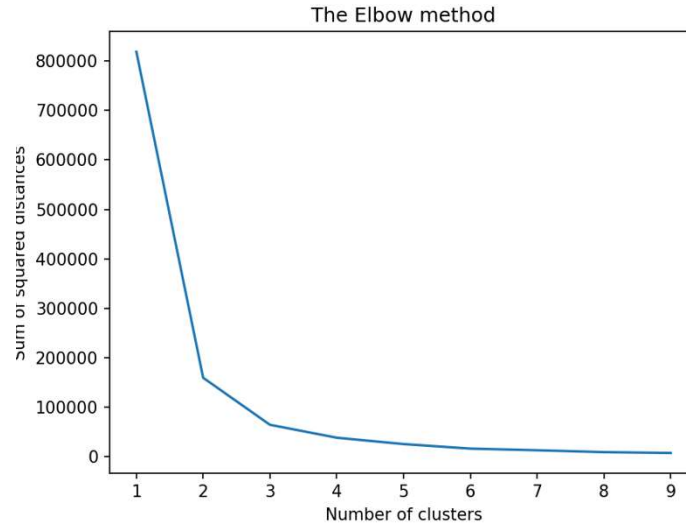
Πράγματι, η βέλτιστη τιμή για το k και εδώ είναι το 3. Τέλος παρουσιάζονται και τα αποτελέσματα του Silhouette score για τον καρπό.



Εικόνα 33: Αποτελέσματα του *Silhouette score* σε διάγραμμα για τον καρπό

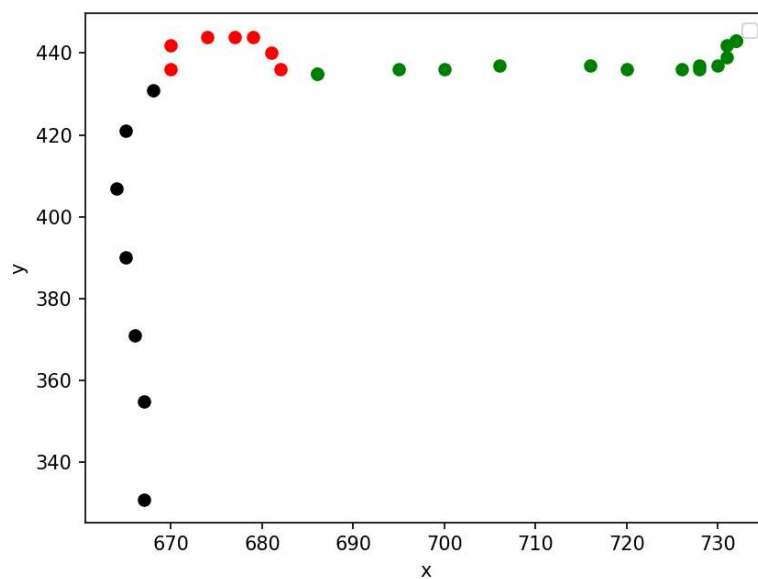
```
For cluster: 2, the silhouette score is: 0.6989516612796198
For cluster: 3, the silhouette score is: 0.6630094099251341
For cluster: 4, the silhouette score is: 0.620857555162776
For cluster: 5, the silhouette score is: 0.5900187216642812
For cluster: 6, the silhouette score is: 0.5708205090215756
For cluster: 7, the silhouette score is: 0.5284527222317263
For cluster: 8, the silhouette score is: 0.5197158534297016
For cluster: 9, the silhouette score is: 0.4779347339270655
```

Εικόνα 34: Αποτελέσματα του *Silhouette score* για τον καρπό

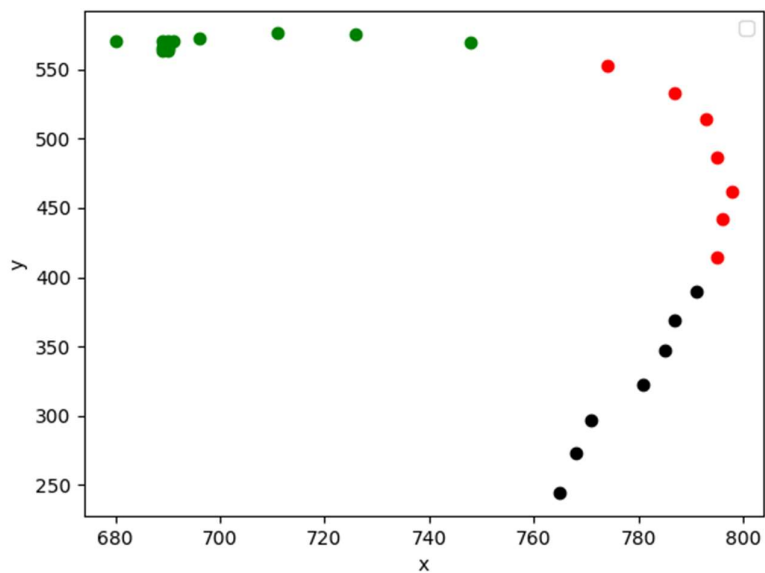


Εικόνα 35: Αποτελέσματα της μεθόδου *Elbow* για τον καρπό

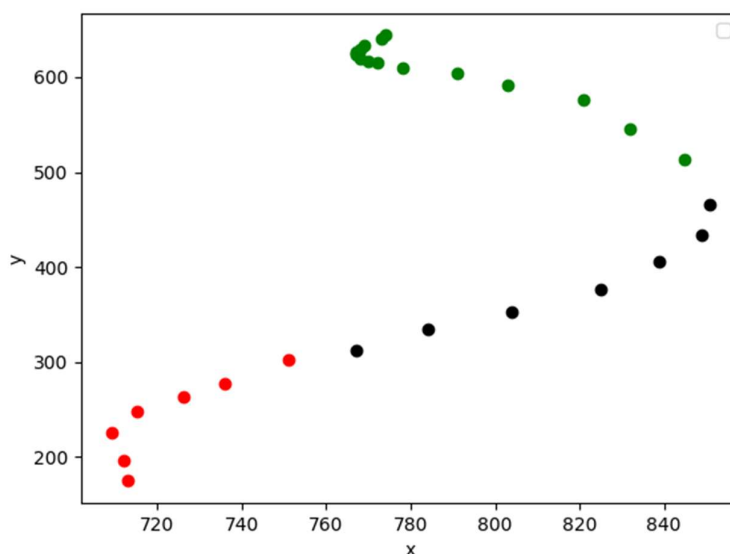
Όπως βλέπουμε και για τον καρπό, το βέλτιστο πλήθος συστάδων είναι το 3. Αυτό συμβαίνει γιατί η κίνηση της βολής αποτελείται 3 φάσης. Η πρώτη φάση είναι το σημείο έναρξης των σημείων του χεριού (ώμος, αγκώνας και καρπός) και πιάνει λίγο και από την έναρξη εκτέλεσης της βολής. Η δεύτερη φάση είναι όταν έχει ξεκινήσει η εκτέλεση της βολής και φτάνει έως και λίγο πριν το τελικό σημείο πριν απελευθερωθεί η μπάλα από το χέρι. Η τρίτη φάση είναι πριν και μετά την απελευθέρωση της μπάλας. Άρα με την βοήθεια της Μηχανικής Μάθησης αποτυπώνονται οι παραπάνω φάσεις και μπορούν να αναλυθούν με μεγαλύτερη ακρίβεια τα λάθη που γίνονται. Επίσης εντοπίζεται και η φάση στην οποία ο αθλητής κάνει λάθος, έτσι μπορούν να γίνουν ακόμα πιο ακριβείς υποδείξεις για την βελτίωση της βολής του αθλητή. Παρακάτω φαίνονται και τα αποτελέσματα της συσταδοποίησης.



Εικόνα 36: Αποτελέσματα του *k-means* για τον ώμο



Εικόνα 37: Αποτελέσματα του *k-means* για τον αγκώνα



Εικόνα 38: Αποτελέσματα του *k-means* για τον καρπό

Συνοψίζοντας, με την χρήση της Μηχανικής Μάθησης και της Όρασης Υπολογιστών καταφέραμε να υλοποιήσουμε μία εφαρμογή με την οποία βελτιστοποιείται η αθλητική απόδοση, έχοντας ως πρότυπο τα δεδομένα ενός επαγγελματία αθλητή. Μπορούν να αναλυθούν οι φάσεις τις κίνησης των βολών οπότε είναι εύκολο να εντοπιστούν τα πρότυπα και οι διαφορές με την χρήση των διαγραμμάτων και του αποτελέσματος της συσταδοποίησης. Το σύνολο του κώδικα, βρίσκεται στο αντίστοιχο repository στον λογαριασμό μου στο GitHub [93].

ΚΕΦΑΛΑΙΟ 5

ΣΥΜΠΕΡΑΣΜΑΤΑ

Πραγματοποιώντας την παραπάνω εφαρμογή, παρατηρήσαμε την σημαντικότητα της ποιότητας των δεδομένων καθώς και της επεξεργασίας τους. Είναι σημαντικό να δοθεί μεγάλη προσοχή κατά την λήψη των δεδομένων καθώς αυτά είναι ο πρώτος παράγοντας για την εκπόνηση μίας επιτυχημένης εφαρμογής Μηχανικής Μάθησης. Οι δυνατότητες της Μηχανικής Μάθησης είναι μεγάλες. Με την χρήση της πειραματικής εφαρμογής που αναπτύχθηκε, καταφέρνουμε και παράγουμε θετικά αποτελέσματα με τα οποία ένας αθλητής μπορεί να βελτιστοποιήσει την απόδοσή του. Με βάση την επεξεργασία που γίνεται στην εφαρμογή και έχοντας τα δεδομένα ενός επαγγελματία αθλητή είναι εφικτή η αλλαγή συγκεκριμένων σημείων κατά την διάρκεια της βολής, τα οποία μπορεί να αλλάξει ο χρήστης ώστε να βελτιστοποιήσει την βολή του.

Η εφαρμογή μπορεί να εμπλουτιστεί μελλοντικά. Θα μπορούσε να υπολογίζει την ομοιότητα του ερασιτέχνη αθλητή με του επαγγελματία αθλητή και να βγάζει ένα ποσοστό ομοιότητας ως αποτέλεσμα, με την χρήση κατάλληλης μετρικής. Επίσης, θα μπορούσαν να αξιοποιηθούν και η κίνηση των ποδιών καθώς είναι σημαντική η δύναμη που αντλεί από το έδαφος ο αθλητής. Τέλος, μπορεί να αξιοποιηθεί και από την φυσική, η κυκλική κίνηση για να υπολογιστεί η δύναμη που παράγεται από το χέρι καθώς και η ταχύτητα που αναπτύσσει η μπάλα μετά την απελευθέρωσή της.

Αναλύοντας τα δεδομένα που παράγονται καθημερινά μπορεί να βελτιστοποιηθεί η διάρκεια ζωής των αισθητήρων καθώς και η ποιότητα των μέσων επικοινωνίας μας. Επίσης η Μηχανική Μάθηση και η Όραση Υπολογιστών μπορούν να χρησιμοποιηθούν σε πολλούς τομείς όπως η Ιατρική και η περίθαλψη. Μέσω καμερών, μπορούν να παρακολουθούνται οι κινήσεις των ασθενών. Οπότε αν η κάμερα εντοπίσει ότι κάποιος ασθενής βρίσκεται στο έδαφος λόγω ατυχήματος, τότε θα ενημερώνονται οι νοσοκόμοι να τον βοηθήσουν. Αυτή η τεχνική ονομάζεται εντοπισμός πτώσης (fall detection) και χρησιμοποιεί την Όραση Υπολογιστών. Επίσης, ο συνδυασμός των δύο τεχνολογιών μπορεί να αξιοποιηθεί από τους προπονητές των ομάδων. Οι κάμερες μπορούν να παρακολουθούν τους αθλητές μίας ομάδας κατά την διάρκεια των αγώνων προκειμένου να παράγονται δεδομένα

επίδοσης (αντοχή, τεχνική κλπ.). Έτσι οι προπονητές μπορούν να εντοπίσουν ποιοι αθλητές έχουν αρκετή αντοχή και ποιοι είναι τεχνικά καλύτεροι, με αποτέλεσμα οι προπονητές να πάρουν καλύτερες αποφάσεις όταν κάνουν αλλαγές ή όταν δημιουργούν τις βασικές ομάδες με τις οποίες θα ξεκινήσουν τον αγώνα. Με την Μηχανική Μάθηση μπορούμε να διαχειριστούμε μεγάλο πλήθος δεδομένων. Με αυτό το χαρακτηριστικό μας δίνεται η δυνατότητα, να αλλάξουμε ριζικά τις ζωές μας προς το καλύτερο. Πλέον μπορούμε να βγάλουμε νόημα από το χάος των συνόλων δεδομένων που παράγονται καθημερινά. Ήρθε ο καιρός να εκμεταλλευτούμε και αυτή την δυνατότητα της Τεχνητής Νοημοσύνης, ώστε να υλοποιήσουμε εφαρμογές που κάνουν την ζωή μας καλύτερη.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] History of artificial intelligence: Revision history. (2022, 5 March). In Wikipedia. Ανακτήθηκε 5 Μαρτίου 2022, από https://en.wikipedia.org/w/index.php?title=History_of_artificial_intelligence&dir=prev&action=history
- [2] Tutorialspoint. (2021). Artificial Intelligence – Overview. Αναρτήθηκε από https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_overview.htm
- [3] ITU. (2021). Artificial Intelligence for good. Ανακτήθηκε από <https://www.itu.int/en/mediacentre/backgrounders/Pages/artificial-intelligence-for-good.aspx>
- [4] Raj, R. (2018, 5 March). What is Artificial Intelligence? In 5 minutes. Ανακτήθηκε από https://www.youtube.com/watch?v=2ePf9rue1Ao&t=2s&ab_channel=RajRamesh
- [5] Nadia, B. (2020, November 30). Artificial Intelligence has changed our world. Ανακτήθηκε από <https://wsimag.com/science-and-technology/64215-artificial-intelligence-has-changed-our-world>
- [6] PwC US. (2021, 5 March) How artificial intelligence is changing the world. Ανακτήθηκε από https://www.youtube.com/watch?v=JEX2BoScTFY&ab_channel=PwCUS
- [7] Tutorialspoint. (2021). Artificial Intelligence - Intelligent Systems. https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligent_systems.htm
- [8] Software Testing Help. (2022). What Is Artificial Intelligence: Definition & Sub-Fields Of AI. Ανακτήθηκε από <https://www.softwaretestinghelp.com/what-is-artificial-intelligence/>
- [9] Shiffman, Daniel. (2012) The nature of code. Ανακτήθηκε από <https://natureofcode.com/book/chapter-10-neural-networks/>
- [10] Neural network. (2022, 5 March). In Wikipedia. Ανακτήθηκε 5 Μαρτίου 2022, από https://en.wikipedia.org/wiki/Neural_network#History
- [11] Neuroscience. (2022, 5 March). In Wikipedia. Ανακτήθηκε 5 Μαρτίου 2022, από <https://en.wikipedia.org/wiki/Neuroscience>

- [12] Tyagi, N. (2021, 23 January). 6 Major Branches of Artificial Intelligence (AI). Ανακτήθηκε από <https://www.analyticssteps.com/blogs/6-major-branches-artificial-intelligence-ai>
- [13] Simplilearn. (2019, 5 March). Neural Network In 5 Minutes | What Is A Neural Network? | How Neural Networks Work | Simplilearn. Ανακτήθηκε από https://www.youtube.com/watch?v=bfmFfD2RIcg&ab_channel=Simplilearn
- [14] Cognifit research. (2022). Ο ανθρώπινος εγκέφαλος. Ανακτήθηκε από <https://www.cognifit.com/el/brain>
- [15] Zilouchian, A. (2001). Fundamentals of neural networks. Intelligent control systems using soft computing methodologies, (1), 1-5.
- [16] Sharma, S. (2017, 9 Sep). What the Hell is Perceptron? The Fundamentals of Neural Networks. Ανακτήθηκε από <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>
- [17] Bernard Marr & Co. (2021). Deep Learning Vs Neural Networks – What's The Difference?. Ανακτήθηκε από <https://bernardmarr.com/deep-learning-vs-neural-networks-whats-the-difference/>
- [18] DeepAI. (2022). Perceptron. Ανακτήθηκε από <https://deepai.org/machine-learning-glossary-and-terms/perceptron>
- [19] Gallant, S. I., & Gallant, S. I. (1993). Neural network learning and expert systems. MIT press.
- [20] Great Learning Team. (2021, 5 September). Types of Neural Networks and Definition of Neural Network. Ανακτήθηκε από <https://www.mygreatlearning.com/blog/types-of-neural-networks/>
- [21] Vidiyala, R. (2020, 17 December). 6 Types of Neural Networks Every Data Scientist Must Know. Ανακτήθηκε από <https://towardsdatascience.com/6-types-of-neural-networks-every-data-scientist-must-know-9c0d920e7fce>
- [22] Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of microbiological methods, 43(1), 3-31.
- [23] Maladkar, K. (2018, January 15). 6 Types of Artificial Neural Networks Currently Being Used in Machine Learning. Ανακτήθηκε από

<https://analyticsindiamag.com/6-types-of-artificial-neural-networks-currently-being-used-in-todays-technology/>

- [24] Erb, R. J. (1993). Introduction to backpropagation neural network computation. *Pharmaceutical research*, 10(2), 165-170.
- [25] Al-Masri, A. (2019, 30 January). How Does Back-Propagation in Artificial Neural Networks Work?. Ανακτήθηκε από <https://towardsdatascience.com/how-does-back-propagation-in-artificial-neural-networks-work-c7cad873ea7>
- [26] Awodele, O., & Jegede, O. (2009). Neural networks and its application in engineering. *Sci IT*, 83-95.
- [27] Kavlakoglu, E. (2020, 27 May). AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?. Ανακτήθηκε από <https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>
- [28] Rancho Labs. (2021). 6 Major Sub-Fields of Artificial Intelligence. Ανακτήθηκε από <https://rancholabs.medium.com/6-major-sub-fields-of-artificial-intelligence-77f6a5b28109>
- [29] Javatpoint. (2011). Subsets of Artificial Intelligence. Ανακτήθηκε από <https://www.javatpoint.com/subsets-of-ai>
- [30] Bhardwaj, A., Di, W., & Wei, J. (2018). *Deep Learning Essentials: Your hands-on guide to the fundamentals of deep learning and neural network modeling*. Packt Publishing Ltd.
- [31] Aravindpai. (2020, 17 February). CNN vs. RNN vs. ANN – Analyzing 3 Types of Neural Networks in Deep Learning. Ανακτήθηκε από <https://www.analyticsvidhya.com/blog/2020/02/cnn-vs-rnn-vs-mlp-analyzing-3-types-of-neural-networks-in-deep-learning/>
- [32] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [33] Chojecki, P. (2019, 18 November). What is Deep Learning?. Ανακτήθηκε από <https://medium.com/swlh/what-is-deep-learning-b2cd80911cbc>
- [34] Goyal, K. (2019, 13 December). Deep Learning vs Neural Networks: Difference Between Deep Learning and Neural Networks. <https://www.upgrad.com/blog/deep-learning-vs-neural-networks-difference->

[between-deep-learning-and-neural-networks/#What is the difference between Deep Learning and Machine Learning](#)

- [35] Intellipaat. (2022). What is Artificial Intelligence?. Ανακτήθηκε από <https://intellipaat.com/blog/what-is-artificial-intelligence/>
- [36] Kyrykovich, A. (2020, February). Deep Neural Networks. Ανακτήθηκε από <https://www.kdnuggets.com/2020/02/deep-neural-networks.html>
- [37] Healthyliving.gr. (2013). Το φως, η όραση και το μάτι. Ανακτήθηκε από <https://www.healthyliving.gr/2013/01/08/%CF%84%CE%BF-%CF%86%CF%89%CF%82-%CE%B7-%CF%8C%CF%81%CE%B1%CF%83%CE%B7-%CE%BA%CE%B1%CE%B9-%CF%84%CE%BF-%CE%BC%CE%AC%CF%84%CE%B9/>
- [38] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017, August). Understanding of a convolutional neural network. In 2017 international conference on engineering and technology (ICET) (pp. 1-6). Ieee.
- [39] Saini, B. (2021, 1 February). Introduction to Deep Learning. Ανακτήθηκε από <https://ai.plainenglish.io/introduction-to-deep-learning-2e7f21cad4e>
- [40] Madhavan, S. (2017, 8 September). Deep learning architectures. Ανακτήθηκε από <https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures/>
- [41] Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019.
- [42] Arras, L., Osman, A., Müller, K. R., & Samek, W. (2019). Evaluating recurrent neural network explanations. arXiv preprint arXiv:1904.11829.
- [43] IBM Cloud Education. (2020). Recurrent Neural Networks. Ανακτήθηκε από <https://www.ibm.com/cloud/learn/recurrent-neural-networks>
- [44] Liddy, E. D. (2001). Natural language processing.
- [45] Allen, J. F. (2003). Natural language processing. In Encyclopedia of computer science (pp. 1218-1222).

- [46] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18(5), 544-551.
- [47] Tanna, V. (2021, 24 March). INTRODUCTION TO DIFFERENT LEVELS OF NATURAL LANGUAGE PROCESSING. Ανακτήθηκε από <https://www.datascienceprophet.com/different-levels-of-nlp/>
- [48] Gutierrez-Garcia, J. O., & López-Neri, E. (2015, July). Cognitive computing: a brief survey and open research challenges. In 2015 3rd international conference on applied computing and information technology/2nd international conference on computational science and intelligence (pp. 328-333). IEEE.
- [49] Chen, M., Herrera, F., & Hwang, K. (2018). Cognitive computing: architecture, technologies and intelligent applications. *Ieee Access*, 6, 19774-19783.
- [50] Botelho, B. (2018, July). Cognitive computing. Ανακτήθηκε από <https://www.techtarget.com/searchenterpriseai/definition/cognitive-computing>
- [51] Wu, J. (2019, 6 June). AI and Cognitive Computing. Ανακτήθηκε από <https://digileaders.com/ai-and-cognitive-computing/>
- [52] Richardson, M. (2021, 23 September). Cognitive Computing vs. AI: 3 Key Differences and Why They Matter. Ανακτήθηκε από <https://www.toolbox.com/tech/artificial-intelligence/articles/cognitive-computing-vs-ai/>
- [53] Makadia, M. (2019, 10 May). What is Cognitive Computing? How are Enterprises benefitting from Cognitive Technology?. Ανακτήθηκε από <https://towardsdatascience.com/what-is-cognitive-computing-how-are-enterprises-benefitfrom-cognitive-technology-6441d0c9067b>
- [54] IBM. (2022). What is computer vision?. Ανακτήθηκε από <https://www.ibm.com/topics/computer-vision>
- [55] Huang, T. (1996). Computer vision: Evolution and promise.
- [56] Danuser, G. (2011). Computer vision in cell biology. *Cell*, 147(5), 973-978.
- [57] Raghavan, V. V., Gudivada, V. N., Govindaraju, V., & Rao, C. R. (2016). *Cognitive computing: Theory and applications*. Elsevier.

- [58] Demirkan, H., Earley, S., & Harmon, R. R. (2017). Cognitive computing. IT professional, 19(4), 16-20.
- [59] Chen, M., Herrera, F., & Hwang, K. (2018). Cognitive computing: architecture, technologies and intelligent applications. Ieee Access, 6, 19774-19783.
- [60] Zion, B. (2012). The use of computer vision technologies in aquaculture—a review. Computers and electronics in agriculture, 88, 125-132.
- [61] Tutorialspoint. (2021). Machine Learning with Python – Basics. Ανακτήθηκε από https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_basics.htm
- [62] Tutorialspoint. (2021). Machine Learning with Python Tutorial. Ανακτήθηκε από https://www.tutorialspoint.com/machine_learning_with_python/index.htm
- [63] NetApp. (2022). What Is Machine Learning?. Ανακτήθηκε από <https://www.netapp.com/artificial-intelligence/what-is-machine-learning/>
- [64] Foote, K. (2021, 3 December). A Brief History of Machine Learning. Ανακτήθηκε από <https://www.dataversity.net/a-brief-history-of-machine-learning/#>
- [65] Murali, A. (2021, 2 August). Understanding Generation Data. Ανακτήθηκε από <https://www.forbes.com/sites/forbestechcouncil/2021/08/02/understanding-generation-data/?sh=295ddb3136b7>
- [66] Bulao, J. (2022, 6 February). How Much Data Is Created Every Day in 2022?. Ανακτήθηκε από <https://techjury.net/blog/how-much-data-is-created-every-day/#gref>
- [67] Barrett, S. (2021, 30 March). How Much Data Is Produced Every Day in 2021. Ανακτήθηκε από <https://the-tech-trend.com/reviews/how-much-data-is-produced-every-day/>
- [68] SeedScientific. (2021). How Much Data Is Created Every Day?. Ανακτήθηκε από <https://seedscientific.com/how-much-data-is-created-every-day/>

- [69] The World Bank. (2021). Crossing borders. Ανακτήθηκε από <https://wdr2021.worldbank.org/stories/crossing-borders/>
- [70] Google Cloud Tech. (2021, 6 March). Can AI make me a better athlete? | Using machine learning to analyze penalty kicks. Ανακτήθηκε από https://www.youtube.com/watch?v=yLrOy2Xedgk&ab_channel=GoogleCloudTech
- [71] Data Magic. (2021, 6 March). Step By Step Machine Learning | Machine Learning Steps | Data Magic. Ανακτήθηκε από https://www.youtube.com/watch?v=GyrhVZnKM00&ab_channel=DataMagic
- [72] Hyperparameter (machine learning). (2022, 6 March). Ανακτήθηκε 6 Μαρτίου 2022, από [https://en.wikipedia.org/wiki/Hyperparameter_\(machine_learning\)](https://en.wikipedia.org/wiki/Hyperparameter_(machine_learning))
- [73] RISKSPAN. (2017, 7 November). Tuning Machine Learning Models. Ανακτήθηκε από <https://riskspan.com/tuning-machine-learning-models/>
- [74] Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. In Journal of physics: conference series (Vol. 1142, No. 1, p. 012012). IOP Publishing.
- [75] El Naqa, I., Li, R., & Murphy, M. J. (Eds.). (2015). Machine learning in radiation oncology: theory and applications. Springer.
- [76] Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9, 381-386.
- [77] Soong, T. T. (2004). Fundamentals of probability and statistics for engineers. John Wiley & Sons.
- [78] Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. Machine Learning, 109(2), 373-440.
- [79] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.
- [80] MediaPipe. (2020). Home - mediapipe. Ανακτήθηκε από <https://google.github.io/mediapipe/>
- [81] Seede. (2008). Getting Started with MediaPipe on reTerminal. Ανακτήθηκε από https://wiki.seeedstudio.com/reTerminal_ML_MediaPipe/
- [82] MediaPipe. (2020). Solutions. Ανακτήθηκε από <https://google.github.io/mediapipe/solutions/solutions.html>

- [83] MediaPipe. (2020). MediaPipe Pose. Ανακτήθηκε από <https://google.github.io/mediapipe/solutions/pose>
- [84] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- [85] Team, T. P. D. (2020). pandas-dev/pandas: Pandas (Version latest). doi:10.5281/zenodo.3509134
- [86] OpenCV. (2022). About. Ανακτήθηκε από <https://opencv.org/about/>
- [87] PyPI. (2020). Xlwt. Ανακτήθηκε από <https://pypi.org/project/xlwt/>
- [88] Docs.python. (2001). Miscellaneous operating system interfaces. Ανακτήθηκε από <https://docs.python.org/3/library/os.html>
- [89] Tatariants, M. (2020, 21 December). AI Pose Estimation in Fitness Application. Ανακτήθηκε από <https://www.unite.ai/ai-pose-estimation-in-fitness-application/>
- [90] Bhardwaj, A. (2020, 26 May). Silhouette Coefficient Validating clustering techniques. Ανακτήθηκε από <https://towardsdatascience.com/silhouette-coefficient-validating-clustering-techniques-e976bb81d10c>
- [91] Saji, B. (2021, 20 January). In-depth Intuition of K-Means Clustering Algorithm in Machine Learning. Ανακτήθηκε από <https://www.analyticsvidhya.com/blog/2021/01/in-depth-intuition-of-k-means-clustering-algorithm-in-machine-learning/>
- [92] Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. doi:10.1109/MCSE.2007.55
- [93] <https://github.com/johny-kampe/Athletic-Performance-Optimization-Diploma-Thesis>