



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS



**Πρόβλεψη δράσης ανθρώπου από την κίνηση του χεριού για
συνεργασία ανθρώπου-ρομπότ**

Από

Γεώργιο Μπαζάκο

Υποβάλλεται

για την εκπλήρωση των προϋποθέσεων λήψης

Μεταπτυχιακού Διπλώματος

στην «Τεχνητή Νοημοσύνη»

στο

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Απρίλιος 2021

Πανεπιστήμιο Πειραιώς, ΕΚΕΦΕ «ΔΗΜΟΚΡΙΤΟΣ». Κάτοχος όλων των δικαιωμάτων

Συγγραφέας Γεώργιος Μπαζάκος

ΔΠΜΣ «Τεχνητή Νοημοσύνη»

Απρίλιος 12, 2021

Έγινε αποδεκτό από

Μαρία Δαγιόγλου
Ερευνήτρια
Επιβλέπουσα

Έγινε αποδεκτό από

Θεόδωρο
Γιαννακόπουλο
Ερευνητή
Επιβλέπων

Έγινε αποδεκτό από

Γεώργιο Βούρο
Καθηγητή
Μέλος Εξεταστικής
Επιτροπής

Πρόβλεψη δράσης ανθρώπου από την κίνηση του χεριού για συνεργασία ανθρώπου-ρομπότ

Από

Γεώργιο Μπαζάκο

Υποβλήθηκε στο ΔΠΜΣ «Τεχνητή Νοημοσύνη» την 12 Απριλίου 2021 ως υποχρέωση για την λήψη Μεταπτυχιακού Διπλώματος Σπουδών

Περίληψη

Στην σημερινή εποχή έχει παρατηρηθεί ραγδαία ανάπτυξη στον κλάδο της ρομποτικής και της τεχνητής νοημοσύνης και αυτό έχει ως αποτέλεσμα την αύξηση της χρήσης των ρομπότ σε διάφορους τομείς της καθημερινότητας του ανθρώπου. Για να επιτευχθεί η χρήση των ρομπότ σε καθημερινές εργασίες πρέπει αυτά να είναι σε θέση να αντιμετωπίσουν μη δομημένα, απρόβλεπτα και συνεχώς μεταβαλλόμενα περιβάλλοντα. Συνεπώς απαιτείται να ενεργούν αυτόνομα, να μαθαίνουν πως να αντιδρούν σε διάφορες μεταβολές του περιβάλλοντος, καθώς και τις επιπτώσεις που έχουν οι ενέργειες τους στο περιβάλλον.

Η παρούσα εργασία μελετά τη συνεργασία ανθρώπου - ρομπότ σε ένα κοινό εργασιακό περιβάλλον. Συγκεκριμένα, για την επίτευξη μιας καλύτερης και πιο φυσικής συνεργασίας βασική προϋπόθεση είναι να μπορούν τα ρομπότ να προβλέψουν τις ενέργειες των ανθρώπων. Σε αυτό βοηθάει η παρατήρηση ότι η αρχή της κίνησης του ανθρώπου περιέχει χρήσιμες πληροφορίες ικανές για την πρόβλεψη της τελικής δράσης αυτού. Για παράδειγμα, η πόζα της παλάμης διαφοροποιείται από την αρχή μιας κίνησης ανάλογα με το μέγεθος του αντικειμένου που πρόκειται να πιάσει κάποιος.

Στα πλαίσια της εργασίας έγινε συλλογή κατάλληλων δεδομένων και αναζητήθηκε κατάλληλη μέθοδος Τεχνητής Νοημοσύνης που θα επιτρέψει σε ένα ρομπότ να προβλέψει πως θα δράσει ο άνθρωπος στον κοινό εργασιακό χώρο μέσα από οπτική πληροφορία (κάμερα) των κινήσεων του χεριού και της παλάμης.

Αρχικά τα δεδομένα τα οποία συλλέχθηκαν αφορούσαν τις κινήσεις διάφορων ανθρώπων οι οποίοι είχαν σαν στόχο να πιάσουν και να μετακινήσουν τρία διαφορετικού μεγέθους αντικείμενα. Από αυτά τα δεδομένα κατασκευάστηκαν δυο σύνολα δεδομένων. Το πρώτο αφορούσε την τρισδιάστατη πληροφορία και το δεύτερο τη δισδιάστατη πληροφορία κάθε κίνησης. Για το δεύτερο σύνολο δεδομένων πέραν από τις τρεις κλάσεις μια για κάθε αντικείμενο δοκιμάστηκε και η προβλέπει ανάμεσα σε δυο κλάσεις για όλους τους πιθανούς συνδυασμούς των τριών κλάσεων. Μετά από την επεξεργασία των δεδομένων για καθαρισμό του θορύβου καθώς και για την οριοθέτηση του κομματιού της κινήσεως προς μελέτη ακολούθησε η εφαρμογή αλγορίθμων μηχανικής μάθησης στα επεξεργασμένα δεδομένα. Οι πέντε αλγόριθμοι μηχανικής μάθησης που εφαρμόστηκαν είναι οι: SVM, Decision Tree, Random Forest, Extra Tree και ο Gradient boosting. Η διαδικασία της εκτίμησης των αποτελεσμάτων έγινε μέσω της διαδικασίας του K-Fold Cross validation. Τα αποτελέσματα έδειξαν ότι ο καλύτερος αλγόριθμος με ποσοστό επιτυχίας 94% ήταν ο Gradient boosting για το δισδιάστατο σύνολο δεδομένων με δυο κλάσεις αυτές του μικρού και του μεγάλου αντικειμένου.

Λέξεις κλειδιά : Αλληλεπίδραση ανθρώπου – ρομπότ, Τεχνητή Νοημοσύνη, αλγόριθμοι μηχανικής μάθησης.

Επιβλέπουσα:
Ακαδημαϊκή Θέση:

Επιβλέπων:
Ακαδημαϊκή Θέση:

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τους επιβλέποντες μου κα Μαρία Δαγιόγλου και κο Θεόδωρο Γιαννακόπουλο, γιατί μου έδωσαν την ευκαιρία να ασχοληθώ με ένα ιδιαίτερα ενδιαφέρον θέμα, είχα την αμέριστη και ουσιαστική επιστημονική βοήθεια και καθοδήγηση τους. Τους ευχαριστώ θερμά για τις εξαιρετικά ωφέλιμες κριτικές παρατηρήσεις τους στην επεξεργασία του θέματος αυτού. Επίσης θα ήθελα να ευχαριστήσω την οικογένεια μου για τη στήριξη τους, την συμπαράσταση τους, καθώς και για την πεποίθηση τους ότι θα τα καταφέρω να πετύχω όλους τους στόχους μου.

Οι απόψεις που εκφράζονται εδώ, τα ευρήματα και τα συμπεράσματα είναι αυτά του συγγραφέως και δεν εκφράζουν τις απόψεις του Πανεπιστημίου Πειραιώς ή του Ινστ. Πληροφορικής και Τηλεπικοινωνιών του ΕΚΕΦΕ «Δημόκριτος».

Περιεχόμενα

1.1	Επιστήμη της ρομποτικής	7
1.1.1	Αλληλεπίδραση ανθρώπου ρομπότ	7
1.1.2	Ασφάλεια στην συνεργασία ανθρώπου ρομπότ.	9
1.1.3	Συνεισφορά	10
1.2	Μηχανική μάθηση	11
1.2.1	Τι είναι η μηχανική μάθηση	11
1.2.2	Ταξινόμηση της μηχανικής μάθησης	12
1.2.3	Διαχωρισμός μηχανικής μάθησης με βάση το βέλτιστο αποτέλεσμα	14
1.3	Αλγόριθμοι μηχανικής μάθησης	16
1.3.1	Μηχανισμοί διανυσμάτων στήριξης - Support vector machines	16
1.3.2	Δέντρα αποφάσεων – Decision tree	20
1.3.3	Τυχαία δάση - Random forests	22
1.3.4	Εξαιρετικά τυχαία δέντρα - Extra Trees - Extremely Randomized Trees	26
1.3.5	Βαθμωτή ενίσχυση - Gradient boosting	27
1.4	Διασταυρωμένη Επικύρωση – Cross Validation	30
1.5	Σχετικές εργασίες	33
2	Μεθοδολογία	35
2.1	Συλλογή δεδομένων	35
2.2	Ανάλυση δεδομένων	37
2.2.1	Ανάλυση οπτικής πληροφορίας	37
2.2.2	Φιλτράρισμα θορύβου	39

2.2.3	Κόψιμο της κίνησης	39
2.2.3.1	Κόψιμο κίνησης για 3D δεδομένα	40
2.2.3.2	Κόψιμο δεδομένων 2D	45
2.2.4	Σύνολα δεδομένων	46
2.3	Καθορισμός χαρακτηριστικών	47
2.4	Γραφήματα αποτελεσμάτων	53
3	Αποτελέσματα	55
3.1	Τρισδιάστατο σύνολο δεδομένων.	55
3.2	Δισδιάστατο σύνολο δεδομένων με τρεις κλάσεις	56
3.3	Δισδιάστατο σύνολο δεδομένων με δυο κλάσεις	64
4	Συμπεράσματα	73
5	Βιβλιογραφία	76

Λίστα Εικόνων/ List of Figures

Εικόνα 1: Πλαίσιο μηχανικής μάθησης[19].....	12
Εικόνα 2:Κατηγορίες μηχανικής μάθησης[20]	13
Εικόνα 3:Υπερπλαίσια σε χώρους χαρακτηριστικών 2 και 3 διαστάσεων[49].	17
Εικόνα 4:Αριστερά χαμηλή τιμή κανονικοποίησης ενώ δεξιά υψηλή[50].	18
Εικόνα 5:Στην αριστερή εικόνα με υψηλό γάμμα λαμβάνονται υπόψη μόνο τα κοντινά σημεία, ενώ στην δεξιά με χαμηλό γάμμα συνεισφέρουν και τα μακρινά[50].	19
Εικόνα 6:Το πρώτο «καλό» περιθώριο απέχει εξίσου από τις δύο κλάσεις, αντίθετα το δεύτερο «κακό» περιθώριο βρίσκεται πολύ κοντά στη μπλε κλάση[50].....	19
Εικόνα 7:Δομή του ταξινομητή Random Forest[51]	23
Εικόνα 8:Απεικόνιση του ταξινομητή Gradient Boosting[52].....	28
Εικόνα 9:Λάθος εκπαίδευσης και επικύρωσης σε σχέση με τις παραμέτρους[53].	31
Εικόνα 10:Ο χώρος του πειράματος.....	35
Εικόνα 11:Τυπική έξοδος του Openpose.	38
Εικόνα 12:Πορεία επεξεργασίας δεδομένων.	38
Εικόνα 13: Απεικόνιση κίνησης καρπού πριν και μετά από την επεξεργασία.	39
Εικόνα 14:Standard deviation 3D δεδομένων. Εικόνα 15:Standard deviation 2D δεδομένων.	40
Εικόνα 16: Ιστόγραμμα Frames 3D Εικόνα 17: Ιστόγραμμα Frames 2D	46
Εικόνα 18: Ιστόγραμμα θέσης καρπού στον άξονα Y για τρεις κλάσεις.....	48
Εικόνα 19: Ιστόγραμμα θέσης καρπού στον άξονα X για τρεις κλάσεις.....	48
Εικόνα 20 Ιστόγραμμα θέσης δείκτη στον άξονα X για τρεις κλάσεις.....	49
Εικόνα 21 Ιστόγραμμα θέσης δείκτη στον άξονα Y για τρεις κλάσεις.....	49
Εικόνα 22: Ιστόγραμμα απόστασης δείκτη αντίχειρα για τις τρεις κλάσεις.....	49
Εικόνα 23 Γραφήματα μεταξύ όλων των χαρακτηριστικών	51
Εικόνα 24: Γράφημα συσχέτισης χαρακτηριστικών.	52
Εικόνα 25 Διάγραμμα απόδοσης αλγορίθμου αλγορίθμου	Εικόνα 26: Παράδειγμα heatmap 53
Εικόνα 27: Διάγραμμα απόδοσης αλγορίθμου σε τρισδιάστατα δεδομένα(SVM).....	56

Εικόνα 28: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(SVM).	57
Εικόνα 29: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(Gradient Boosting).	59
Εικόνα 30:: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(Decision Tree).	60
Εικόνα 31: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(Random Forest).	61
Εικόνα 32: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(Extra Tree).	63
Εικόνα 33: Συγκεντρωτικό διάγραμμα απόδοσης αλγορίθμων σε δισδιάστατα δεδομένα με τρεις κλάσεις	63
Εικόνα 34: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με δυο κλάσει(SVM)..	68
Εικόνα 35: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με δυο κλάσει(Gradient Boosting).	69
Εικόνα 36: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με δυο κλάσεις(Decision Tree).	70
Εικόνα 37: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με δυο κλάσεις(Random Forest).	71
Εικόνα 38:: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με δυο κλάσεις (Extra Tree).	72
Εικόνα 39: Συγκεντρωτικό διάγραμμα απόδοσης αλγορίθμων σε δισδιάστατα δεδομένα με δυο κλάσεις.	72
Εικόνα 40: Αποτελέσματα Predicting Object Size from Hand Kinematics	74

Λίστα Πινάκων/ List of Tables

Πίνακας 1	41
Πίνακας 2	43
Πίνακας 3	45
Πίνακας 4	55
Πίνακας 5	57
Πίνακας 6	58
Πίνακας 7	59
Πίνακας 8	61
Πίνακας 9	62
Πίνακας 10	64
Πίνακας 11	64
Πίνακας 12.....	64
Πίνακας 13.....	65
Πίνακας 14.....	65
Πίνακας 15.....	65
Πίνακας 16.....	65
Πίνακας 17.....	66
Πίνακας 18.....	66
Πίνακας 19.....	66
Πίνακας 20	66
Πίνακας 21.....	67
Πίνακας 22	67
Πίνακας 23	67
Πίνακας 24	67
Πίνακας 25	73
Πίνακας 26	74

Βιβλιογραφική επισκόπηση

1.1 Επιστήμη της ρομποτικής

Η επιστήμη της ρομποτικής είναι ένας κλάδος που εμπεριέχει τη μηχανολογία, την πληροφορική, την ηλεκτρολογία και άλλες επιμέρους επιστήμες. Επιδίδεται στη κατασκευή, την λειτουργία, τον χειρισμό των ρομπότ αλλά και τον προγραμματισμό τους. Οι τεχνολογίες αυτές αξιοποιούνται για την ανάπτυξη μηχανών που θα δύναται να λειτουργούν σαν υποκατάστατα του ανθρώπου. Η δημιουργία μηχανών οι οποίες θα μπορούσαν να λειτουργούν αυτόνομα και θα λαμβάνουν αποφάσεις χωρίς την καθοδήγηση του ανθρώπου δεν είχε πραγματική άνθηση μέχρι τον 20ο αιώνα.

Σύμφωνα με το Robot Institute of America (1979)¹, ως ρομπότ μπορούμε να ορίσουμε έναν μηχανισμό σχεδιασμένο ώστε, μέσω προγραμματιζόμενων κινήσεων, να μεταφέρει υλικά, τεμάχια, εργαλεία ή εξειδικευμένες συσκευές, με σκοπό την επιτέλεση ποικίλων εργασιών.

1.1.1 Αλληλεπίδραση ανθρώπου ρομπότ

Η αλληλεπίδραση ανθρώπου ρομπότ (Human-Robot Interaction - HRI) αποτελεί διεπιστημονικό πεδίο το οποίο έχει συνεισφορές από την αλληλεπίδραση ανθρώπου – υπολογιστή, την τεχνητή νοημοσύνη, τη ρομποτική, την κατανόηση της φυσικής γλώσσας, του σχεδιασμού και των κοινωνικών επιστημών.

Συγκεκριμένα ο τομέας της Συνεργασίας Ανθρώπου-Ρομπότ (Human-Robot Collaboration HRC) είναι η μελέτη διαδικασιών ανάμεσα σε ανθρώπους και ρομπότ που συνεργάζονται για την επίτευξη κοινών στόχων. Πολλές νέες εφαρμογές για ρομπότ απαιτούν να συνεργάζονται με άτομα ως ικανά μέλη ομάδων. Υπάρχει ένα αυξανόμενο ενδιαφέρον στον τομέα της συνεργατικής ρομποτικής ώστε να γίνει δυνατό για τους ανθρώπους να διδάσκουν ρομπότ διαφορετικούς τύπους δεξιοτήτων [1]. Προκειμένου τα ρομπότ να μάθουν από την ανθρώπινη επίδειξη, πρέπει πρώτα να είναι σε θέση να μάθουν να αναγνωρίζουν νόημα της κάθε ενέργειας για την επίτευξη συγκεκριμένου στόχου. Ο

¹ <http://www.robotiksistem.com/robotinfo.html>

πιο αποτελεσματικός τρόπος για να γίνει αυτό είναι να εξοπλίσει το ρομπότ με την ικανότητα να κατανοεί τις φυσικές ενέργειες του ανθρώπου, τη λεκτική καθώς και μη λεκτική επικοινωνία του [2]. Στη συνέχεια, πρέπει να μάθει να αναγνωρίζει διάφορες εργασίες μέσω των αντικειμένων που χρησιμοποιούνται σε αυτές. Αυτό θα βοηθήσει στην υλοποίηση ενός προγράμματος εργασιών που μπορεί να χρησιμοποιήσει για να κατανοήσει πώς μπορούν να χρησιμοποιηθούν διαφορετικές ενέργειες για την επίτευξη του ίδιου στόχου [3]. Τέλος, πρέπει να είναι σε θέση να μάθει πώς να χρησιμοποιεί τις ανθρώπινες ενέργειες με το δικό του σώμα με σκοπό να πετύχει το ίδιο αποτέλεσμα με τον άνθρωπο. Υπάρχουν παραδείγματα συστημάτων που μαθαίνουν ταυτόχρονα τόσο στόχους επιπέδου εργασίας όσο και δράσεις επιπέδου κινητήρων από ανθρώπινους δασκάλους [4]. Σε ένα δυναμικό περιβάλλον, η μίμηση δεν είναι αρκετή, αλλά το ρομπότ πρέπει να μιμηθεί αυτό που παρατηρεί για να κατανοήσει πώς να τροποποιήσει τις ενέργειές του σε μελλοντικές ελαφρώς διαφορετικές καταστάσεις [5]. Παρέχοντας στα ρομπότ την ικανότητα να κατανοούν τις ανθρώπινες ενέργειες και τις επικοινωνιακές δεξιότητες, μπορούμε να αναπτύξουμε ρομπότ που συνεργάζονται με ανθρώπους σε διάφορες εργασίες. Σε ένα σενάριο όπου το ρομπότ βοηθά έναν άνθρωπο στη συναρμολόγηση, πρέπει να ξέρει πότε να παρέχει βοήθεια. Αυτή η βοήθεια μπορεί να είναι είτε η ανταπόκριση σε ένα αίτημα για βοήθεια από τους χρήστες είτε η αυτόματη ανίχνευση του πότε πρέπει να βοηθήσει [6]. Οι άνθρωποι μπορούν να βοηθήσουν τα ημιαυτόνομα ρομπότ συναρμολόγησης με εργασίες χαμηλού επιπέδου, όπως ανίχνευση αντικειμένου και προσαρμογή κατά την εκτέλεση των δράσεων, ή εργασίες υψηλού επιπέδου όπως αποδοχή δράσης και αλλαγή της ακολουθίας εργασιών [7].

Η Συνεργασία Ανθρώπου-Ρομπότ απαιτεί επιπλέον από τα δύο μέρη να αναφέρονται σε αντικείμενα στον κοινόχρηστο χώρο. Οι άνθρωποι μπορούν να χρησιμοποιήσουν έναν συνδυασμό λεκτικών περιγραφών, χειρονομίες και βλέμματα για να ξεχωρίσουν ένα από τα πολλά αντικείμενα [8]. Ένας άνθρωπος που εκτελεί μια εργασία μπορεί να μην μπορεί να χρησιμοποιήσει χειρονομίες για να ζητήσει κάποιο αντικείμενο από ένα ρομπότ. Ωστόσο, δεδομένου ότι οι λεκτικές περιγραφές μπορεί να είναι δύσκολο να αποσαφηνιστούν, το ρομπότ χρειάζεται τρόπους για να επιτύχει κοινό έδαφος για το αντικείμενο που στοχεύει ο χρήστης [9]. Οι δυσκολίες αυτές μπορούν να επιλυθούν με λεκτική γείωση (verbal grounding) [10], με υπόδειξη [11] ή χρησιμοποιώντας βλέμμα [12].

Φυσικά, τα ρομπότ μπορούν επίσης να χρησιμοποιούν μη ανθρώπινους τρόπους για να δείξουν τον σκοπό τους, όπως η προβολή των προθέσεων τους στο κοινό περιβάλλον [13] ή χρησιμοποιώντας ελαυξημένη ή μικτή πραγματικότητα [14].

1.1.2 Ασφάλεια στην συνεργασία ανθρώπου ρομπότ.

Αναπόφευκτα, το ζήτημα της ασφάλειας κατά τη συνεργασία ανθρώπων και ρομπότ, είναι πρωτεύουσας σημασίας. Σύμφωνα με τα συμβατικά πρότυπα της βιομηχανίας, τα ρομπότ οφείλουν να λειτουργούν κάτω από συγκεκριμένες συνθήκες, μακριά από ανθρώπους και συχνά πίσω από προστατευτικούς φράχτες και εμπόδια.

Η παρουσία ανθρώπων στο χώρο δράσης των ρομπότ απαγορεύεται ρητά κατά τη λειτουργία τους, καθώς οι ταχύτητες με τις οποίες κινούνται, σε συνδυασμό με το πολλές φορές μεγάλο μέγεθος και βάρος τους, αποτελούν σημαντικό κίνδυνο για τη σωματική ακεραιότητα των εργαζόμενων.

Στα συνεργατικά περιβάλλοντα όμως, οι άνθρωποι και ρομπότ καλούνται να εργαστούν στον ίδιο χώρο και να εκτελούν παράλληλα εργασίες. Ακόμα, στα πλαίσια της αποδοτικής και διαισθητικής συνεργασίας, οι χειριστές των ρομπότ χρειάζεται να έρθουν σε φυσική επαφή μαζί τους, για να τα καθοδηγήσουν ή για να τους δώσουν κάποια εντολή. Επομένως, η ανάπτυξη επιπλέον μέτρων και προτύπων είναι αναγκαία για την εγγύηση της ασφάλειας των εργαζόμενων. Η έρευνα στον συγκεκριμένο τομέα είναι επίσης έντονη και συνεχώς ενεργές είναι οι διαδικασίες θέσπισης προτύπων ασφάλειας από διεθνείς φορείς και οργανισμούς. Οι βασικές μέθοδοι που χρησιμοποιούνται για την ανάπτυξη συστημάτων ασφάλειας είναι:

- Ο έλεγχος της ταχύτητας του ρομπότ ανάλογα με την απόσταση του ανθρώπου [15]. Πολλές φορές ορίζονται ζώνες επικινδυνότητας γύρω από το ρομπότ, η είσοδος στις οποίες, προκαλεί σε πρώτο επίπεδο τη μείωση της ταχύτητας του ρομπότ (reduced mode) μέχρι και την πλήρη διακοπή της κίνησής του (emergency mode) στις εγγύτερες ζώνες.
- Η μεταβολή της τροχιάς κίνησης των ρομπότ ανάλογα με τη θέση του ανθρώπου (reactive motion planning) [15]. Σύνθετοι αλγόριθμοι πραγματικού χρόνου χρησιμοποιούνται για τον εντοπισμό πιθανών εμποδίων στην τροχιά της κίνησης και την άμεση μεταβολή όταν αυτό κρίνεται αναγκαίο.

- Άμεση διακοπή της κίνησης σε περίπτωση ανίχνευσης σύγκρουσης. Σε αντίθεση με τις προηγούμενες μεθόδους που αφορούν την αποφυγή της σύγκρουσης (precollision), σε αυτή την περίπτωση καθορίζεται η συμπεριφορά του ρομπότ μετά, ή κατά, της σύγκρουσης (postcollision) [16]. Τα ρομπότ εφοδιάζονται με αισθητήρια όργανα που μπορούν να ανιχνεύσουν την επαφή, με αισθητήρες δύναμης (force sensors), είτε με χωρητικούς αισθητήρες (capacitive sensors), οι οποίοι αντιλαμβάνονται ηλεκτρική αγωγιμότητα σε πολύ κοντινή τους απόσταση.

Σε κάθε περίπτωση, οι συσκευές (αισθητήρες, λογικοί ελεγκτές, κλπ) και το λογισμικό των συστημάτων ασφάλειας, καθώς και η μεταφορά δεδομένων στο εσωτερικό τους αλλά και με τρίτες συσκευές υπόκεινται σε αυστηρά πρότυπα, και οφείλουν να περνούν συστηματικούς ελέγχους και να λαμβάνουν πιστοποιήσεις. Η αξιοπιστία και η εγγύηση της συνεπούς λειτουργίας τους είναι απαραίτητη, καθώς η απουσία της μπορεί να αποβεί μοιραία τόσο για τους εργαζόμενους όσο και για τον εξοπλισμό.

1.1.3 Συνεισφορά

Η ανάπτυξη μεθόδων που επιτρέπουν την αλληλεπίδραση μεταξύ του ανθρώπινου δυναμικού με τον ρομποτικό εξοπλισμό αποτελεί βασική προτεραιότητα της βιομηχανίας, καθώς όπως έχει παρατηρηθεί, το κομμάτι του HRC μπορεί να συμβάλλει σημαντικά στην ασφαλή χρήση των ρομπότ σε ένα κοινό εργασιακό περιβάλλον.

Η παρούσα εργασία συνεισφέρει στο κομμάτι της πρόβλεψης της κίνησης ενός ανθρώπου από ένα ρομπότ. Προβλέποντας τις ενέργειες και τις κινήσεις του ανθρώπου, ένα ρομπότ μπορεί να σχεδιάσει την επόμενη κίνηση του. Αυτό είναι ιδιαίτερα χρήσιμο σε περιβάλλοντα που αλλάζουν συνεχώς, επομένως ο εκ το πρότερον προγραμματισμός κάποιων δεδομένων κινήσεων είναι αδύνατος. Με την ενσωμάτωση της γνώσης προβλέψεων και του προγραμματισμού με προληπτικό τρόπο, καθίσταται δυνατή η αποφυγή ατυχημάτων.

Μέσω της χρήσης είτε ρητών είτε έμμεσων ενδείξεων, το ρομπότ μπορεί να κάνει τους επιδιωκόμενους στόχους και κινήσεις του πιο ξεκάθαρες στους ανθρώπους όπου καλείται να συνεργαστεί, γεγονός που με τη σειρά του διευκολύνει την ικανότητα των ανθρώπων να επιλέγουν ενέργειες και κινήσεις που διατηρούν την ασφάλεια.

Τέλος μέσω της πρόβλεψης τα ρομπότ μπορούν επιπλέον να αποτρέψουν ατυχήματα που δεν σχετίζονται μόνο με την συνεργασία ανάμεσα σε αυτά και τον άνθρωπο αλλά και ατυχήματα που έχουν να κάνουν με την απροσεξία του ανθρώπου όπως αποτροπή επαφής με ηλεκτροφόρο καλώδιο.

1.2 Μηχανική μάθηση

1.2.1 Τι είναι η μηχανική μάθηση

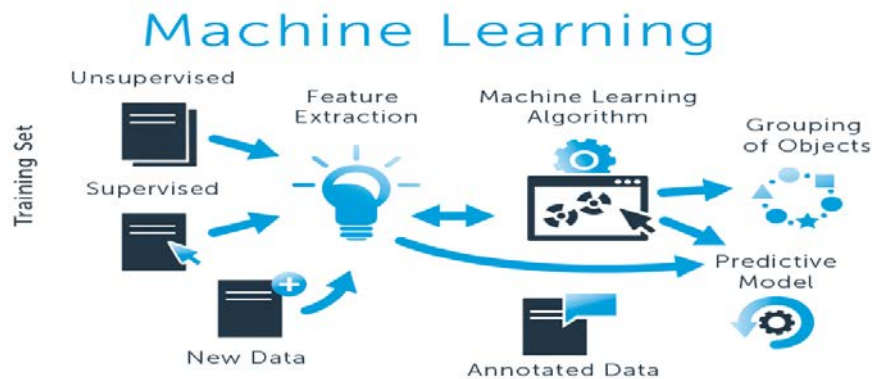
Είναι ένας από τους τομείς της επιστήμης των υπολογιστών που είναι συνδεδεμένος με την τεχνητή νοημοσύνη και έχει ως σκοπό την κατασκευή την χρήση αλγορίθμων για την μετατροπή δεδομένων σε μοντέλα που μπορούν στη συνέχεια να χρησιμοποιηθούν και στο μέλλον. Το χαρακτηριστικό της μηχανικής μάθησης είναι η ικανότητα του συστήματος με την χρήση αλγορίθμων να αποκτά γνώση από τα δεδομένα και να βελτιώνεται με σκοπό να προβεί σε προβλέψεις.[17]

Ο Arthur Samuel το 1959, πραγματοποίησε μία αναφερόμενη δήλωση για τη μηχανική μάθηση δηλώνοντας: «Η μηχανική μάθηση είναι το πεδίο σπουδών που δίνει στους υπολογιστές τη δυνατότητα να μάθουν χωρίς να προγραμματίζονται ρητά» [18]. Επίσης Ο Mitchell το 1997: «Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία E ως προς κάποια κλάση εργασιών T και μέτρο απόδοσης P , αν η απόδοσή του σε εργασίες από το T , όπως μετριέται από το P , βελτιώνεται μέσω της εμπειρίας E ». [18].

Οι επιστήμονες της Τεχνητής Νοημοσύνης (ΤΝ) αναφέρουν τη μηχανική μάθηση ως το φαινόμενο κατά το οποίο ένα σύστημα βελτιώνει την απόδοσή του κατά την εκτέλεση μίας συγκεκριμένης εργασίας, δίχως να υπάρχει ανάγκη να προγραμματιστεί από την αρχή. Βάσει του ορισμού αυτού, η μηχανική μάθηση επιδιώκει τη δημιουργία μηχανών που να είναι ικανές να μαθαίνουν, να βελτιώνουν την απόδοσή τους σε κάποιους τομείς μέσω της αξιοποίησης προηγούμενης γνώσης.

Η μηχανική μάθηση μπορεί να χρησιμοποιηθεί στις υπολογιστικές εφαρμογές, όπου ο σχεδιασμός και ο ρητός προγραμματισμός των αλγορίθμων είναι ανέφικτος. Διαφέρει από την εξόρυξη δεδομένων, η οποία επικεντρώνεται περισσότερο στην εξερευνητική ανάλυση των δεδομένων (μη επιβλεπόμενη μάθηση). Η μηχανική μάθηση βοηθάει στην ραγδαία εξέλιξη του πεδίου της ανάλυσης δεδομένων, διότι μέσω αυτής μπορούν να

κατασκευαστούν πολυπλοκότερα μοντέλα και αλγόριθμους για μεγάλους όγκους δεδομένων από ότι πριν. (εικόνα 1). Τα αναλυτικά μοντέλα επιτρέπουν στους ερευνητές, τους επιστήμονες δεδομένων, τους μηχανικούς και τους αναλυτές να παράγουν αξιόπιστες αποφάσεις και αποτελέσματα και να αναδείξουν αλληλοσχετίσεις μέσω της μάθησης από ιστορικές σχέσεις και τάσεις στα δεδομένα.



Εικόνα 1: Πλαίσιο μηχανικής μάθησης[19]

1.2.2 Ταξινόμηση της μηχανικής μάθησης

Υπάρχουν πολλές τεχνικές μηχανικής μάθησης που χρησιμοποιούνται διαφορετικά με βάση τη φύση του προβλήματος και χωρίζονται σε τρία είδη:

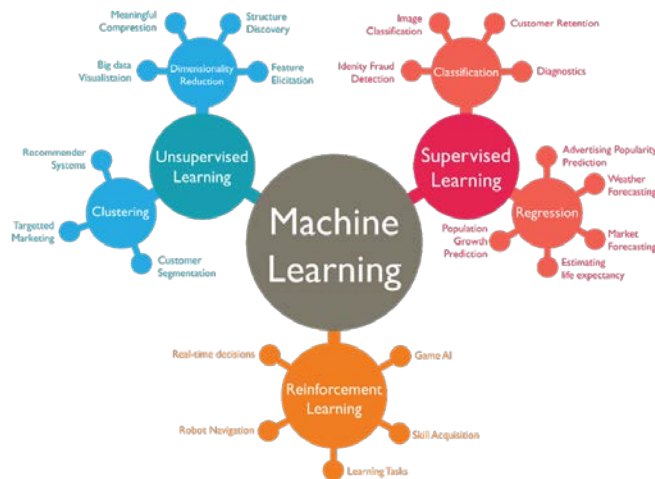
1. Ενισχυτική μάθηση (reinforcement learning)
2. Επιβλεπόμενη μάθηση (Supervised learning)
3. Μη επιβλεπόμενη μάθηση (Unsupervised learning)

Η διαφορά ανάμεσα στην επιβλεπόμενη και τη μη επιβλεπόμενη μάθηση έγκειται στο ότι επιβλεπόμενη μάθηση το σύστημα που προγραμματίζεται εκπαιδεύεται χρησιμοποιώντας σύνολα δεδομένων που γνωρίζουμε από πριν τις τιμές εξόδων. Στόχος της επιβλεπόμενης μάθησης είναι η δημιουργία μιας συνάρτησης με της οποίας να πλησιάζει όσο τον δυνατόν καλύτερα την σχέση εισόδου-εξόδου υπάρχει στο σύνολο δεδομένων που τις δίνεται. Η μη επιβλεπόμενη μάθηση δεν περιέχει πληροφορίες όσον αφορά τις τιμές εξόδων και στόχος είναι να βρίσκει συσχετίσεις χρησιμοποιώντας τη φυσική δομή που υπάρχει στο σύνολο δεδομένων που τις δίνεται. Επειδή η μη επιβλεπόμενη μάθηση βασίζεται στα δεδομένα και τις ιδιότητές τους, προκύπτει ότι η μη επιβλεπόμενη μάθηση έχει ως βάση τα δεδομένα (data driven). Τα αποτελέσματα από μια μη επιβλεπόμενη μαθησιακή εργασία ελέγχονται από τα δεδομένα και τον τρόπο με τον οποίο διαμορφώνονται

Οι κυριότερες τεχνικές Μηχανικής Μάθησης με επίβλεψη είναι:

- Μάθηση Εννοιών (Concept Learning)
- Δέντρα Απόφασης (Decision Trees)
- Μάθηση Κανόνων (Rule Learning)
- Μάθηση κατά Περίπτωση (Instance Based Learning)
- Μάθηση κατά Bayes
- Γραμμική Παρεμβολή (Linear Regression)
- Νευρωνικά Δίκτυα (Neural Networks)
- Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines)

Όσο αναφορά την ενισχυτική μάθηση, σε αυτή δίνεται η δυνατότητα ο αλγόριθμος να μαθαίνει μια στρατηγική ενεργειών μέσα από την αλληλεπίδραση με το περιβάλλον. Ο αλγόριθμος ενίσχυσης αναγνωρίζει δοκιμάζοντας διάφορους τρόπους για την επίλυση ενός προβλήματος και ενώ στην αρχή κάνει λάθη, στη συνέχεια ο αλγόριθμος μάθησης μαθαίνει να κάνει λιγότερα λάθη κάθε φορά που παρουσιάζεται το ίδιο πρόβλημα. Για το λόγο αυτό η ενισχυτική μάθηση είναι γνωστή και ως μάθηση από λάθη. Χρησιμοποιείται κυρίως σε προβλήματα σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ και η βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους.



Εικόνα 2: Κατηγορίες μηχανικής μάθησης [20]

1.2.3 Διαχωρισμός μηχανικής μάθησης με βάση το βέλτιστο αποτέλεσμα

Η μηχανική μάθηση χωρίζεται σε περιπτώσεις ανάλογα με το επιθυμητό αποτέλεσμα του συστήματος στα παρακάτω είδη:

- **Ομαδοποίηση (Clustering):** Η ομαδοποίηση (clustering) είναι μία τεχνική της μηχανικής μάθησης που περιλαμβάνει την ομαδοποίηση σημείων δεδομένων. Είναι ένας τύπος μεθόδου μάθησης χωρίς επίβλεψη. Χρησιμοποιείται ως διαδικασία για την εύρεση ουσιαστικής δομής, επεξηγηματικών υποκείμενων διαδικασιών, γενετικών χαρακτηριστικών και ομαδοποιήσεων που μπορεί να υπάρχουν σε ένα σύνολο παραδειγμάτων. Δεδομένου ενός συνόλου σημείων δεδομένων, μπορούμε να χρησιμοποιήσουμε έναν αλγόριθμο ομαδοποίησης για να ταξινομήσουμε κάθε σημείο δεδομένων σε μια συγκεκριμένη ομάδα. Θεωρητικά, τα σημεία δεδομένων που βρίσκονται στην ίδια ομάδα πρέπει να έχουν παρόμοιες ιδιότητες ή χαρακτηριστικά, ενώ τα σημεία δεδομένων σε διαφορετικές ομάδες θα πρέπει να έχουν πολύ ανόμοιες ιδιότητες ή χαρακτηριστικά. Στην Επιστήμη Δεδομένων, μπορεί να χρησιμοποιηθεί η ανάλυση συμπλέγματος για να εξορυχθούν πολύτιμες πληροφορίες από τα δεδομένα, βλέποντας σε ποιες ομάδες εμπίπτουν τα σημεία δεδομένων όταν εφαρμοστεί ένας αλγόριθμος ομαδοποίησης.
- **Μείωση διαστασιμότητας (Dimensionality Reduction):** Είναι η διαδικασία με την οποία επιτυγχάνεται η μείωση του αριθμού των υπό εξέταση χαρακτηριστικών μεταβλητών του συνόλου δεδομένων. Καθώς ο αριθμός των χαρακτηριστικών αυξάνεται, το μοντέλο γίνεται πιο περίπλοκο. Όσο υψηλότερος είναι ο αριθμός των χαρακτηριστικών, τόσο πιο δύσκολα γίνεται η απεικόνιση του συνόλου εκπαίδευσης και στη συνέχεια και η επεξεργασία του. Ένα μοντέλο μηχανικής εκμάθησης που εκπαιδεύεται σε μεγάλο αριθμό χαρακτηριστικών, εξαρτάται όλο και περισσότερο από τα δεδομένα στα οποία εκπαιδεύτηκε και με τη σειρά του υπερ-εκπαιδεύει, με αποτέλεσμα την κακή απόδοση σε πραγματικά δεδομένα, ξεπερνώντας τον σκοπό του. Όσο λιγότερα χαρακτηριστικά διαθέτει το σύνολο δεδομένων εκπαίδευσης, τόσο λιγότερες υποθέσεις κάνει το μοντέλο και τόσο απλό γίνεται το μοντέλο.

Με τη μείωση διαστάσεων επιτυγχάνονται τα εξής:

1. Βελτίωση ακρίβειας μοντέλου
 2. Μικρότερη υπολογιστική δύναμη
 3. Εκπαίδευση αλγορίθμων με πιο γρήγορους ρυθμούς
 4. Μεγαλύτερος αποθηκευτικός χώρος
 5. Αφαίρεση περιττών θορύβων και λειτουργιών.
- Εκτίμηση πυκνότητας (Density Estimation): είναι η κατασκευή ενός εκτιμητή της συνάρτησης πυκνότητας με βάση την κατανομή των δεδομένων.
 - Παλινδρόμηση (Regression): Η ανάλυση παλινδρόμησης (regression) χρησιμοποιείται σε στατιστικά στοιχεία για την εύρεση τάσεων στα δεδομένα. Πιο συγκεκριμένα είναι μια αξιόπιστη μέθοδος προσδιορισμού των μεταβλητών που επηρεάζουν ένα θέμα ενδιαφέροντος. Η διαδικασία εκτέλεσης μιας παλινδρόμησης επιτρέπει να προσδιοριστεί με σιγουριά ποιοι παράγοντες έχουν μεγαλύτερη σημασία, ποιοι παράγοντες μπορούν να αγνοηθούν και πώς αυτοί οι παράγοντες επηρεάζουν ο ένας τον άλλον.
 - Κατηγοριοποίηση (Classification): Η κατηγοριοποίηση (classification) είναι ένας τύπος εποπτευόμενης μάθησης και αποτελεί μία από τις πιο διαδεδομένες ενέργειες της μηχανικής μάθησης. Καθορίζει την κλάση στην οποία ανήκουν τα στοιχεία δεδομένων και χρησιμοποιείται καλύτερα όταν η έξοδος έχει πεπερασμένες και διακριτές τιμές. Πιο αναλυτικά τα δεδομένα εισόδου χωρίζονται σε δύο ή περισσότερες κλάσεις, και η μηχανή πρέπει να κατασκευάσει ένα μοντέλο, το οποίο θα αντιστοιχίζει τα δεδομένα σε μία ή περισσότερες (multi-label) κλάσεις, συνεπώς αφορά στη δημιουργία μοντέλων πρόβλεψης διακριτών τάξεων (κλάσεων/κατηγοριών).

1.3 Αλγόριθμοι μηχανικής μάθησης

Σε αυτό το κεφάλαιο θα αναλυθεί ο τρόπος λειτουργίας των αλγορίθμων μηχανικής μάθησης που θα χρησιμοποιηθούν για την υλοποίηση της διπλωμάτης εργασίας.

1.3.1 Μηχανισμοί διανυσμάτων στήριξης - **Support vector machines**

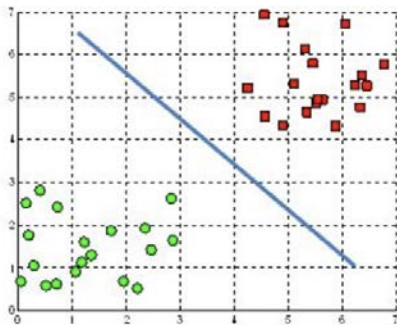
Οι μηχανισμοί διανυσμάτων στήριξης (SVM) χρησιμοποιούνται για την αντιμετώπιση πολλών προβλημάτων. Πιο συγκεκριμένα, χρησιμοποιούνται ευρέως στην κατηγοριοποίηση ενός κειμένου και την ταξινόμηση των εικόνων. Σε αποτελέσματα πειραμάτων διαπιστώνουμε, ότι τα SVM επιτυγχάνουν υψηλότερη ακρίβεια και τα συστήματα κατακερματισμού εικόνας χρησιμοποιούν μια τροποποιημένη έκδοση του αλγορίθμου. Ακόμη, στην αναγνώριση χειρόγραφων τα SVM έχουν επιδείξει μεγάλη ακρίβεια. Ένας τομέας ο οποίος έχει επωφεληθεί πολύ από αυτά είναι ο τομέας της βιολογίας στον οποίο έχουν πετύχει να ταξινομήσουν έως και το 90% των ενώσεων των πρωτεϊνών[21].

Η ταξινόμηση των δεδομένων χρησιμοποιείται ευρέως στην μηχανική μάθηση. Η κύρια ιδέα γύρω από τους μηχανισμούς διανυσμάτων υποστήριξης SVM είναι η ίδια με τα πολλαπλών στρώσεων νευρωνικά δίκτυα. Το πιο βασικό χαρακτηριστικό των SVM είναι οι συναρτήσεις πυρήνα (kernel Function). Το χαρακτηριστικό αυτό μπορεί να είναι γραμμικές, πολυωνυμικές και εκθετικές συναρτήσεις. Πρόκειται για συναρτήσεις, οι οποίες παίρνουν σαν είσοδο, χώρους χαμηλών διαστάσεων και τους μετασχηματίζουν σε χώρους υψηλότερων διαστάσεων, δηλαδή μετατρέπουν μη διαχωρίσιμα προβλήματα σε διαχωρίσιμα μέσω εξαιρετικά πολυσύνθετων μετασχηματισμών. Η διαδικασία που ακολουθείται για να επιτευχθεί αυτό είναι αρχικά η εύρεση του βέλτιστου υπερ-πλαισίου (hyperplane) για τα γραμμικά διαχωρίσιμα μοτίβα. Έπειτα, γίνεται η επέκταση σε μοτίβα που δεν είναι γραμμικά διαχωρίσιμα με μετασχηματισμούς των αρχικών δεδομένων, ώστε να χαρτογραφηθούν στο νέο χώρο. Αυτό είναι γνωστό σαν το τέχνασμα του πυρήνα (kernel trick) [22].

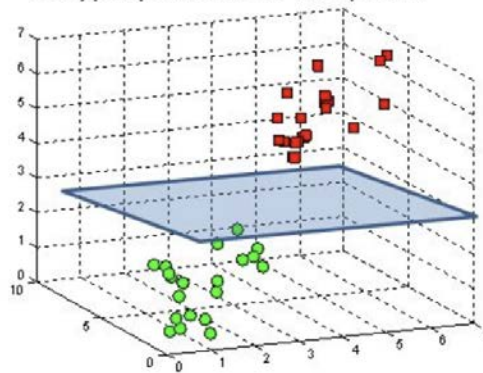
Ας υποθέσουμε, ότι ορισμένα δεδομένα - σημεία ανήκουν σε μία από τις δύο κλάσεις και ο στόχος είναι να αποφασιστεί σε ποια κατηγορία θα ταξινομηθεί ένα νέο σημείο. Στην περίπτωση του SVM αλγόριθμου, ένα σημειακό δεδομένο θεωρείται σαν ένα p -διάστατο

διάνυσμα (μία λίστα από p αριθμούς). Το σημαντικό θέμα είναι αν υπάρχει η δυνατότητα διαχωρισμού του συγκεκριμένου σημείου σ' ένα $(p-1)$ -διάστατο υπερπλάνιο. Στη γεωμετρία με τον όρο υπερπλάνιο (hyperplane) αναφερόμαστε σ' έναν υπόχωρο, του οποίου η διάσταση είναι κατά μία λιγότερη από τον περιβάλλοντα χώρο του. Ουσιαστικά, τα υπερπλάγια είναι όρια απόφασης, που βοηθούν στην ταξινόμηση των νέων σημείων. Τα σημεία που εμπίπτουν σε κάθε πλευρά του υπερπλαισίου μπορούν να αποδοθούν σε διαφορετικές κατηγορίες. Επίσης, η διάσταση του υπερπλαισίου εξαρτάται από τον αριθμό των χαρακτηριστικών (features). Εάν ο αριθμός των χαρακτηριστικών εισόδου είναι 2, τότε το υπερπλάνιο είναι μόνο μια γραμμή. Εάν ο αριθμός των χαρακτηριστικών εισόδου είναι 3, τότε το υπερπλάνιο γίνεται ένα δισδιάστατο επίπεδο. Είναι δύσκολο να φανταστεί κάποιος τον χώρο, όταν ο αριθμός των χαρακτηριστικών ξεπερνά τα τρία.

A hyperplane in \mathbb{R}^2 is a line



A hyperplane in \mathbb{R}^3 is a plane



Εικόνα 3: Υπερπλάγια σε χώρους χαρακτηριστικών 2 και 3 διαστάσεων[49].

Τα διανύσματα στήριξης είναι σημεία, που είναι πλησιέστερα στο υπερπλάνο και επηρεάζουν τη θέση και τον προσανατολισμό του ίδιου του υπερπλαισίου. Χρησιμοποιώντας αυτούς τους φορείς στήριξης, μεγιστοποιούμε το περιθώριο (margin) του ταξινομητή. Συνεπώς, η διαγραφή των διανυσμάτων στήριξης θα αλλάξει τη θέση του υπερπλαισίου. Αυτά είναι τα σημεία που μας βοηθούν να κατασκευάσουμε τον SVM. [23] [24]

Υπάρχουν αρκετοί ρυθμιστικοί παράγοντες σχετικά με τον αλγόριθμο SVM. Αρχικά, πρέπει να αναφερθεί η παράμετρος κανονικοποίησης, ή αλλιώς παράμετρος C , η οποία ουσιαστικά καθορίζει πόσο θέλουμε να αποφευχθεί η εσφαλμένη ταξινόμηση κάθε δείγματος εκπαίδευσης. Για τις μεγάλες τιμές του C , η βελτιστοποίηση θα επιλέξει ένα υπερπλάνο μικρότερου περιθωρίου (smaller - margin hyperplane). Αντίθετα, μια πολύ μικρή τιμή του C θα αναγκάσει το βελτιστοποιητή να αναζητήσει ένα μεγαλύτερου περιθωρίου υπερπλάνο (larger-margin hyperplane), ακόμη και αν αυτό ταξινομεί εσφαλμένα περισσότερα σημεία. Μια καλή επιλογή για το βέλτιστο υπερπλάνο είναι αυτή που επιτρέπει το μεγαλύτερο διαχωρισμό ή περιθώριο (margin) μεταξύ των δύο τάξεων. Έτσι, επιλέγεται το υπερπλάνο, ώστε η απόσταση από αυτό στο κοντινότερο σημείο από κάθε πλευρά να μεγιστοποιείται. Ένα τέτοιο υπερπλάνο ονομάζεται υπερπλάνο ανώτατου περιθωρίου (maximum -margin hyperplane) και ο γραμμικός ταξινομητής που ορίζει είναι γνωστός ως ταξινομητής μέγιστου περιθωρίου, ή αλλιώς νευρώνας perceptron βέλτιστης σταθερότητας.



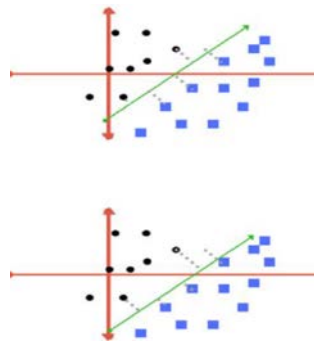
Εικόνα 4: Αριστερά χαμηλή τιμή κανονικοποίησης ενώ δεξιά υψηλή [50].

Μια ακόμη παράμετρος σημαντική για τον SVM είναι η παράμετρος Γάμμα η οποία ορίζει το κατά ποσό ο αλγόριθμος θα επηρεάζεται από ένα μοναδικό σημείο. Για μικρό Γάμμα, τα σημεία που είναι μακριά από την πιθανή γραμμή διαχωρισμού χρησιμοποιούνται στον υπολογισμό αυτής σε αντίθεση με ένα μεγάλο Γάμμα που χρησιμοποιούνται μόνο τα κοντινά σημεία αυτής.



Εικόνα 5: Στην αριστερή εικόνα με υψηλό γάμμα λαμβάνονται υπόψη μόνο τα κοντινά σημεία, ενώ στην δεξιά με χαμηλό γάμμα συνεισφέρουν και τα μακρινά[50].

Η τελευταία παράμετρος για τον SVM είναι το περιθώριο(margin). Ένα καλό περιθώριο είναι αυτό που διαχωρίζει τα περισσότερα σημεία και για τις δύο τάξεις καθώς και επιτρέπει στα σημεία να είναι στις σωστές τάξεις χωρίς να διασχίζουν την άλλη κατηγορία. [24]



Εικόνα 6: Το πρώτο «καλό» περιθώριο απέχει εξίσου από τις δύο κλάσεις, αντίθετα το δεύτερο «κακό» περιθώριο βρίσκεται πολύ κοντά στη μπλε κλάση[50].

Όπως αναφέρθηκε και προηγουμένως, όσο μεγαλύτερο είναι το περιθώριο, τόσο μικρότερο είναι το σφάλμα γενίκευσης του αλγορίθμου. Η συνάρτηση απώλειας, που βοηθάει να μεγιστοποιηθεί το περιθώριο, είναι η συνάρτηση εύκαμπτου συνδέσμου (hinge loss function).

$$l(y) = \max (0.1 - t * y)$$

1.3.2 Δέντρα αποφάσεων – Decision tree

Τα δέντρα ταξινόμησης χρησιμοποιούνται ως αλγόριθμοι μηχανικής μάθησης. Σε αυτή την μέθοδο χρησιμοποιούμε ως είσοδο ένα διάνυσμα τιμών, με κάποιες ιδιότητες, και σαν έξοδο ο αλγόριθμος μας επιστρέφει μια τιμή που μπορεί να είναι είτε διακριτή είτε συνεχής. Όταν η έξοδος είναι διακριτή το πρόβλημα μας κατατάσσεται στη κατηγορία ταξινόμησης. Όταν η έξοδος δεν είναι μια διακριτή τιμή τότε το πρόβλημα μας κατατάσσεται σαν πρόβλημα παλινδρόμησης[25].

Τα δέντρα ταξινόμησης χρησιμοποιούνται για την πρόβλεψη της τιμής της μεταβλητής που θέλουμε να μοντελοποιήσουμε, έχοντας ως βάση τις τιμές των θεωρούμενων ανεξάρτητων μεταβλητών. Έτσι ένα δέντρο απόφασης αναπαριστά μια διαδικασία λήψης απόφασης, όπου για κάθε πιθανό σημείο ή κατάσταση έχουμε ένα κόμβο. Για κάθε επιλογή τώρα που μπορεί να γίνει σε ένα σημείο απόφασης αναπαριστάτε ένα «κόμβο-παιδί». Κάθε κόμβος ορίζει μια συνθήκη ελέγχου της τιμής κάποιου χαρακτηριστικού των περιπτώσεων. Κάθε κλαδί που φεύγει από ένα κόμβο αντιστοιχεί σε μια διαφορετική διακριτή τιμή του χαρακτηριστικού που σχετίζεται με τον κόμβο.

Στα κλαδιά καταλήγουν οι τελικοί κόμβοι που ανήκουν σε ένα μόνο σύνολο, όπου είναι οι τελικές αποφάσεις ή ενέργειες.

Για να διασπάσουμε ένα κόμβο (parent) με N εγγραφές σε k παιδιά u_i , ο αριθμός των εγγραφών είναι $N(u_i)$ με:

$$N(u_i) \text{ με } \sum N(u_i) = N$$

Για να επιλεγθεί η διάσπαση με το μεγαλύτερο κέρδος, υπολογίζουμε το μέγιστο Δ από την σχέση

$$\Delta = I(\text{parent}) - \sum_{i=1}^k \frac{N(u_i)}{N} * i(u_i)$$

Τα μέτρα μη-καθαρότητας στην κατάσταση ενός δέντρου είναι τα εξής :

1) η εντροπία , 2) το ευρετήριο Gini και 3) το λάθος ταξινόμησης.

1) Η Εντροπία (Entropy), ορίζει το μετρώ της αταξίας ενός συστήματος και ορίζεται από τη εξής σχέση:

$$E(S) = - \sum_{i=1}^c p_i * \log_2(p_i)$$

Όπου S είναι ένα σύνολο δεδομένων pI που είναι το ποσοστό των παραδειγμάτων του S που ανήκουν στην κατηγορία I και c διαφορετικές κατηγορίες.

Το κέρδος πληροφορίας (information gain) αναφέρεται στη μείωση εντροπίας ενός συνόλου εκπαίδευσης S με παράμετρο διαχωρισμού A. Τονίζεται ότι όταν η εντροπία πληροφορίας (ή πληροφοριακή εντροπία) μειώνεται τότε η πυκνότητα πληροφορίας αυξάνεται, δηλαδή η περιγραφή γίνεται περισσότερο συμπαγής.

Η εντροπία βοηθά ώστε να προσδιοριστεί η κατάλληλη μεταβλητή που οδηγεί σε περισσότερο συμπαγές δέντρο.

Το κέρδος πληροφορίας G (S,A) ορίζεται ως

$$G(S, A) = E(S) - \sum_{V(A)} |S_u| * \frac{E(S_u)}{|S|}$$

Όπου E(S) είναι η εντροπία πληροφορίας του υπό εξέταση κόμβου, A είναι μια ανεξάρτητη μεταβλητή, με τιμές V(A) με τις οποίες γίνεται ο επόμενος διαχωρισμός, u είναι μια από τις (δυνατές) τιμές A και Su είναι το πλήθος των εγγραφών (με A=u) και E(SU) είναι η εντροπία πληροφορίας του υπό εξέταση κόμβου ως προς τη τιμή A=u.

2) Το ευρετήριο Gini για κάθε κόμβο t , ορίζεται ως:

$$GINI(t) = 1 - \sum_{j=1}^c [p(\frac{j}{t})]^2$$

Όπου p (j / t) σχετική συχνότητα της κλάσης j στο κόμβο t (ποσοστό εγγραφών της κλάσης j στο κόμβο t) και c αριθμός κλάσεων. Το αναμενόμενο κέρδος τώρα, σε πληροφορία μετά τον έλεγχο στο χαρακτηριστικό A ισούται με την πληροφορία που μας έλειπε πριν τον έλεγχο μείον τη αναμενόμενη πληροφορία που θα μας λείπει μετά τον έλεγχο.

Η ελάχιστη τιμή που μπορεί να πάρει είναι (0.0) , όταν όλες οι εγγραφές ανήκουν σε μια κλάση, ενώ η μέγιστη τιμή (1-1/c), όταν όλες οι εγγραφές είναι ομοιόμορφα κατανομημένες στις κλάσεις.

Στους αλγορίθμους CART,SLIQ, και SPRINT χρησιμοποιείται ο παρακάτω τύπος

$$GINI_{split} = \sum_{i=1}^k \frac{n_i}{n} * GINI(i)$$

Όπου n_i είναι ο αριθμός εγγραφών του παιδιού i και n είναι ο αριθμός εγγραφών του κόμβου p . Αυτός ο τύπος χρησιμοποιείται για την ποιότητα διαχωρισμού εγγραφών, όταν ένας κόμβος p διασπάται σε k κόμβους.

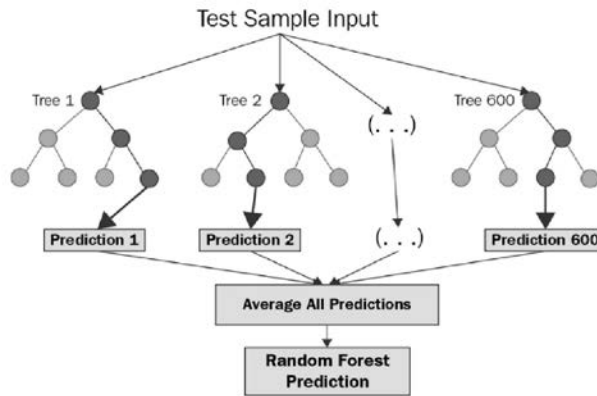
1.3.3 Τυχαία δάση - Random forests

Τα τυχαία δάση αποφάσεων αποτελούν επίσης μια μέθοδος μάθησης. Χρησιμοποιούνται τόσο για την ταξινόμηση όσο και την παλινδρόμηση. Ο τρόπος που λειτουργούν κατά την εκπαίδευση είναι η κατασκευή πολλαπλών δέντρων αποφάσεων. Μέσα από την διαδικασία παραγωγής αυτών των δέντρων εξάγουν τη σωστή κλάση ή τη μέση πρόβλεψη των επιμέρους δένδρων (παλινδρόμηση). Τα τυχαία δάση αποφάσεων δημιουργήθηκαν με σκοπό να διορθώσουν το πρόβλημα των δέντρων αποφάσεων, δηλαδή την υπερπροσαρμογή (overfitting) του συνόλου εκπαίδευσης (training set data).

Ο πρώτος αλγόριθμος για τα τυχαία δάση αποφάσεων δημιουργήθηκε από τον Tin Kam Ho, χρησιμοποιώντας την τυχαία μέθοδο του υποχώρου , η οποία στη διατύπωση του Ho είναι ένας τρόπος να εφαρμοστεί η προσέγγιση "στοχαστικής διάκρισης" στην ταξινόμηση που προτείνει ο Eugene Kleinberg.

Ο Ho διαπίστωσε, ότι τα δάση των οποίων τα δέντρα χωρίζονται με λοξά υπερ-πλαίσια, μπορούν να αποκτήσουν αρκετά μεγάλη ακρίβεια, καθώς μεγαλώνουν χωρίς να υποφέρουν από υπερπροσαρμογή, αρκεί τα δάση να χωρίζονται τυχαία, ώστε να επηρεάζονται μόνο από επιλεγμένα χαρακτηριστικά. Σημαντικό είναι το γεγονός, ότι ένας πιο περίπλοκος ταξινομητής, για παράδειγμα ένα μεγαλύτερο δάσος, γίνεται πιο ακριβής σχεδόν μονοτονικά και αυτό είναι σε πλήρη αντίθεση με την κοινή πεποίθηση ότι η πολυπλοκότητα ενός ταξινομητή μπορεί να αυξηθεί μόνο σ' ένα ορισμένο επίπεδο ακρίβειας, πριν μειωθεί εξαιτίας της υπερπροσαρμογής. Η ερμηνεία της αντοχής της

δασικής μεθόδου στην υπερπροσαρμογή μπορεί να βρεθεί στη θεωρία των στοχαστικών διακρίσεων του Kleinberg. [26]



Εικόνα 7: Δομή του ταξινομητή Random Forest[51]

Το τυχαίο δάσος, αποτελείται από ένα πλήθος ατομικών δέντρων αποφάσεων, που λειτουργούν ως σύνολο. Κάθε μεμονωμένο δέντρο στο τυχαίο δάσος κάνει μια πρόβλεψη κλάσης και η κλάση με τις περισσότερες ψήφους γίνεται η τελική πρόβλεψη του μοντέλου. Η ισχύς του πλήθους είναι η βασική ιδέα πίσω από τα τυχαία δάση. Σύμφωνα με την επιστήμη δεδομένων, ο λόγος που το μοντέλο των τυχαίων δασών λειτουργεί τόσο καλά είναι ότι ένας αρκετά μεγάλος αριθμός ασυσχέτιστων μοντέλων, συγκεκριμένα δέντρων, μπορούν να παράγουν προβλέψεις που είναι ακριβέστερες από οποιαδήποτε μεμονωμένη πρόβλεψη. Η επιτυχία του μοντέλου είναι η χαμηλή συσχέτιση μεταξύ των δέντρων. Ο λόγος γι' αυτό το αποτέλεσμα είναι, ότι τα δέντρα προστατεύονται μεταξύ τους από τα ατομικά λάθη, εφόσον δεν κάνουν πάντα λάθος στην ίδια κατεύθυνση. Αναλυτικότερα, ενώ το αποτέλεσμα κάποιων δέντρων είναι λανθασμένο, πολλών άλλων είναι σωστό. Όπως είναι λογικό, η μέθοδος του τυχαίου δάσους πρέπει να διασφαλίσει, ότι η συμπεριφορά κάθε δέντρου δεν είναι αρκετά συσχετισμένη με τη συμπεριφορά οποιουδήποτε άλλου δέντρου στο ίδιο μοντέλο. [27]

Σε αυτό το σημείο θα πρέπει να αναφερθούν μερικές σημαντικές τεχνικές. Αρχικά, θα περιγράψουμε την μέθοδο bootstrapping, η οποία είναι μια τεχνική στατιστικής αναδειγματοληψίας, που περιλαμβάνει τυχαία δειγματοληψία ενός συνόλου δεδομένων με αντικατάσταση. Χρησιμοποιείται συχνά ως μέσο ποσοτικοποίησης της αβεβαιότητας που συνδέεται μ' ένα μοντέλο μηχανικής μάθησης. Το bootstrapping είναι εξαιρετικά

χρήσιμο, καθώς επιτρέπει τη δημιουργία νέων δειγμάτων από έναν πληθυσμό χωρίς να χρειάζεται να προχωρήσει και να συλλέξει πρόσθετα "στοιχεία εκπαίδευσης". Η διαδικασία είναι να επαναλαμβάνεται η δειγματοληψία των δεδομένων με αντικατάσταση από το αρχικό σύνολο εκπαίδευσης, ώστε να παράγονται πολλαπλά ξεχωριστά υποσύνολα εκπαίδευσης. Στη συνέχεια αυτά χρησιμοποιούνται, για να επιτρέψουν τις μεθόδους μετα-μάθησης (meta-learner), να μειώσουν τη διακύμανση των προβλέψεων, βελτιώνοντας έτσι σημαντικά την απόδοση της πρόβλεψης. Είναι αξιόλογο, να αναφέρουμε, ότι σ' αυτήν την περίπτωση δεν υποδιαιρούμε το σύνολο εκπαίδευσης σε μικρότερα τμήματα, για να εκπαιδεύουμε κάθε δέντρο σε διαφορετικό κομμάτι. Αντίθετα, αν έχουμε ένα δείγμα μεγέθους N , κρατάμε το πλήθος σταθερό και συνεχίζουμε να εκπαιδεύουμε κάθε δέντρο με σύνολο εκπαίδευσης μεγέθους N . Αλλά αντί για τα αρχικά δεδομένα εκπαίδευσης, παίρνουμε ένα τυχαίο δείγμα μεγέθους N κάνοντας κάθε φορά αντικατάσταση. Για παράδειγμα, εάν υπήρχε μια λίστα, όπου τα δεδομένα εκπαίδευσης ήταν τα ακόλουθα [1, 2, 3, 4, 5, 6], τότε θα μπορούσαμε να δώσουμε σ' ένα από τα δέντρα του δάσους την ακόλουθη λίστα [1, 2, 2, 3, 6, 6]. Παρατηρήστε ότι και στις δύο περιπτώσεις, πριν και μετά της αντικατάστασης, οι δύο λίστες εξακολουθούν να έχουν μήκος 6. Όπως αναφέρθηκε παραπάνω, ένα από τα κύρια μειονεκτήματα των δέντρων αποφάσεων είναι ότι υποφέρουν από εκτιμήσεις μεγάλης διακύμανσης. Αυτό σημαίνει, ότι η προσθήκη ενός μικρού αριθμού πρόσθετων παρατηρήσεων εκπαίδευσης μπορεί να μεταβάλει δραματικά την απόδοση πρόβλεψης ενός ήδη εκπαιδευμένου δέντρου, μολονότι τα δεδομένα εκπαίδευσης δεν θα αλλάξουν σε μεγάλο βαθμό.

Αυτό έρχεται σε αντίθεση με έναν εκτιμητή χαμηλής διακύμανσης, όπως είναι η γραμμική παλινδρόμηση, η οποία δεν επηρεάζεται από την προσθήκη επιπλέον δεδομένων, τα οποία είναι σχετικά κοντά με τα υπόλοιπα δεδομένα. Ένας τρόπος για να μετριαστεί αυτό το πρόβλημα είναι μέσω του bagging. Η λογική της μεθόδου αυτής είναι να συνδυαστούν πολλαπλά δέντρα, τα οποία είναι όλα προσαρμοσμένα σε ξεχωριστά δείγματα (bootstrapped samples) και υπολογίζουν τον μέσο όρο των προβλέψεων τους για να μειώσουν τη συνολική διακύμανση αυτών των προβλέψεων. Δηλαδή, εάν ληφθεί ο μέσος όρος των παρατηρήσεων αυτών, η διακύμανση μειώνεται κατά παράγοντα ίσο με τον αριθμό των παρατηρήσεων. Ωστόσο, όταν υπάρχει μόνο ένα σύνολο εκπαίδευσης

σημαίνει, ότι είναι δύσκολο να δημιουργηθούν πάρα πολλά ανεξάρτητα σύνολα εκπαίδευσης. Εδώ απαιτείται η τεχνική του bagging.

Αν B είναι τα ξεχωριστά δείγματα εκκίνησης (bootstrapped samples), που έχουν δημιουργηθεί με ξεχωριστούς εκτιμητές μοντέλων $f^b(x)$, παίρνοντας τον μέσο όρο αυτών έχουμε έναν νέο εκτιμητή χαμηλής πλέον διακύμανσης $f_{avg}(x)$, όπου

$$\widehat{F}_{avg}(x) = \frac{1}{B} * \sum_{b=1}^B \widehat{f}^b(x)$$

Ένα από τα βασικά πλεονεκτήματα του bagging είναι ότι δεν κάνει υπερπροσαρμογή το μοντέλο όταν αυξάνεται ο αριθμός των δειγμάτων εκκίνησης B . [28]

Τα τυχαία δάση ακολουθούν μια παρόμοια τεχνική με τη μέθοδο bagging, την τυχειότητα των χαρακτηριστικών (feature Randomness), στην οποία τυχαία επιλέγεται ένα υποσύνολο από τα p -διάστατα χαρακτηριστικά σε κάθε ξεχωριστό δέντρο που αναπτύσσεται. Αυτό είναι επιθυμητό, ώστε να συμπεριληφθούν όσο το δυνατόν περισσότερα χαρακτηριστικά (features), προκειμένου να συγκεντρωθούν περισσότερες πληροφορίες για το μοντέλο. Αυτό έχει ως στόχο τη σκόπιμη αποφυγή πολύ ισχυρών προγνωστικών χαρακτηριστικών που οδηγούν σε παρόμοιους διαχωρισμούς στα δέντρα. Σε ένα κανονικό δέντρο αποφάσεων, κατά τη διάρκεια του διαχωρισμού ενός κόμβου, κάθε πιθανό χαρακτηριστικό εξετάζεται, ώστε τελικά να επιλεγεί αυτό που παράγει τον μεγαλύτερο διαχωρισμό μεταξύ των παρατηρήσεων στον αριστερό κόμβο έναντι εκείνων που βρίσκονται στον δεξιό κόμβο. Ο αλγόριθμος των τυχαίων δέντρων ακολουθεί μια εντελώς αντίθετη προσέγγιση. Συγκεκριμένα, κάθε δέντρο επιλέγει μόνο από ένα τυχαίο υποσύνολο χαρακτηριστικών. Αυτό οδηγεί σε ακόμη περισσότερες παραλλαγές μεταξύ των δέντρων του μοντέλου και τελικά προκύπτει χαμηλότερη συσχέτιση μεταξύ των δέντρων και μεγαλύτερη διαφοροποίηση. Καταλήγοντας έτσι, σε δέντρα που όχι μόνο εκπαιδεύονται σε διαφορετικά σύνολα δεδομένων, αλλά και χρησιμοποιούν διαφορετικά χαρακτηριστικά κατά τη λήψη αποφάσεων.

1.3.4 Εξαιρετικά τυχαία δέντρα - Extra Trees - Extremely Randomized Trees

Ένα μοναδικό δέντρο αποφάσεων, όπως αναφέρθηκε παραπάνω, παρουσιάζει συνήθως υψηλή απόκλιση, καθώς υπερπροσαρμόζει (overfit) τα δεδομένα, από τα οποία εκπαιδεύεται. Αυτό πρακτικά συμβαίνει, γιατί μαθαίνει από ένα μόνο μονοπάτι των αποφάσεων. Συμπεραίνοντας έτσι ότι οι προβλέψεις από ένα δέντρο απόφασης συνήθως δεν είναι ακριβείς για την κατηγοριοποίηση των νέων δεδομένων. Μια εξέλιξη της παραπάνω μεθόδου είναι το τυχαίο δάσος, όπως αναλυτικά έχει περιγραφεί, το οποίο παρουσιάζει μεσαία απόκλιση και εισάγει την έννοια της τυχαιότητας. [29]

Τα Εξαιρετικά Τυχαία Δέντρα (Extremely Randomized Trees) έχουν γενικά χαμηλή απόκλιση. Ο κύριος στόχος τους είναι η περαιτέρω τυχαιοποίηση της δομής των δέντρων, όπου η επιλογή του βέλτιστου σημείου κοπής είναι υπεύθυνη για την διακύμανση του επαγόμενου δέντρου. Μοιάζουν πολύ με τα τυχαία δάση, καθώς οικοδομούν πολλαπλά δέντρα (n estimators) και κόβουν κόμβους χρησιμοποιώντας τυχαία υποσύνολα χαρακτηριστικών, αλλά έχουν δύο βασικές διαφορές. Πρώτον τα εξαιρετικά τυχαία δάση δεν ακολουθούν την τεχνική bootstrapping για τις παρατηρήσεις, που σημαίνει ότι ο αλγόριθμος χρησιμοποιεί δείγματα χωρίς αντικατάσταση. Δεύτερον οι κόμβοι χωρίζονται βασιζόμενοι σε τυχαίους διαχωρισμούς μεταξύ ενός τυχαίου υποσυνόλου των χαρακτηριστικών που επιλέγονται στον κάθε κόμβο. Στα Εξαιρετικά Τυχαία Δέντρα, η τυχαιότητα δεν προέρχεται από το bootstrapping των δεδομένων, αλλά πηγάζει από τυχαίους διαχωρισμούς όλων των παρατηρήσεων και ο αλγόριθμος προσπαθεί αντί να βρει ένα βέλτιστο σημείο κοπής για καθένα από τα τυχαία επιλεγμένα χαρακτηριστικά K κάθε κόμβου, να επιλέξει ένα σημείο κοπής τυχαία. [30]

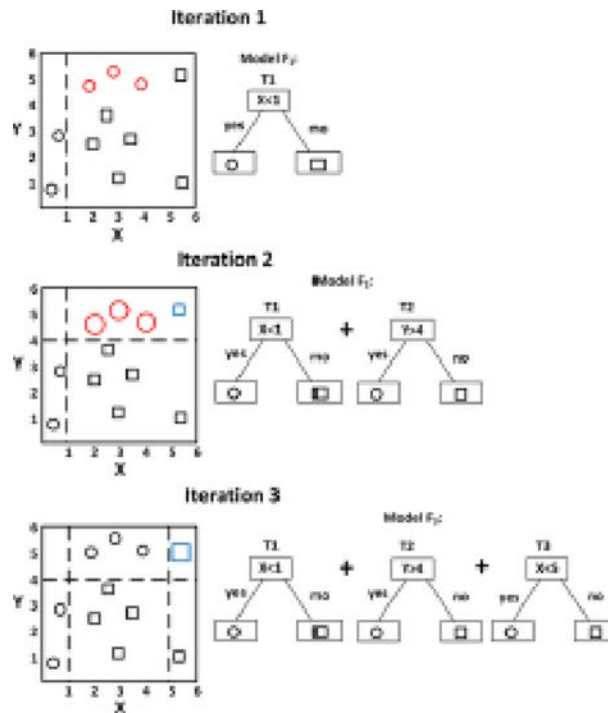
Αυτή η μέθοδος οδηγεί συχνά σε αυξημένη ακρίβεια χάρη στην εξομάλυνση και ταυτοχρόνως μειώνει σημαντικά την υπολογιστική επιβάρυνση που συνδέεται με τον προσδιορισμό του βέλτιστου σημείου κοπής των δέντρων στα τυχαία δάση.

Από στατιστικής απόψεως, η απόρριψη της ιδέας του bootstrapping οδηγεί σε ένα πλεονέκτημα όσον αφορά το συστηματικό σφάλμα (bias), ενώ η τυχαιοποίηση των σημείων τομής έχει συχνά εξαιρετικό αποτέλεσμα στην μείωση της διακύμανσης. Αυτή η μέθοδος έχει αποφέρει προηγμένα αποτελέσματα σε αρκετά μεγάλα πολύπλοκα προβλήματα. [29]

1.3.5 Βαθμωτή ενίσχυση - Gradient boosting

Η τεχνική της ενίσχυσης (Boosting) προέκυψε από την ιδέα, πως ένας αδύναμος ταξινομητής μπορεί να τροποποιηθεί, για να γίνει καλύτερος. Αποτελεί μια μέθοδο μετατροπής των αδύναμων ταξινομητών σε ισχυρούς. Μια αδύναμη υπόθεση ή ένας ανίσχυρος ταξινομητής ορίζεται ως εκείνος του οποίου η απόδοση είναι τουλάχιστον ελαφρώς καλύτερη από μια τυχαία επιλογή. Με την τεχνική της ενίσχυσης (boosting), κάθε νέο δέντρο είναι μια προσαρμογή σε μια τροποποιημένη έκδοση του αρχικού συνόλου δεδομένων. [31]

Ο αλγόριθμος gradient boosting μπορεί να αποσαφηνιστεί εύκολα, αφού πρώτα αναφερθούμε στον αλγόριθμο AdaBoost. Ο αλγόριθμος AdaBoost αρχίζει με την εκπαίδευση ενός δέντρου αποφάσεων, στο οποίο σε κάθε παρατήρηση αποδίδεται ίσο βάρος, δηλαδή όλες οι παρατηρήσεις είναι όμοια ζυγισμένες. Μετά την αξιολόγηση του πρώτου δέντρου, αυξάνουμε τα βάρη εκείνων των παρατηρήσεων που ήταν δύσκολο να ταξινομηθούν σωστά και μειώνουμε τα βάρη για τις παρατηρήσεις, οι οποίες ευχερώς ταξινομήθηκαν. Κατά συνέπεια, το δεύτερο δέντρο που δημιουργείται αναπτύσσεται σ' αυτά τα νέα σταθμισμένα δεδομένα. Στόχος μας είναι να βελτιωθούν οι προβλέψεις του πρώτου δέντρου. Το νέο μοντέλο, επομένως, είναι «Δέντρο 1 + Δέντρο 2». Κατόπιν, υπολογίζεται το σφάλμα ταξινόμησης απ' αυτό το νέο μοντέλο δύο δέντρων και δημιουργείται ένα τρίτο δέντρο για την πρόβλεψη των αναθεωρημένων υπολοίπων (residuals). Η εν λόγω διαδικασία επαναλαμβάνεται για ένα συγκεκριμένο αριθμό επαναλήψεων. Τα μεταγενέστερα δέντρα βοηθούν στην ταξινόμηση των παρατηρήσεων, που δεν είναι σωστά ταξινομημένες από τα προηγούμενα δέντρα. Οι προβλέψεις του τελικού μοντέλου είναι, συνεπώς, το σταθμισμένο άθροισμα των προβλέψεων, που προέκυψαν από τα προηγούμενα μοντέλα (δέντρα).



Εικόνα 8: Απεικόνιση του ταξινομητή Gradient Boosting[52].

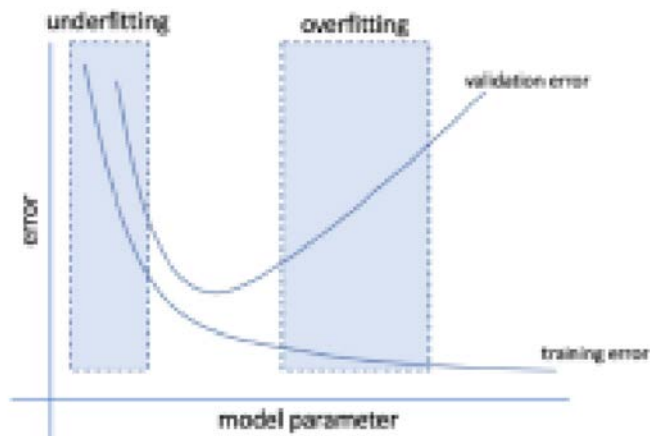
Σχετικά με τον αλγόριθμο Gradient Boosting, είναι σημαντικό να αναφέρουμε ότι λειτουργεί εκπαιδεύοντας πολλά μοντέλα με βαθμιαίο, προσθετικό και διαδοχικό τρόπο. Αρχικά, όπως στην περίπτωση του AdaBoost, χρησιμοποιούνται πολύ σύντομα δέντρα απόφασης, που έχουν μόνο ένα μοναδικό διαχωρισμό, το οποίο ονομάζεται decision stump. Η μεγάλη διαφορά μεταξύ του AdaBoost και του Gradient Boosting είναι ο τρόπος με τον οποίο οι δύο αλγόριθμοι εντοπίζουν τις ατέλειες των αδύναμων ταξινομητών (στη συγκεκριμένη περίπτωση των δέντρων αποφάσεων). Ενώ το μοντέλο AdaBoost αναγνωρίζει τις αδυναμίες χρησιμοποιώντας υψηλού βάρους δεδομένα, το μοντέλο gradient boosting κάνει το ίδιο τοποθετώντας βαθμίδες (gradients) στη συνάρτηση απώλειας (loss function). Έτσι, πίσω από τον αλγόριθμο GB βρίσκεται η επαναλαμβανόμενη μόχλευση των μοτίβων στα υπόλοιπα (residuals), η ενίσχυση ενός μοντέλου με αδύναμες προβλέψεις και τελικά η βελτίωση του. Μόλις φτάσουμε σ' ένα στάδιο, όπου τα υπολοιπα (residuals) δεν έχουν κανένα πρότυπο, που θα μπορούσε να διαμορφωθεί, η μοντελοποίηση σταματάει, διαφορετικά ίσως προκύψει υπερπροσαρμογή (overfitting) των δεδομένων. Αλγοριθμικά, ελαχιστοποιούμε τη συνάρτηση απώλειας, ώστε να φτάσει στις ελάχιστες τιμές της. [32]

Η συνάρτηση απώλειας είναι ένα μέτρο που υποδεικνύει πόσο καλοί είναι οι συντελεστές (coefficients) του μοντέλου στην προσαρμογή των υποκείμενων δεδομένων [33]. Για να κατανοήσουμε λοιπόν τη συνάρτηση απώλειας, πρέπει να εξετάσουμε το τι προσπαθούμε να βελτιστοποιήσουμε. Για παράδειγμα, αν προσπαθούμε να προβλέψουμε τις τιμές πώλησης ακινήτων, τότε η συνάρτηση απώλειας θα βασίζεται στο σφάλμα μεταξύ των πραγματικών και των προβλεπόμενων τιμών των κατοικιών. Γενικά, η συνάρτηση απώλειας που χρησιμοποιείται εξαρτάται από τον τύπο του προβλήματος που επιλύεται. Ο αλγόριθμος υποστηρίζει πολλές τυπικές συναρτήσεις απώλειας, αλλά και ο χρήστης μπορεί να ορίσει τη δική του. Για παράδειγμα, στην παλινδρόμηση μπορεί να χρησιμοποιηθεί το τετραγωνικό σφάλμα σαν συνάρτηση απώλειας, ενώ για τα προβλήματα ταξινόμησης μπορεί να εξεταστεί η λογαριθμική απώλεια (logarithmic loss). Ένα πλεονέκτημα του gradient boosting είναι ότι δεν χρειάζεται να προκύψει ένας νέος αλγόριθμος ενίσχυσης για κάθε συνάρτηση απώλειας, που μπορεί να θέλει να χρησιμοποιήσει ο χρήστης, αν αυτός πρόκειται για ένα αρκετά γενικό πλαίσιο που είναι ικανό να χειριστεί οποιαδήποτε διαφοροποιημένη συνάρτηση απώλειας.[34] Ένα από τα μεγαλύτερα κίνητρα για τη χρήση της του αλγορίθμου gradient boosting είναι ότι επιτρέπει τη βελτιστοποίηση μιας συνάρτησης κόστους που καθορίζεται από το χρήστη, αντί για μια συγκεκριμένη συνάρτηση απώλειας που συνήθως προσφέρει λιγότερο έλεγχο και δεν ανταποκρίνεται ουσιαστικά στις εφαρμογές του πραγματικού κόσμου.

1.4 Διασταυρωμένη Επικύρωση – Cross Validation

Η τεχνική cross - validation είναι μια μέθοδος στατιστική, χρησιμοποιείται για την εκτίμηση της ικανότητας των μοντέλων μηχανικής μάθησης. Η επικύρωση είναι μία από τις σημαντικές τεχνικές, που χρησιμοποιεί ένας επιστήμονας δεδομένων, αφού υπάρχει πάντα η ανάγκη να κατακυρωθεί η σταθερότητα του μοντέλου μηχανικής μάθησης και να εξεταστεί το πόσο καλά θα γενικευόταν σε κάποια νέα δεδομένα. Συνήθως χρειαζόμαστε μια διαβεβαίωση, ότι το μοντέλο έχει καταλάβει τα περισσότερα από τα μοτίβα των δεδομένων σωστά και ότι δεν έχει συμπεριλάβει πολύ θόρυβο. Με απλά λόγια, πρέπει να σιγουρευτούμε ότι το μοντέλο έχει το χαμηλότερο συστηματικό σφάλμα (bias) και διακύμανση (variance). [35]

Πρόκειται για μια τεχνική επικύρωσης για την αξιολόγηση του τρόπου γενίκευσης των αποτελεσμάτων μιας στατιστικής ανάλυσης σ' ένα ανεξάρτητο σύνολο δεδομένων. Χρησιμοποιείται σε περιπτώσεις, στις οποίες ο στόχος είναι η πρόβλεψη και κάποιος θέλει να υπολογίσει με ακρίβεια πόσο εύστοχο είναι το μοντέλο. Ο στόχος της διασταυρούμενης επικύρωσης είναι να καθοριστεί το σύνολο δεδομένων για τη δοκιμή του μοντέλου στη φάση της εκπαίδευσης, το οποίο ονομάζεται σύνολο επικύρωσης, έτσι ώστε να περιοριστούν προβλήματα, όπως η υπερπροσαρμογή (overfitting), η ελλιπής προσαρμογή (underfitting) και να δημιουργηθεί μια εικόνα για το πώς το μοντέλο θα γενικευθεί σ' ένα ανεξάρτητο σύνολο δεδομένων. Είναι σημαντικό τα σύνολα επικύρωσης και εκπαίδευσης να είναι από την ίδια κατανομή. Με τον όρο της ελλιπούς προσαρμογής (underfitting) γίνεται λόγος για τη μη λήψη επαρκών μοτίβων από τα δεδομένα. Το μοντέλο είναι ελλιπείς απέναντι στο σύνολο εκπαίδευσης και το σύνολο δοκιμών. Αντίθετα, στην περίπτωση της υπερπροσαρμογής (overfitting) εννοούμε τόσο τη λήψη περιττού θορύβου, αλλά και τη καταγραφή μοτίβων, τα οποία δεν μπορούν να γενικευτούν «καλά» σε άγνωστα δεδομένα. Το μοντέλο ανταποκρίνεται εξαιρετικά «καλά» στο σύνολο εκπαίδευσης, αλλά «άσχημα» στο σύνολο δοκιμών.



Εικόνα 9: Λάθος εκπαίδευσης και επικύρωσης σε σχέση με τις παραμέτρους[53].

Υπάρχουν πολλές τεχνικές, για να εφαρμοστεί η μέθοδος cross validation. Στην πρώτη στρατηγική, τα δεδομένα χωρίζονται σε δύο ομάδες, στο σύνολο εκπαίδευσης και δοκιμής, ώστε το ένα σύνολο να μην αλληλεπικαλύπτει το άλλο. Αυτός είναι ο λόγος, για τον οποίο είναι σημαντικό, να μην υπάρχουν διπλά δείγματα στο σύνολο δεδομένων. Στη συνέχεια, η εκπαίδευση του μοντέλου γίνεται στο σύνολο εκπαίδευσης, ενώ το σύνολο δοκιμής χρησιμοποιείται για σκοπό επικύρωσης. Αν δεν υπάρχουν αρκετά δεδομένα, για να εκπαιδευτεί το μοντέλο, η κατάργηση ενός μέρους του συνόλου εκπαίδευσης (training set), για να χρησιμοποιηθεί σαν σύνολο επικύρωσης (validation set), θέτει πρόβλημα ελλιπούς προσαρμογής (underfitting). Ακολουθώντας, εάν διαθέτουμε περιορισμένα δεδομένα, κινδυνεύουμε να χάσουμε σημαντικά μοτίβα και τάσεις στο σύνολο δεδομένων εκπαίδευσης, γεγονός που αυξάνει την πιθανότητα μεγάλου συστηματικού σφάλματος (bias). Επίσης, τι συμβαίνει στην περίπτωση που ο διαχωρισμός που γίνεται δεν είναι τυχαίος; Για παράδειγμα, αν το σύνολο των δεδομένων έχει μόνο άτομα από ένα συγκεκριμένο κράτος, αμειβόμενους μ' ένα συγκεκριμένο μισθό, μόνο γυναίκες ή μόνο άτομα ορισμένης ηλικίας. Αυτό θα οδηγήσει σε υπερπροσαρμογή (overfitting), καθώς δεν είναι βέβαιο ποια δεδομένα θα καταλήξουν στο σύνολο επικύρωσης. Έτσι το αποτέλεσμα μπορεί να είναι εντελώς διαφορετικό για διαφορετικά σύνολα. Εάν το πλήθος των δεδομένων είναι μεγάλο και τα δείγματα εκπαίδευσης κι ελέγχου ακολουθούν την ίδια κατανομή, τότε αυτό αποφεύγεται και η εν λόγω προσέγγιση είναι αποδεκτή.

Μια ακόμα τεχνική είναι το k cross validation. Κατά τη διάρκεια αυτής της διαδικασίας, λαμβάνεται μια μοναδική παράμετρος που ονομάζεται k, η οποία αναφέρεται στον αριθμό

των ομάδων, στις οποίες πρόκειται να χωριστεί ένα δοσμένο σύνολο δεδομένων. Ως εκ τούτου, η διαδικασία συχνά ονομάζεται *k-fold cross-validation*. Είναι μια δημοφιλής μέθοδος, επειδή είναι απλή στην κατανόησή της και γενικά έχει ως αποτέλεσμα μια λιγότερο μεροληπτική ή λιγότερο αισιόδοξη εκτίμηση της ικανότητας του μοντέλου από άλλες μεθόδους. Αυτή η προσέγγιση περιλαμβάνει την τυχαία διαίρεση του συνόλου των παρατηρήσεων σε ομάδες k , ή πτυχές, περίπου ίσου μεγέθους. Η πρώτη πτυχή αντιμετωπίζεται ως σύνολο επικύρωσης και η μέθοδος εφαρμόζεται στις υπόλοιπες $k - 1$ πτυχές. Στην επικύρωση *k Fold*, τα δεδομένα χωρίζονται σε υποσύνολα k . Η διαδικασία επικύρωσης επαναλαμβάνεται k φορές, έτσι ώστε κάθε φορά, ένα από τα υποσύνολα k χρησιμοποιείται σαν σύνολο επικύρωσης, ενώ τα υπόλοιπα υποσύνολα $k-1$ σχηματίζουν ένα σύνολο εκπαίδευσης. Το σφάλμα υπολογίζεται κατά μέσον όρο σε όλες τις δοκιμές k , για να υπολογιστεί η συνολική αποτελεσματικότητα (*effectiveness*) του μοντέλου. Όπως είναι φανερό, κάθε σημείο, δηλαδή κάθε δεδομένο, βρίσκεται σ' ένα σύνολο επικύρωσης ακριβώς μία φορά, ενώ σε ένα σύνολο εκπαίδευσης $k-1$ φορές. Αυτό μειώνει σημαντικά το συστηματικό σφάλμα (*bias*), καθώς χρησιμοποιούμε τα περισσότερα δεδομένα για προσαρμογή (*fitting*), αλλά μειώνει σημαντικά και τη διακύμανση (*variance*), καθώς τα περισσότερα από τα δεδομένα χρησιμοποιούνται και στο σύνολο επικύρωσης. Επίσης, πρόκειται για μια αρκετά χρονοβόρα διαδικασία. [35] [36]

1.5 Σχετικές εργασίες

Πολλές εργασίες έχουν γίνει, όσον αφορά την μελέτη και την πρόβλεψη της κίνησης του ανθρώπου. Επίσης έχουν χρησιμοποιηθεί διαφορετικές τεχνικές, για να προσεγγιστούν τα προβλήματα ταξινόμησης και εξαγωγής των χαρακτηριστικών που είναι απαραίτητα σε μια τέτοια εργασία. Στην υποενότητα αυτή θα παρουσιαστούν διάφορες έρευνες οι οποίες έχουν εκπονηθεί στον τομέα αυτό.

Αρχικά η ερευνά των Ansuini et al[37] με τίτλο *Predicting Object Size from Hand Kinematics: A Temporal Perspective* είχε σκοπό την κατανόηση του πως επηρεάζεται η κίνηση του χεριού ενός ανθρώπου ανάλογα με το αντικείμενο το οποίο καλείται να πιάσει. Αυτή η έρευνα αποτέλεσε βάση για την παρούσα διπλωματική. Στα πλαίσια αυτής της έρευνας μετρήθηκαν η ταχύτητα του κάπρου και το άνοιγμα της λαβής των ανθρώπων στην διάρκεια του χρόνου της κίνησης προς διάφορα μεγέθη αντικειμένων. Τα μεγέθη των αντικειμένων ήταν δυο ένα μεγάλο (γκρέιπφρουτ) και ένα μικρό (φουντούκι). Τα δεδομένα της έρευνας στηρίχθηκαν στις προσπάθειες δεκαπέντε ανθρώπων να πιάσουν τα δυο αυτά αντικείμενα. Οι κινήσεις τους καταγράφονταν από οχτώ motion capture κάμερες ρυθμισμένες στα 100 Hz. Τα αποτελέσματα της έρευνάς τους έδειξαν ότι μπορούμε με πολύ μεγάλη ακρίβεια 90% να προβλέψουμε την κίνηση ενός ανθρώπου σε πολύ αρχικό στάδιο περίπου στο 40% αυτής.

Άλλες έρευνες για την πρόβλεψη ανθρώπινης κίνησης στην ρομποτική.

Πολλές εργασίες σχετικά με την πρόβλεψη κίνησης στη ρομποτική έχουν χρησιμοποιήσει γραφικά μοντέλα (graphical models). Για παράδειγμα, οι Kulic et al. περιγράφουν μια προσέγγιση για on-line, σταδιακή μάθηση της πλήρους κίνησης του σώματος ενός ανθρώπου, από την παρατήρηση της μέσα από τα αρχικά στάδια της χρησιμοποιώντας Hidden Markov Models. Η εφαρμογή του μοντέλου το οποίο κατασκεύασαν μπορεί να εφαρμοστεί τόσο για την αναγνώριση όσο και για την πρόβλεψη μιας κίνησης [38]. Οι Koppula και Saxena επικεντρώθηκαν στην πρόβλεψη της κίνησης χρησιμοποιώντας conditional random fields [39]. Μια άλλη προσέγγιση που χρησιμοποιείται συνήθως για την πρόβλεψη της ανθρώπινης κίνησης είναι το Inverse Optimal Control (IOC), το οποίο στοχεύει να βρει μια συνάρτηση κόστους που βασίζεται στην παρατηρούμενη συμπεριφορά. Για παράδειγμα, οι Berret et al. διερεύνησαν τις λειτουργίες κόστους για τον σχεδιασμό κίνησης του βραχίονα[40]. Οι συγγραφείς αναφέρουν ότι στις κινήσεις

αυτές σχετίζονται στενά οι δαπάνες της μηχανικής ενέργειας με την ομαλότητα της κίνησης σε επίπεδο αρθρώσεων. Οι Mainprice et al. διερεύνησαν την πρόβλεψη των ανθρώπινων κινήσεων σε κοινόχρηστους χώρους εργασίας [41]. Η μέθοδος που χρησιμοποίησαν ήταν η προσπάθεια αποφυγής εμποδίων μέσω συνεργασίας χρησιμοποιώντας επαναληπτικό επανασχεδιασμό (iterative replanning). Στον IOC, οι βιοκινηματικές διεργασίες αντιπροσωπεύονται συνήθως από απλοποιημένα μοντέλα, τα οποία δεν είναι σε θέση να συλλάβουν πλήρως τις σύνθετες βιομηχανικές συμπεριφορές της ανθρώπινης κίνησης σε όλο το σώμα και να προβλέψουν με ακρίβεια μια παρατηρούμενη κίνηση.

Πρόβλεψη ανθρώπινης κίνησης μέσο νευρικών δικτύων.

Οι πρόσφατες εργασίες για την πρόβλεψη της ανθρώπινης κίνησης επικεντρώνονται σε επαναλαμβανόμενες αρχιτεκτονικές νευρωνικών δικτύων (RNN). Οι Fragkiadaki et al. πρότειναν ένα μοντέλο βασισμένο σε RNN που ενσωματώνει μη γραμμικά δίκτυα κωδικοποιητή και αποκωδικοποιητή (encoder and decoder) πριν και μετά από επαναλαμβανόμενα επίπεδα [42]. Το μοντέλο τους είναι σε θέση να χειριστεί την εκπαίδευση σε πολλά θέματα και τομείς δραστηριοτήτων. Οι Jain et al. εισήγαγαν μια μέθοδο ενσωμάτωσης δομικών στοιχείων σε μια αρχιτεκτονική RNN [43]. Οι αυτόματοι κωδικοποιητές μπορούν επίσης να χρησιμοποιηθούν για τον καθαρισμό της πρόβλεψης [44]. Οι Martinez et al. εισήγαγαν μια προσέγγιση με βάση gated recurrent unit (GRU) με εναπομένουσα σύνδεση στη λειτουργία βρόχου και έδειξε ότι αυτό ξεπερνά τις προηγούμενες μεθόδους που βασίζονται σε RNN [45]. Οι Pavllo et al. βελτίωσαν περαιτέρω την πρόβλεψη με βάση το RNN αλλάζοντας την αναπαράσταση της γωνίας άρθρωσης σε τεταρτημόρια [45], [46]. Ωστόσο, αυτό επιβαρύνεται με επιπλέον επίπεδα κανονικοποίησης και ποινές κανονικοποίησης. Πρόσφατα, οι Wang και Feng εισήγαγαν ένα επαναλαμβανόμενο μοντέλο κωδικοποιητή-αποκωδικοποιητή ταχύτητας θέσης (VRED) [47]. Το μοντέλο τους προσθέτει μια επιπλέον σύνδεση ταχύτητας ως είσοδο στο κελί GRU στην επαναλαμβανόμενη δομή. Η πρόβλεψη κίνησης που βασίζεται σε επαναλαμβανόμενα νευρωνικά δίκτυα υπόσχεται καλά αποτελέσματα για την πρόβλεψη βραχυπρόθεσμης κίνησης. Ωστόσο, τα μοντέλα εκπαιδεύονται μόνο σε ανθρώπινα δεδομένα. Ο χειρισμός περιβαλλοντικών περιορισμών δεν είναι ακόμη δυνατός και θα απαιτούσε μεγάλες ποσότητες δεδομένων εκπαίδευσης.

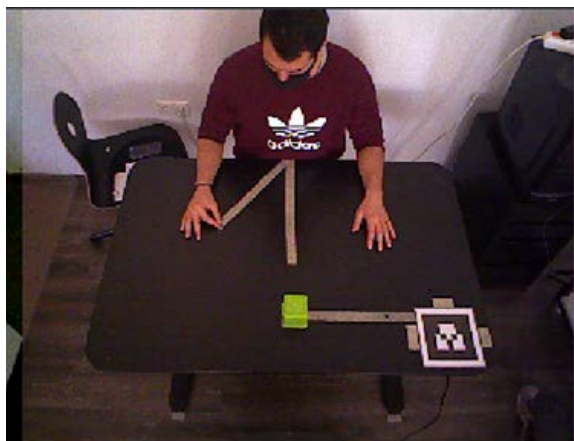
2 Μεθοδολογία

Στο κεφάλαιο αυτό περιγράφεται αρχικά η διαδικασία συλλογής δεδομένων και η προεπεξεργασία τους ώστε να είναι κατάλληλα για χρήση από τους αλγόριθμους μηχανικής μάθησης. Επίσης αναφέρονται οι αλγόριθμοι μηχανικής μάθησης που χρησιμοποιήθηκαν.

2.1 Συλλογή δεδομένων

Για την κατασκευή ενός αλγορίθμου πρόβλεψης καταγράψαμε δεδομένα από διάφορες κινήσεις ανθρώπων χρησιμοποιώντας έναν RGB-D οπτικό αισθητήρα. Από τους συμμετέχοντες ζητήθηκε να φθάσουν, να πιάσουν και να μετακινήσουν τρία διαφορετικά σε μέγεθος αντικείμενα αριστερά από την αρχική τους θέση. Στο πείραμα αυτό πήραν μέρος οκτώ άνθρωποι (5 Γυναίκες/ 3 Αντρες). Οι συμμετέχοντες ήταν ηλικίας από 20-40 ετών. Όλοι ήταν δεξιόχειρες με κανονική όραση ή διορθωμένη ώστε να είναι κανονική. Το πείραμα έγινε σύμφωνα με όλες τις προβλεπόμενες διαδικασίες για την συλλογή δεδομένων. Τέλος όλοι οι συμμετέχοντες έδωσαν συγκατάθεση για τη συμμετοχή τους στο πείραμα.

Ο χώρος του πειράματος αποτελούνταν από ένα τραπέζι μαύρου χρώματος, ένα σκαμπό ρυθμιζόμενου ύψους καθώς και το αντικείμενο που καλούνταν να μετακινήσει κάθε φορά ο συμμετέχων (εικ.10).



Εικόνα 10: Ο χώρος του πειράματος

Όλοι οι συμμετέχοντες καθόντουσαν σε ένα ρυθμιζόμενο ύψους σκαμπό ώστε να μπορούν να έχουν τοποθετημένο το δεξί τους χέρι στο τραπέζι με το καρπό και τον αγκώνα να σχηματίζουν σχεδόν ορθή γωνία με τον ώμο. Επίσης έπρεπε να έχουν τον αντίχειρά τους και τον δείκτη τους ενωμένους στο προκαθορισμένο σημείο έναρξης.

Τα αντικείμενα που χρησιμοποιήθηκαν στο πείραμα ήταν τρεις κύβοι. Ένας μπλε κύβος με πλευρά μήκους 2.5 cm, ένας κόκκινος κύβος με πλευρά μήκους 5.5 cm και ένας πράσινος κύβος με πλευρά μήκους 7.5 cm.

Η αρχική θέση του αντικείμενου ήταν σε απόσταση 63 cm από το κέντρο του κορμού του συμμετέχοντος. Η απόσταση που έπρεπε να μετακινήσει το προαναφερθέν αντικείμενο ήταν 30 cm αριστερά από την αρχική του θέση. Από τους συμμετέχοντες ζητήθηκε να πραγματοποιήσουν τις προσπάθειές τους με φυσική ταχύτητα χρησιμοποιώντας το δεξί τους χέρι. Κάθε προσπάθεια τους βιντεοσκοπούνταν και ελεγχόταν ώστε να πληρούνται όλα τα προ-απαιτούμενα. Κάθε συμμετέχον έκανε συνολικά 90 προσπάθειες οι οποίες χωρίζονταν σε 3 σετ των 10 προσπαθειών με κάθε αντικείμενο.

Πριν την συλλογή δεδομένων κάθε συμμετέχον πραγματοποιούσε κάποιες δοκιμαστικές προσπάθειες ώστε να εξοικειωθεί με το πείραμα. Τα αντικείμενα εναλλάσσονταν τυχαία μετά από 10 προσπάθειες. Για να καταγραφεί μια προσπάθεια χρειάζονταν από 7 έως 15 sec (ανάλογα με την αρχική και τελική παύση). Ανάμεσα στις προσπάθειες υπήρχε μία παύση ~15 sec. Το διάλειμμα αυτό αφορούσε τον χρόνο προετοιμασίας για την επόμενη καταγραφή. Όλη η διαδικασία συνολικά για κάθε συμμετέχοντα κρατούσε περίπου 1 ώρα.

Για την καταγραφή της κίνησης χρησιμοποιήθηκε μια RGB-D κάμερα, τοποθετημένη έτσι ώστε να κοιτάει το τραπέζι του πειράματος σχεδόν κάθετα. Η συχνότητα δειγματοληψίας ήταν τα 60 HZ.

2.2 Ανάλυση δεδομένων

2.2.1 Ανάλυση οπτικής πληροφορίας

Το λειτουργικό σύστημα με το οποίο έγινε η επεξεργασία όλων των βίντεο με τις κινήσεις των συμμετεχόντων ήταν τα Linux (Ubuntu 18.04). Επίσης, για την ανάλυση χρησιμοποιήθηκε και το λειτουργικό Robot Operating System (ROS²). Η έκδοση του ROS ήταν η Melodic. Η καταγραφή της οπτικής πληροφορίας έγινε μέσω του ROS. Η RGB και point cloud πληροφορία αποθηκεύτηκαν σε μορφή rosbags.

Κάθε rosbag υπέστη κατάλληλη επεξεργασία για εξαγωγή των 2Δ ή 3Δ συντεταγμένων των αρθρώσεων του χεριού του ανθρώπου κατά τη διάρκεια της κίνησης μέσω της συλλογής ROS πακέτων `openpose_3D_localization`³. Συγκεκριμένα στο `openpose_3D_localization` χρησιμοποιείται το Openpose[48] το οποίο εντοπίζει διάφορες αρθρώσεις του ανθρώπινου σώματος σε μια εικόνα (εικ.11). Οι αρθρώσεις που επιλέχθηκαν για να καταγραφεί η κίνησή τους ήταν ο καρπός του δεξιού χεριού, ο αντίχειρας και ο δείκτης κάθε συμμετέχοντα.

Για να διασφαλιστεί πως δεν θα χαθούν δεδομένα στην παραπάνω διαδικασία λόγω της φύσεως της διαδικασίας χρησιμοποιήθηκε το σύνολο του ROS `bag_read_service`⁴. Το σύνολο αυτό διαβάζει από ένα rosbag ένα μήνυμα τη φορά από κάθε topic και κάνει publish κάθε φορά αυτό το μήνυμα. Όταν ολοκληρωθεί η επεξεργασία ενός μηνύματος το σύνολο του στέλνει το επόμενο. Η χρήση του συνόλου αυτού είναι αναγκαία καθώς ο χρόνος που απαιτείται για την επεξεργασία ενός frame από το OpenPose είναι αρκετά μεγάλος, με αποτέλεσμα να χάνονται δεδομένα γιατί δεν έχει τελειώσει η επεξεργασία των προηγούμενων frames. Η πληροφορία που παρέχει το OpenPose είναι 2Δ. Για την εξαγωγή των 3Δ συντεταγμένων το `openpose_3D_localization` αξιοποιεί την πληροφορία του point cloud.

Τα δεδομένα που προέρχονται από την προαναφερθείσα διαδικασία αποτελούνται από συντεταγμένες που έχουν ως σημείο αναφοράς το κέντρο της κάμερας γεγονός που τα καθιστά δυσνόητα καθώς και μη συγκρίσιμα μεταξύ των διαφόρων συμμετεχόντων. Μη

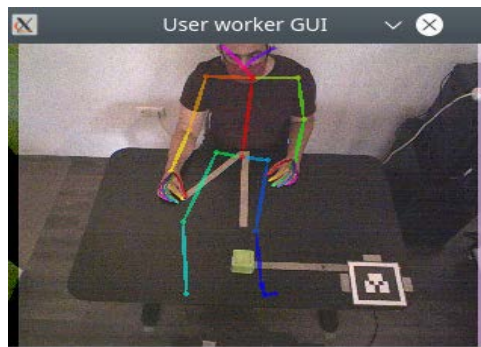
² <https://www.ros.org>

³ https://github.com/Roboskel-Manipulation/openpose_3D_localization

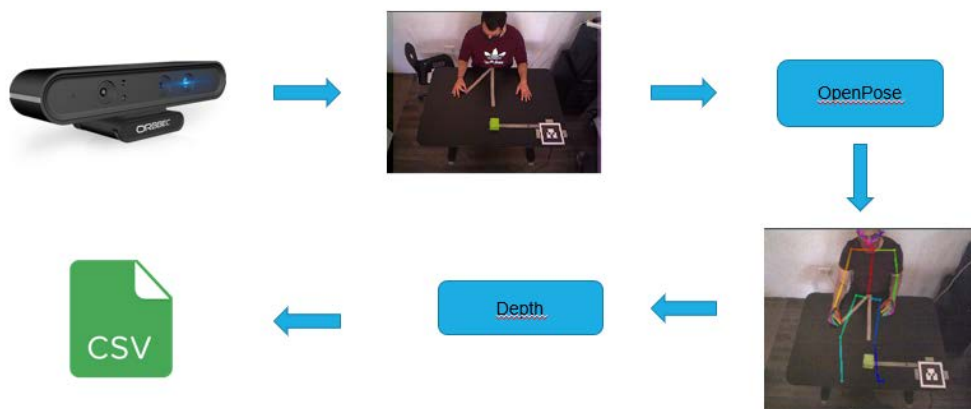
⁴ https://github.com/Roboskel-Manipulation/bag_read_service

συγκρίσιμα γίνονται διότι τα πειράματα δεν λαμβάνουν χώρα την ίδια μέρα με αποτέλεσμα ο χώρος του πειράματος να μεταβάλλεται. Για να είναι επομένως εφικτή η ύπαρξη ενός σταθερού σημείου αναφοράς ώστε να κατανοήσουμε καλύτερα την κίνηση και να εξάγουμε καλύτερά συμπεράσματα θέσαμε σαν σημείο αναφοράς το κέντρο ενός marker στην άκρη του τραπεζιού και μετατρέποντας όλα τα δεδομένα σε κοινό σημείο αναφοράς (και αυτή η υπηρεσία είναι διαθέσιμη στο `openpose_3D_localization`). Τα αποτελέσματα αυτής της μετατροπής καταγράφηκαν σε καινούργια `rosbag`

Τέλος, για να αξιοποιηθεί το σύνολο αυτής της πληροφορίας μετατρέψαμε τα τελικά `rosbag` σε `csv` αρχεία. Αυτό το κάνουμε μέσω της βιβλιοθήκης για του ROS `rosbag_pandas`⁵.



Εικόνα 11:Τυπική έξοδος του Openpose.

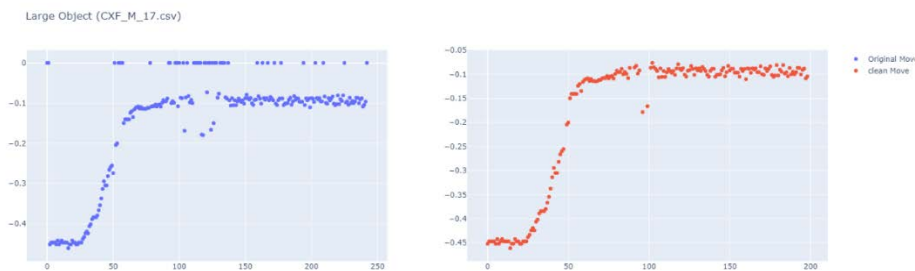


Εικόνα 12:Πορεία επεξεργασίας δεδομένων.

⁵ https://github.com/eurogroep/rosbag_pandas

2.2.2 Φιλτράρισμα θορύβου

Μέσω της απεικόνισης των δεδομένων διαπιστώθηκε ότι περιέχουν πολύ θόρυβο. Ο θόρυβος που εντοπίστηκε αρχικά στα δεδομένα τα οποία συλλέχθηκαν σχετίζεται με τις αποτυχίες στην πληροφορία του βάθους (είτε μη διαθέσιμη πληροφορία, είτε σφάλμα στην μέτρηση). Όταν αυτό συμβαίνει η έξοδος που προκύπτει είναι οι συντεταγμένες (0,0,0). Αυτό σε πολλές από τις κινήσεις έφτανε να αποτελεί έως και το 20% της κίνησης. Σαν πρώτο καθάρισμα των δεδομένων αφαιρέσαμε αυτά τα μηδενικά καθώς και τυχόν σημεία τα οποία εμφανιζόντουσαν εκτός της κίνησης (εικ.12). Σημεία εκτός της κίνησης θεωρήθηκαν σημεία τα οποία δεν ακολουθούν την καμπύλη της κίνηση και έχουν προκληθεί από λάθος της πληροφορίας βάθους. Για τον εντοπισμό αυτών των σημείων τέθηκε ως κριτήριο η απόσταση που έχει ένα σημείο από το προηγούμενο του. Πιο συγκεκριμένα εάν ένα σημείο απέχει από το προηγούμενο του κάτω από 0.055 cm θεωρείται εντός της κίνησης, διαφορετικά θεωρείται εκτός αυτής.

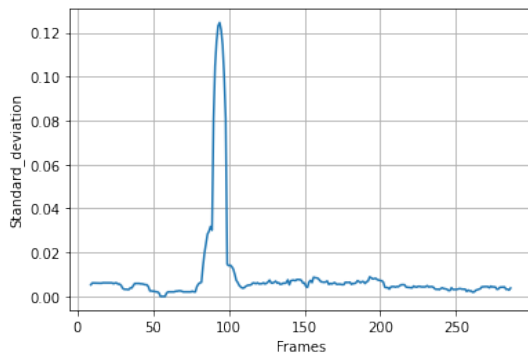


Εικόνα 13: Απεικόνιση κίνησης καρπού πριν και μετά από την επεξεργασία.

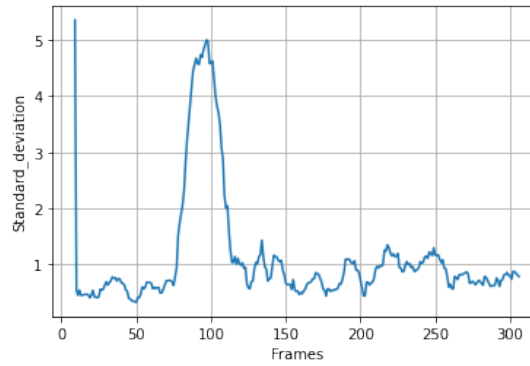
2.2.3 Κόψιμο της κίνησης

Έχοντας την τελική μορφή της κίνησης, το επόμενο βήμα επεξεργασίας ήταν η αφαίρεση των σημείων της κίνησης που δεν αφορούν το πρόβλημά μας. Μας ενδιαφέρει το κομμάτι της κίνησης από την έναρξή της (περίπου 2 με 3 δευτερόλεπτα από την έναρξη του βίντεο) μέχρι την πρώτη επαφή του συμμετέχοντος με το αντικείμενο. Το δεύτερο κομμάτι αφαιρέθηκε καθώς στα πλαίσια αυτής της μελέτης μας ενδιέφερε να κατανοήσουμε μόνο την κίνηση έως το πιάσιμο του αντικειμένου. Για να κοπεί η κίνηση χρειάστηκε να οριστούν οι δύο αυτές χρονικές στιγμές, η αρχή της κίνησης και το τέλος αυτής. Για να οριστούν αυτές οι χρονικές στιγμές αρχικά χρησιμοποιήθηκε η ταχύτητα του καρπού κάθε συμμετέχοντα. Όμως αυτό δεν ήταν δυνατόν καθώς στα δεδομένα της ταχύτητας υπήρχε

πολύ θόρυβο για να καταστεί εφικτή η εξαγωγή ορισμένων χρήσιμων συμπερασμάτων. Τελικά ορίστηκαν οι δυο αυτές χρονικές στιγμές από την τυπική απόκλιση της θέσης του καρπού στον χώρο με κυλιόμενο παράθυρο δέκα σημείων(εικ.14)(εικ.15).



Εικόνα 14:Standard deviation 3D δεδομένων.



Εικόνα 15:Standard deviation 2D δεδομένων.

Για να τεθούν τα όρια της τυπικής απόκλισης τα οποία θα σήμαιναν την αρχή και το τέλος της κίνησης, ώστε να είναι όσο το δυνατόν πιο ακριβής για όλες τις κινήσεις ανεξαρτήτως συμμετέχοντα και αντικειμένου, κατασκευάστηκε ένα dataset αποτελούμενο από πενήντα κινήσεις τυχαία επιλεγμένες από όλους τους συμμετέχοντες. Από τις κινήσεις αυτές μέσω της παρατήρησης των βίντεο εντοπίστηκε πότε ακριβώς αρχίζει και σταματά η κίνηση που μας ενδιαφέρει. Καταγράψαμε τον χρόνο καθώς και το frame στο οποίο έγιναν (groundtruth). Γνωρίζοντας αυτά, έγινε δυνατή η αξιολόγηση των κομμένων κινήσεων συγκρίνοντας το frame που προκύπτει από το αυτόματο κόψιμο μέσω διαφορετικών τιμών της τυπικής απόκλισης με αυτό που παρατηρούμε στο groundtruth.

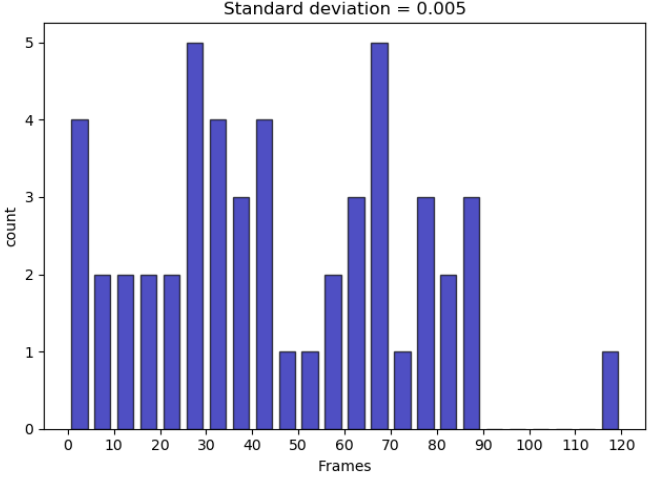
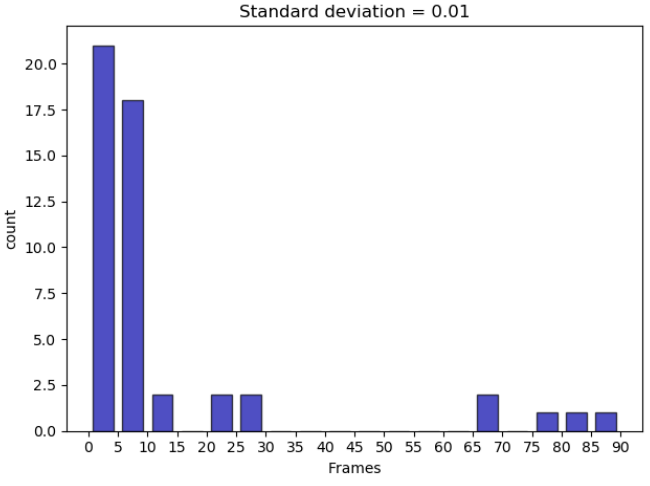
2.2.3.1 Κόψιμο κίνησης για 3D δεδομένα

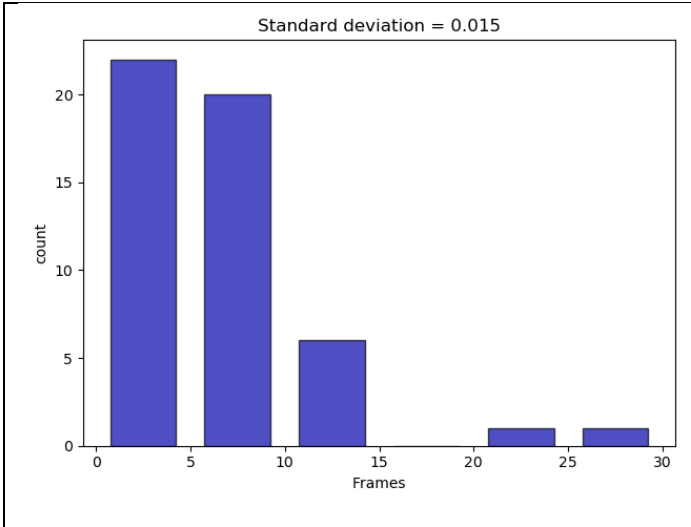
Για τον ορισμό του σημείου στο οποίο έγινε η αρχή της κίνησης δοκιμάστηκαν διάφορες τιμές τυπικής απόκλισης και μετρήθηκε το μέσο απόλυτο σφάλμα το οποίο προκύπτει από κάθε μια από αυτές ώστε να επιλεγθεί την καταλληλότερη για το πείραμα. Τα αποτελέσματα από τις δοκιμές φαίνονται στον παρακάτω πίνακα(πιν.1).

Για να οριστεί το τέλος της κίνησης εκτελέστηκε η αντίστοιχη διαδικασία για τις ίδιες τιμές τυπικής απόκλισης όπου κάτω από αυτές θα σήμαινε ότι η κίνηση έχει τελειώσει. Αυτή τη

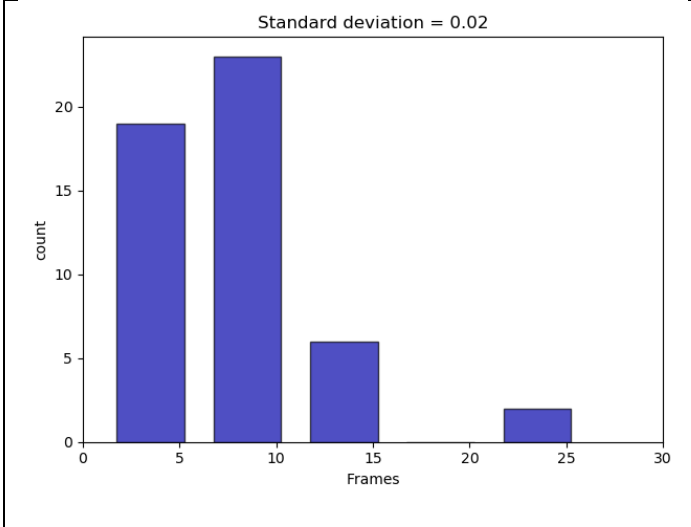
φορά η σύγκριση των τιμών της τυπικής απόκλιση δεν ξεκινά από την αρχή αλλά από το σημείο όπου αυτή είχε την μέγιστη τιμή της. Αυτό έγινε γιατί παρατηρήθηκε πως σε όλες τις κινήσεις το τέλος τους γινόταν μετά από αυτή την τιμή. Τα αποτελέσματα από τις δοκιμές φαίνονται στον παρακάτω πίνακα(πιν.2).

Πίνακας 1

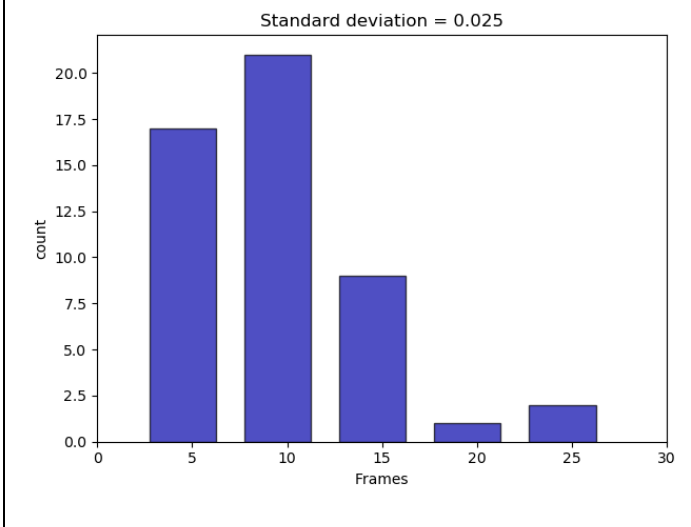
Ιστογράμμο τυπικής απόκλισης	Μέσο απόλυτο σφάλμα (Frames)																																										
 <p>Standard deviation = 0.005</p> <table border="1"> <caption>Data for Standard deviation = 0.005 histogram</caption> <thead> <tr> <th>Frames</th> <th>Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>4</td></tr> <tr><td>5</td><td>2</td></tr> <tr><td>10</td><td>2</td></tr> <tr><td>15</td><td>2</td></tr> <tr><td>20</td><td>2</td></tr> <tr><td>25</td><td>5</td></tr> <tr><td>30</td><td>4</td></tr> <tr><td>35</td><td>3</td></tr> <tr><td>40</td><td>4</td></tr> <tr><td>45</td><td>1</td></tr> <tr><td>50</td><td>1</td></tr> <tr><td>55</td><td>2</td></tr> <tr><td>60</td><td>3</td></tr> <tr><td>65</td><td>5</td></tr> <tr><td>70</td><td>1</td></tr> <tr><td>75</td><td>3</td></tr> <tr><td>80</td><td>2</td></tr> <tr><td>85</td><td>3</td></tr> <tr><td>90</td><td>0</td></tr> <tr><td>115</td><td>1</td></tr> </tbody> </table>	Frames	Count	0	4	5	2	10	2	15	2	20	2	25	5	30	4	35	3	40	4	45	1	50	1	55	2	60	3	65	5	70	1	75	3	80	2	85	3	90	0	115	1	<p>44.66</p>
Frames	Count																																										
0	4																																										
5	2																																										
10	2																																										
15	2																																										
20	2																																										
25	5																																										
30	4																																										
35	3																																										
40	4																																										
45	1																																										
50	1																																										
55	2																																										
60	3																																										
65	5																																										
70	1																																										
75	3																																										
80	2																																										
85	3																																										
90	0																																										
115	1																																										
 <p>Standard deviation = 0.01</p> <table border="1"> <caption>Data for Standard deviation = 0.01 histogram</caption> <thead> <tr> <th>Frames</th> <th>Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>21</td></tr> <tr><td>5</td><td>18</td></tr> <tr><td>10</td><td>2</td></tr> <tr><td>20</td><td>2</td></tr> <tr><td>25</td><td>2</td></tr> <tr><td>30</td><td>0</td></tr> <tr><td>35</td><td>0</td></tr> <tr><td>40</td><td>0</td></tr> <tr><td>45</td><td>0</td></tr> <tr><td>50</td><td>0</td></tr> <tr><td>55</td><td>0</td></tr> <tr><td>60</td><td>0</td></tr> <tr><td>65</td><td>2</td></tr> <tr><td>70</td><td>0</td></tr> <tr><td>75</td><td>0</td></tr> <tr><td>80</td><td>1</td></tr> <tr><td>85</td><td>1</td></tr> <tr><td>90</td><td>1</td></tr> </tbody> </table>	Frames	Count	0	21	5	18	10	2	20	2	25	2	30	0	35	0	40	0	45	0	50	0	55	0	60	0	65	2	70	0	75	0	80	1	85	1	90	1	<p>13.3</p>				
Frames	Count																																										
0	21																																										
5	18																																										
10	2																																										
20	2																																										
25	2																																										
30	0																																										
35	0																																										
40	0																																										
45	0																																										
50	0																																										
55	0																																										
60	0																																										
65	2																																										
70	0																																										
75	0																																										
80	1																																										
85	1																																										
90	1																																										



6.18

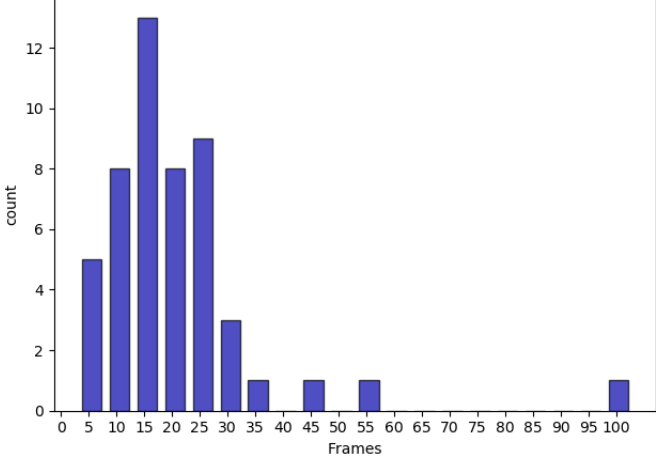
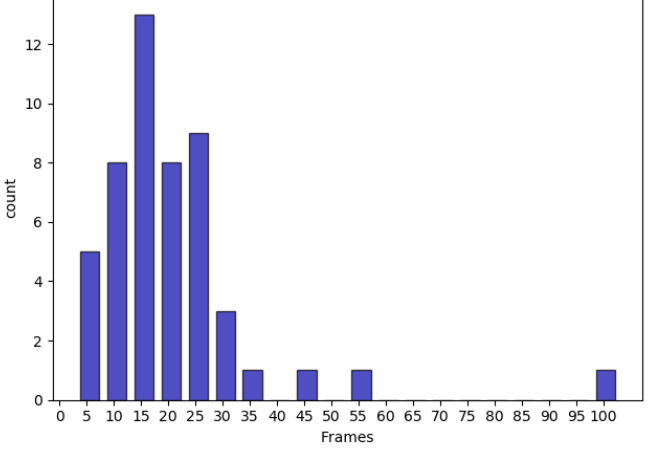


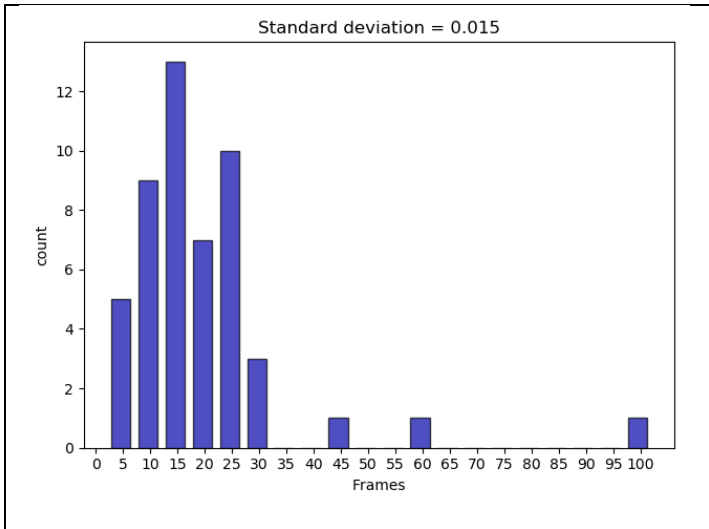
7.38



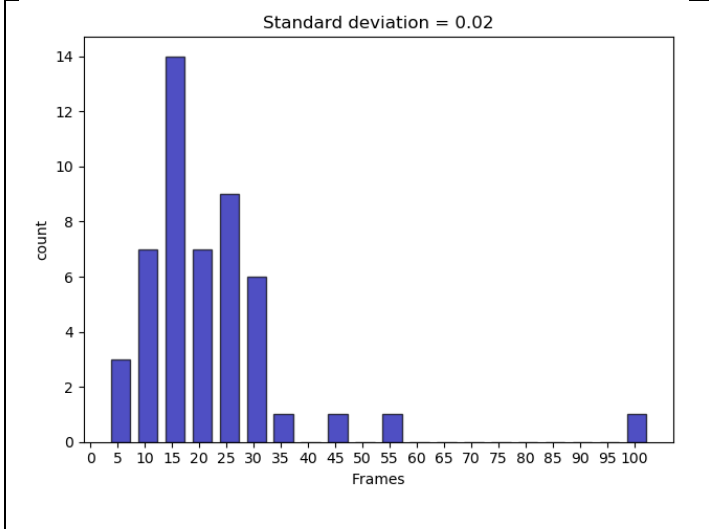
8.8

Πίνακας 2

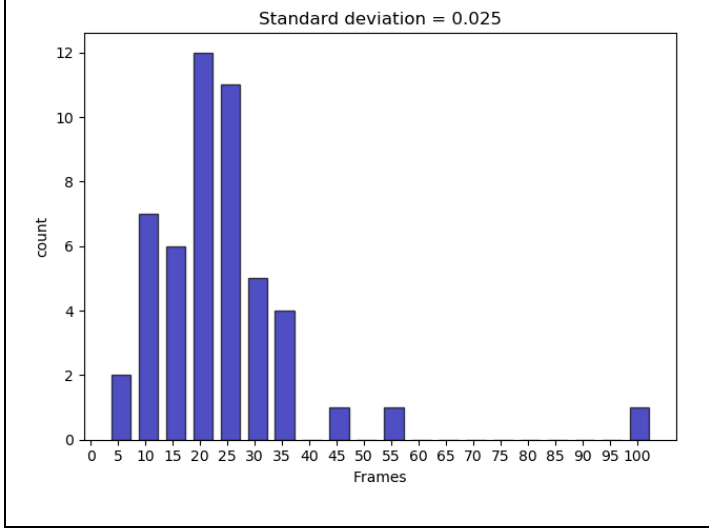
Ιστόγραμμα τυπικής απόκλισης	Μέσο απόλυτο σφάλμα (Frames)																						
<p>Standard deviation = 0.005</p>  <table border="1"> <caption>Data for Standard deviation = 0.005</caption> <thead> <tr> <th>Frames</th> <th>Count</th> </tr> </thead> <tbody> <tr><td>5</td><td>5</td></tr> <tr><td>10</td><td>8</td></tr> <tr><td>15</td><td>13</td></tr> <tr><td>20</td><td>8</td></tr> <tr><td>25</td><td>9</td></tr> <tr><td>30</td><td>3</td></tr> <tr><td>35</td><td>1</td></tr> <tr><td>45</td><td>1</td></tr> <tr><td>55</td><td>1</td></tr> <tr><td>100</td><td>1</td></tr> </tbody> </table>	Frames	Count	5	5	10	8	15	13	20	8	25	9	30	3	35	1	45	1	55	1	100	1	<p>19.06</p>
Frames	Count																						
5	5																						
10	8																						
15	13																						
20	8																						
25	9																						
30	3																						
35	1																						
45	1																						
55	1																						
100	1																						
<p>Standard deviation = 0.01</p>  <table border="1"> <caption>Data for Standard deviation = 0.01</caption> <thead> <tr> <th>Frames</th> <th>Count</th> </tr> </thead> <tbody> <tr><td>5</td><td>5</td></tr> <tr><td>10</td><td>8</td></tr> <tr><td>15</td><td>13</td></tr> <tr><td>20</td><td>8</td></tr> <tr><td>25</td><td>9</td></tr> <tr><td>30</td><td>3</td></tr> <tr><td>35</td><td>1</td></tr> <tr><td>45</td><td>1</td></tr> <tr><td>55</td><td>1</td></tr> <tr><td>100</td><td>1</td></tr> </tbody> </table>	Frames	Count	5	5	10	8	15	13	20	8	25	9	30	3	35	1	45	1	55	1	100	1	<p>22.18</p>
Frames	Count																						
5	5																						
10	8																						
15	13																						
20	8																						
25	9																						
30	3																						
35	1																						
45	1																						
55	1																						
100	1																						



18.92



21.52



23.72

Τέλος, μετρήθηκε το ποσοστό επιτυχίας του προγράμματος για κάθε μια από τις παραπάνω τιμές της τυπικής απόκλισης. Για να μετρηθεί το ποσοστό αυτό υπολογίστηκε η απόλυτη διαφορά των πραγματικών frames αρχής και τέλους (groundtruth) με αυτά που δίνει το πρόγραμμα αντίστοιχα. Εάν η διαφορά αυτή ήταν λιγότερη από δέκα τότε θεωρούνταν ότι το πρόγραμμα βρήκε σωστά είτε την αρχή είτε το τέλος. Στον παρακάτω πίνακα(πιν.3) παρουσιάζονται τα ποσοστά επιτυχίας του προγράμματος στο σύνολο των κινήσεων για τις διαφορετικές τιμές της τυπικής απόκλισης.

Πίνακας 3

Τυπική απόκλιση	0.005	0.01	0.015	0.02	0.025
Start accuracy	14 %	82%	94%	84%	70%
Stop accuracy	52 %	38%	24%	14%	6%

Με βάση τα αποτελέσματα αυτά αποφασίστηκε ότι η τιμή της τυπικής απόκλισης για τον ορισμό της αρχής πρέπει να είναι πάνω από 0.015 ενώ η τιμή της τυπικής απόκλισης για τον ορισμό του τέλους πρέπει να είναι κάτω από 0.005.

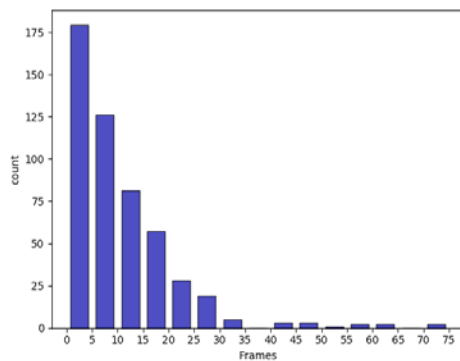
2.2.3.2 Κόψιμο δεδομένων 2D

Για το κόψιμο των δισδιάστατων κινήσεων χρησιμοποιήθηκε η ίδια μέθοδος με αυτή των δεδομένων 3D όπως περιγράφηκε προηγουμένως. Σαν αρχή ορίστηκε το σημείο όπου η τυπική απόκλιση ξεπερνά για πρώτη φορά την τιμή 1,5 ενώ το σημείο του τέλους ορίστηκε το σημείο όπου μετά από την αρχή της κίνησης για πρώτη φορά η τιμή της τυπικής απόκλισης πέσει κάτω από 1,5.

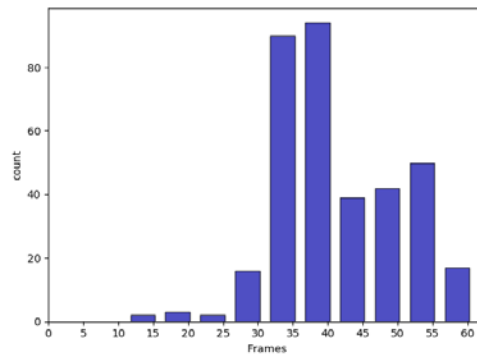
2.2.4 Σύνολα δεδομένων

Για το πείραμα κατασκευάστηκαν δύο διαφορετικά σύνολα δεδομένων. Το πρώτο σύνολο αποτελείται από την τρισδιάστατη πληροφορία κάθε κίνησης και αφορά τον πρώτο κύκλο πειραμάτων, ενώ το δεύτερο περιέχει μόνο την δισδιάστατη πληροφορία αυτής. Στο δεύτερο σύνολο εκτός από την δοκιμή για τις τρεις κλάσεις (μικρή - μεσαία – μεγάλη) έγιναν δοκιμές και για δυο κλάσεις. Αυτές οι δοκιμές κάλυπταν όλους τους συνδυασμούς των τριών κλάσεων. Παρακάτω εμφανίζεται η κατανομή των frames σε κάθε κίνηση σε μορφή ιστογράμματος για τα δυο σύνολα δεδομένων.

3D



2D



Εικόνα 16: Ιστόγραμμα Frames 3D

Εικόνα 17: Ιστόγραμμα Frames 2D

Όπως παρατηρούμε τα frames στο σύνολο δεδομένων 3D είναι πολύ λίγα. Ο λόγος που συμβαίνει αυτό είναι γιατί κατά τη μετατροπή από δισδιάστατη πληροφορία σε τρισδιάστατη χάνονται πάρα πολλά frames. Τα εναπομείναντα frames δεν είναι αρκετά για να κατασκευαστεί ένας αξιόλογος αλγόριθμος πρόβλεψης όπως θα δούμε στη συνέχεια. Αυτός είναι και ο λόγος που κατασκευάστηκε και το δεύτερο σύνολο δεδομένων. Όπως είναι φανερό τα εναπομείναντα frames είναι πολύ περισσότερα. Έχοντας διαθέσιμα αρκετά δεδομένα κατέστη εφικτό να κατασκευαστούν αρκετά καλοί αλγόριθμοι πρόβλεψης.

2.3 Καθορισμός χαρακτηριστικών

Για την εκπαίδευση του αλγορίθμου πρέπει να επιλεγθούν ποια χαρακτηριστικά των δεδομένων θα χρησιμοποιηθούν από όλα τα διαθέσιμα.

Τα χαρακτηριστικά που χρησιμοποιήθηκαν για την εκπαίδευση του αλγορίθμου είναι :

- Ο αύξων αριθμός του Frame
- Η θέση του καρπού στον άξονα X σε κάθε Frame
- Η θέση του καρπού στον άξονα Y σε κάθε Frame
- Η θέση του καρπού στον άξονα Z σε κάθε Frame
- Η θέση του αντίχειρα άξονα X σε κάθε Frame
- Η θέση του αντίχειρα άξονα Y σε κάθε Frame
- Η θέση του αντίχειρα άξονα Z σε κάθε Frame
- Η θέση του δείκτη άξονα X σε κάθε Frame
- Η θέση του δείκτη άξονα Y σε κάθε Frame
- Η θέση του δείκτη άξονα Z σε κάθε Frame
- Η απόσταση του αντίχειρα και του δείκτη σε κάθε Frame
- Η ταχύτητα του καρπού στον χώρο
- Η τυπική απόκλιση

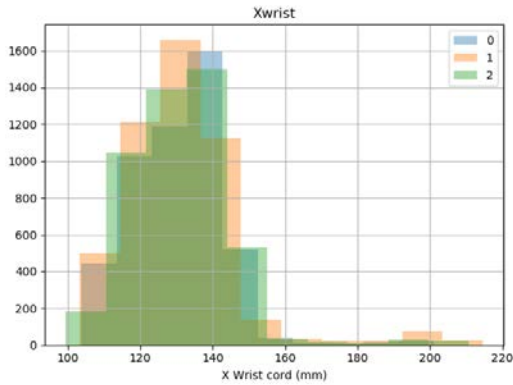
Τα ίδια χαρακτηριστικά χρησιμοποιήθηκαν και στα δυο σύνολα δεδομένων. Για το δεύτερο σύνολο όπως είναι προφανές δεν χρησιμοποιήθηκε η διάσταση Z.

Frame

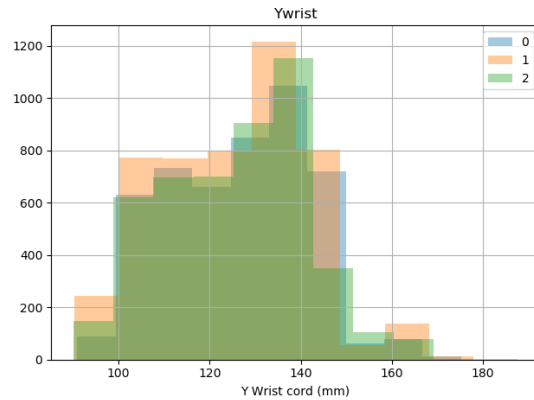
Το χαρακτηριστικό «Frame» αναφέρεται στο Frame που μετρήθηκαν τα εκάστοτε χαρακτηριστικά. Με το συγκεκριμένο χαρακτηριστικό δίνουμε στον αλγόριθμο την πληροφορία του χρόνου.

Θέση του καρπού στους άξονες X,Y, Z

Η θέση του καρπού αφορά της συντεταγμένες του κάπρου στους τρεις ή στους δυο άξονες ανάλογα το σύνολο δεδομένων, σε κάθε χρονική στιγμή.



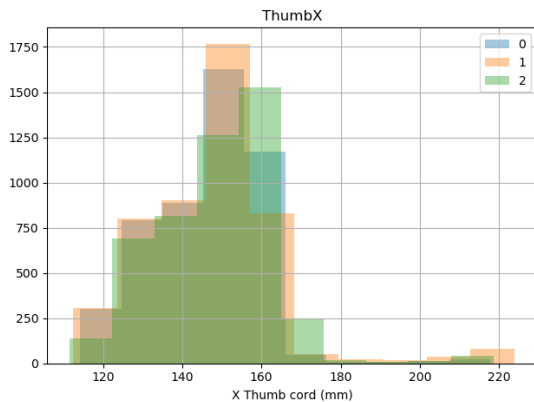
Εικόνα 18: Ιστόγραμμα θέσης καρπού στον άξονα Y για τρεις κλάσεις



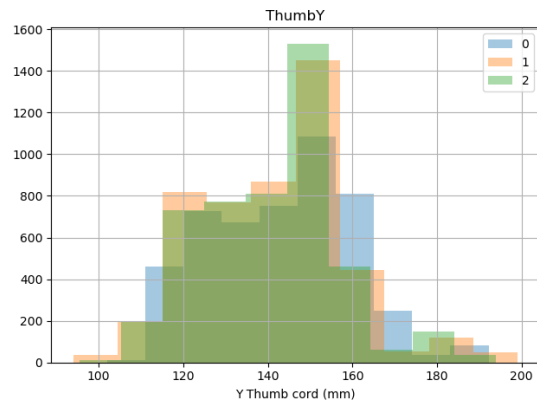
Εικόνα 19: Ιστόγραμμα θέσης καρπού στον άξονα X για τρεις κλάσεις

Θέση του αντίχειρα στους άξονες X,Y,Z

Η θέση του καρπού αφορά της συντεταγμένες του αντίχειρα στους τρεις ή στους δυο άξονες ανάλογα το σύνολο δεδομένων, σε κάθε χρονική στιγμή.



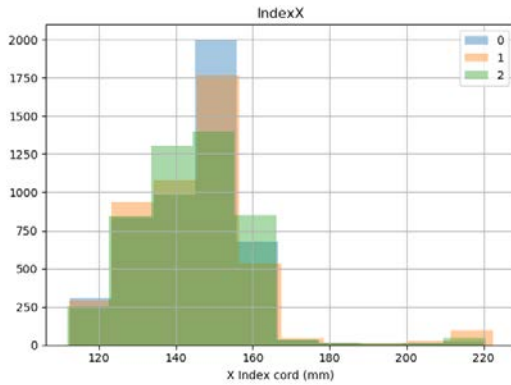
Εικόνα 18: Ιστόγραμμα θέσης αντίχειρα στον άξονα X για τρεις κλάσεις



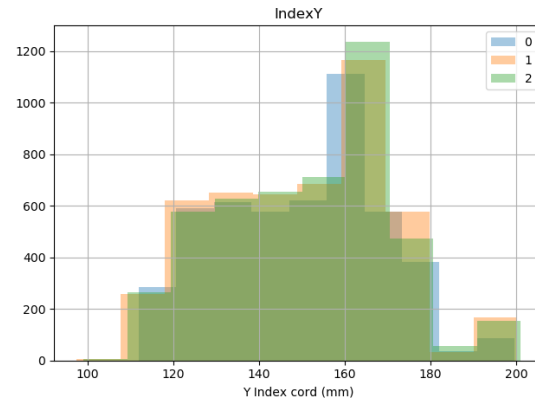
Εικόνα 19: Ιστόγραμμα αντίχειρα καρπού στον άξονα Y για τρεις κλάσεις

Θέση του δείκτη στους άξονες X,Y,Z

Η θέση του καρπού αφορά της συντεταγμένες του δείκτη στους τρεις ή στους δυο άξονες ανάλογα το σύνολο δεδομένων, σε κάθε χρονική στιγμή.



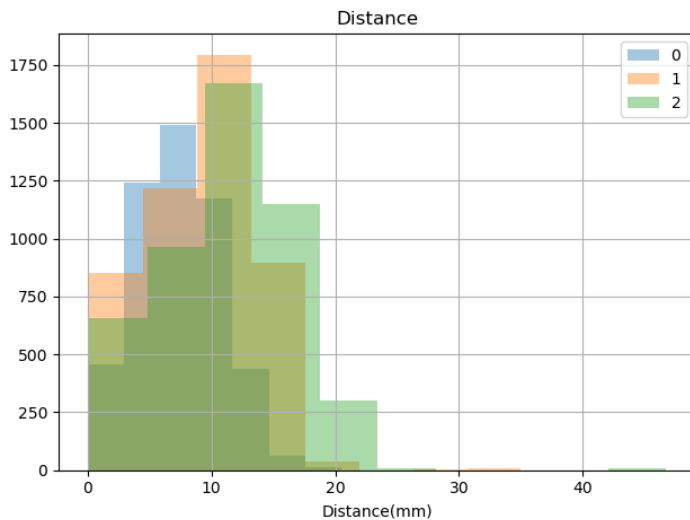
Εικόνα 20 Ιστόγραμμα θέσης δείκτη στον άξονα X για τρεις κλάσεις.



Εικόνα 21 Ιστόγραμμα θέσης δείκτη στον άξονα Y για τρεις κλάσεις.

Η απόσταση του αντίχειρα και του δείκτη σε κάθε Frame

Η απόσταση του αντίχειρα και του δείκτη αφορά την ευκλείδεια απόσταση που έχουν στον χώρο οι άκρες των δυο αυτών δακτύλων σε κάθε χρονική στιγμή.

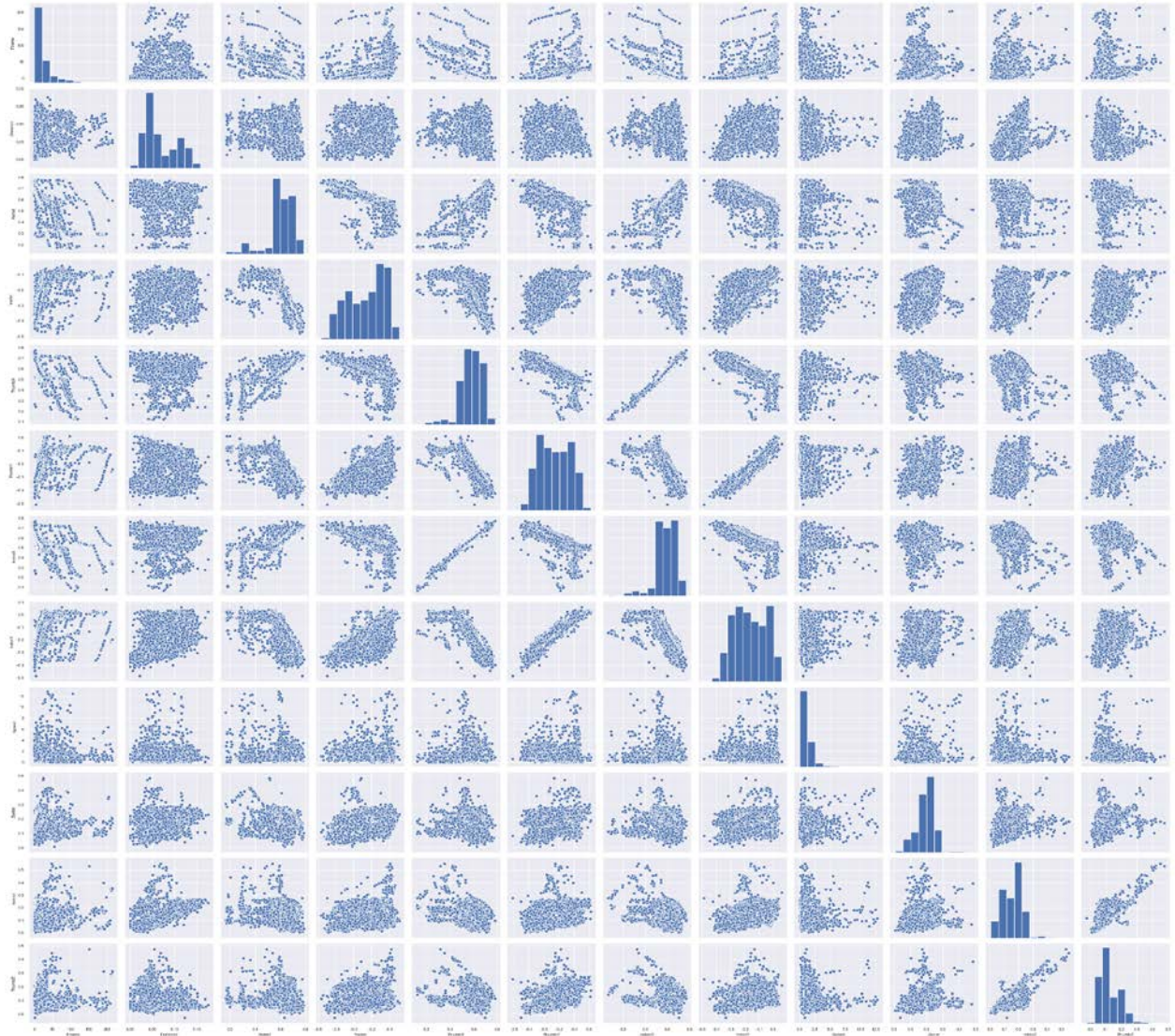


Εικόνα 22: Ιστόγραμμα απόστασης δείκτη αντίχειρα για τις τρεις κλάσεις

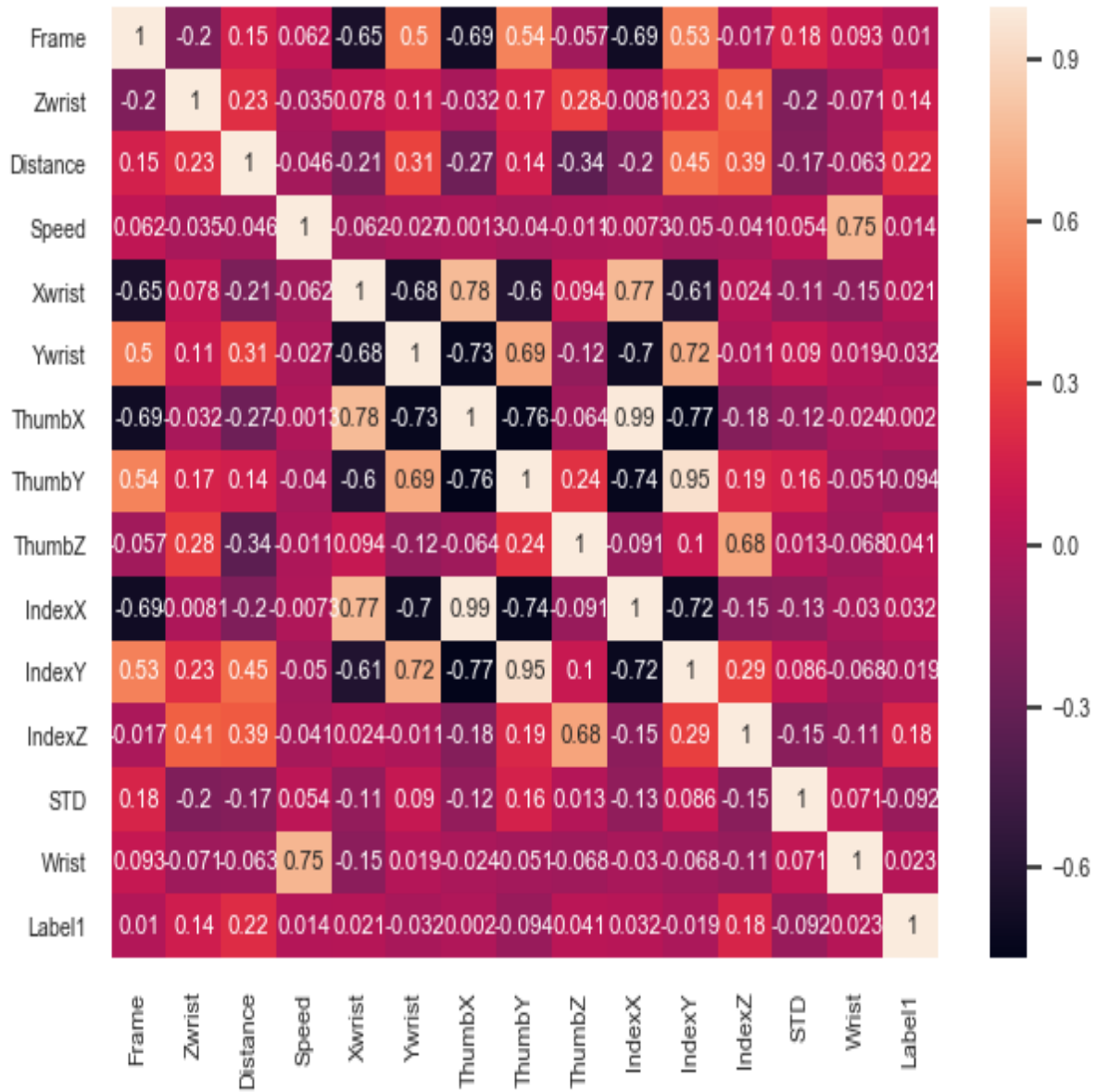
Η ταχύτητα του καρπού στον χώρο

Η ταχύτητα του καρπού στον χώρο αφορά την ταχύτητα που μετακινήθηκε ο καρπός από την θέση που βρισκόταν στο προηγούμενο frame στην θέση στην οποία βρίσκεται την χρονική στιγμή που γίνεται η μέτρηση.

Για να υπάρξει μια καλύτερη κατανόηση των χαρακτηριστικών απεικονίστηκαν το κάθε ένα ξεχωριστά σε σχέση με κάθε ένα από τα εναπομείναντα(εικ.23). Επίσης για τον ίδιο λόγο εξετάσαμε και τον βαθμό συσχέτισης τους(εικ.24). Όπως παρατηρούμε στο κέντρο υπάρχει μεγάλη συσχέτιση ανάμεσα στα χαρακτηριστικά που χρησιμοποιούμε. Ο λόγος που παρατηρείται αυτό είναι γιατί αυτά αφορούν τις συντεταγμένες των δακτύλων ενός χεριού επομένως είναι προφανές ότι θα έχουν υψηλό βαθμό συσχέτισης



Εικόνα 23 Γραφήματα μεταξύ όλων των χαρακτηριστικών



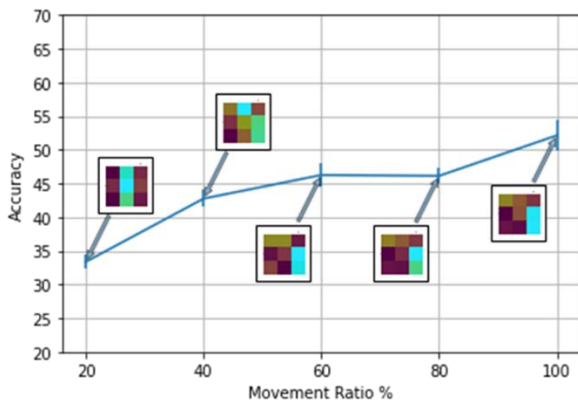
Εικόνα 24: Γράφημα συσχέτισης χαρακτηριστικών.

2.4 Γραφήματα αποτελεσμάτων

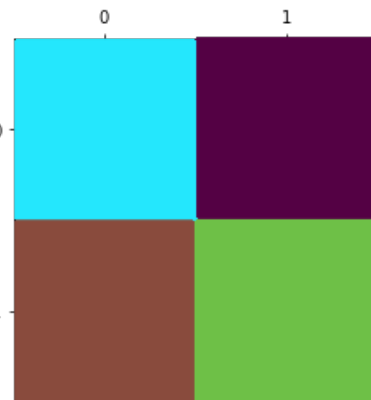
Σκοπός της διπλωματικής όπως αναφέρθηκε και προηγούμενος είναι η αναζήτηση ενός αλγορίθμου μηχανικής μάθησης που θα επιτρέψει σε ένα ρομπότ να προβλέψει πως θα δράσει ο άνθρωπος στον κοινό εργασιακό χώρο μέσα από οπτική πληροφορία. Επομένως έχοντας σκοπό την πρόβλεψη δεν μας ενδιαφέρει τόσο το συνολικό ποσοστό επιτυχίας του αλγορίθμου, δηλαδή στο σύνολο των δεδομένων τα ποσά κατέταξε σωστά. Μας ενδιαφέρει να έχουμε ένα ποσοστό επιτυχίας κατά την διάρκεια της κίνησης.

Για να γίνει αυτό, αρχικά μετρήθηκε από κάθε κίνηση το σύνολο των διαθέσιμων δεδομένων. Στη συνέχεια, χωρίστηκε σε πέντε κομμάτια και καταγράφηκε πόσα δεδομένα αντιστοιχούν σε κάθε κομμάτι της κίνησης. Τέλος κατά την διαδικασία των αποτελεσμάτων γνωρίζοντας για κάθε κομμάτι πόσα δεδομένα αντιστοιχούν μετρούσαμε το πόσα από αυτά ο αλγόριθμος κατέταξε στην σωστή κλάση. Με αυτό καταφέραμε να έχουμε μια εικόνα της ακρίβειας κάθε αλγορίθμου για τα διάφορα ποσοστά της κίνησης.

Στην παρακάτω εικόνα βλέπουμε ένα τυπικό γράφημα της ακρίβειας ενός αλγορίθμου.



Εικόνα 25 Διάγραμμα απόδοσης αλγορίθμου



Εικόνα 26: Παράδειγμα heatmap αλγορίθμου

Όπως παρατηρούμε στο γράφημα για κάθε αλγόριθμο έχουμε υπολογίσει την ακρίβεια του σε πέντε διαστήματα. Αυτά είναι $[0, 20]$, $(20, 40]$, $(40, 60]$, $(60, 80]$ και $(80, 100]$. Όπως αναφέρθηκε για να γίνει αυτό υπολογίσαμε τον αριθμό των συνολικών frames σε ένα διάστημα και μετρήσαμε το ποσοστό επιτυχίας στο διάστημα αυτό. Πέραν της ακρίβειας υπολογίσαμε και για κάθε διάστημα το πόσα δεδομένα κατατάσσονται σε κάθε κλάση. Όσο το χρώμα στον πίνακα που απεικονίζεται στο γράφημα πλησιάζει το γαλάζιο τόσο πιο πολλά δεδομένα έχουν μπει σε αυτήν την κλάση. Αντιθέτως όσο πιο κοντά είναι στο σκούρο καφέ τόσο τα δεδομένα στον πίνακα πλησιάζουν το μηδέν. Οι στήλες του πίνακα

δείχνουν την σωστή κλάση ενώ οι γραμμές την κλάση που κατέταξε ο αλγόριθμος τα δεδομένα. Επομένως ένας καλός αλγόριθμος είναι αυτός που έχει στην διαγώνιο του ανοιχτό χρώμα.

3 Αποτελέσματα

Σε αυτό το κεφάλαιο θα παρουσιαστούν τα αποτελέσματα των πειραμάτων για τα δυο σύνολα δεδομένων(τρισδιάστατο – δισδιάστατο) για όλους τους αλγορίθμους ξεχωριστά.

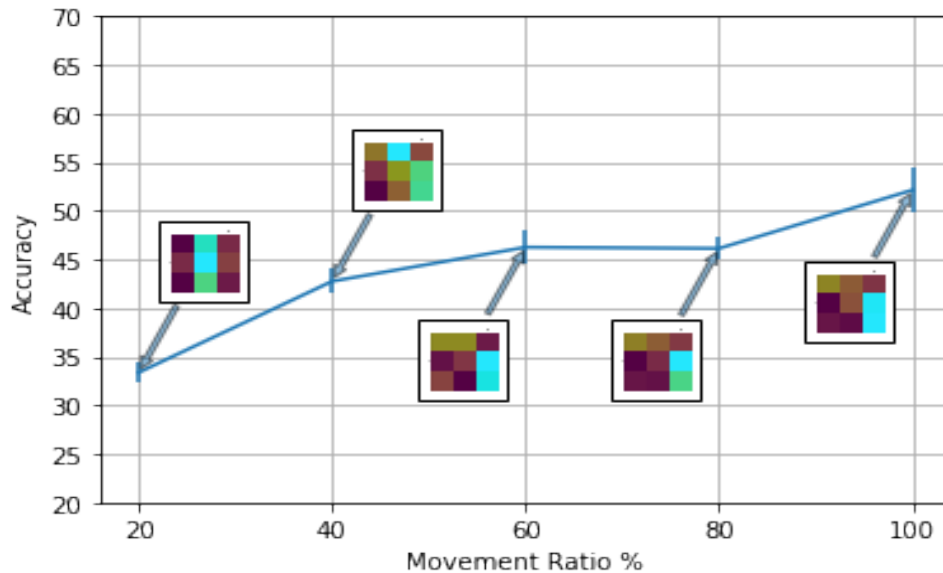
3.1 Τρισδιάστατο σύνολο δεδομένων.

Στα τρισδιάστατα δεδομένα δοκιμάστηκε μόνο ο SVM καθώς όπως προαναφέρθηκε διαπιστώθηκε ότι σε αυτά υπήρχε πολύ λίγη διαθέσιμη πληροφορία. Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα του αλγορίθμου για κάθε κομμάτι από την διαδικασία K-Fold καθώς επίσης και το τυπικό σφάλμα του μέσου όρου.

Πίνακας 4

SVM					
	20%	40%	60%	80%	100%
AVERAGE	33.37	42.75	46.25	46.12	52.12
STD	2.77	3.49	4.77	3.39	6.44
SE	0.98	1.23	1.68	1.20	2.27

Όπως παρατηρούμε τα αποτελέσματα για την αρχή της κίνησης (20%-40%) είναι 33% και 42% αντίστοιχα. Όσον αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 46% και φτάνει στο 52% στο τέλος της κίνησης. Οι τιμές αυτές προέρχονται από τον μέσο ορό των αποτελεσμάτων του αλγορίθμου σε κάθε κομμάτι της κίνησης, και έχουν τυπική απόκλιση από 2.7 στην αρχή της κίνησης ως 6.4 στο τέλος αυτής. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης καθώς η απόκλιση των τιμών από την διαδικασία του K-fold δεν είχαν μεγάλη απόκλιση. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 27: Διάγραμμα απόδοσης αλγορίθμου σε τρισδιάστατα δεδομένα(SVM).

Όπως βλέπουμε ο αλγόριθμος δεν καταφέρνει να ξεχωρίσει τις τρεις κλάσεις και τοποθετεί σωστά τα δεδομένα κυρίως στην κλάση μεγάλο αντικείμενο. Προς το τέλος της κίνησης ο αλγόριθμος μπερδεύει τελείως την κλάση μεσαίο αντικείμενο και τοποθετεί τα δεδομένα της στην μεγάλη κλάση.

3.2 Δισδιάστατο σύνολο δεδομένων με τρεις κλάσεις

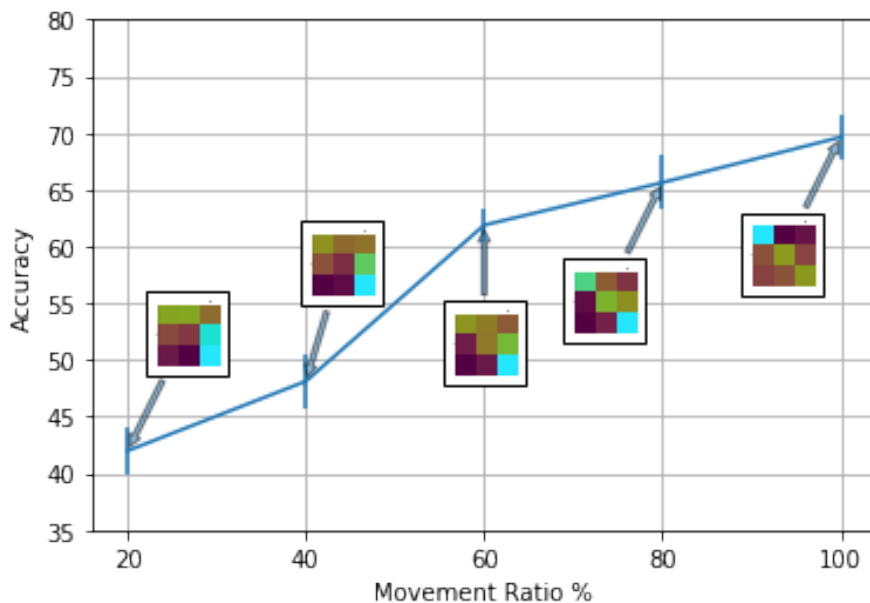
Για τα δεδομένα των δύο διαστάσεων με τρεις κλάσεις δοκιμάστηκαν εκτός από τον SVM και άλλοι τέσσερις αλγόριθμοι. Αυτοί ήταν οι Decision Tree, Random Forest, Extra Tree και ο Gradient. Παρακάτω βλέπουμε τους πίνακες και τα διαγράμματα της ακρίβειας κάθε αλγορίθμου ξεχωριστά καθώς και το τυπικό σφάλμα του μέσου όρου για τα διάφορα κομμάτια της κίνησης σε κάθε αλγόριθμο.

SVM

Πίνακας 5

SVM					
	20%	40%	60%	80%	100%
AVERAGE	42	48.12	61.87	65.62	69.62
STD	5.63	6.62	4.32	6.69	5.42
SE	1.99	2.34	1.52	2.36	1.91

Όπως παρατηρούμε τα αποτελέσματα για την αρχή της κίνησης (20%-40%) είναι 42% και 48% αντίστοιχα. Όσον αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 61% και φτάνει στο 69% στο τέλος της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 4 ως 6.7 στο τέλος αυτής. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης καθώς η απόκλιση των τιμών από την διαδικασία του K-fold δεν είχαν μεγάλη απόκλιση. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 28: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(SVM).

Όπως βλέπουμε στην αρχή της κίνησης, ο αλγόριθμος δεν καταφέρνει να ξεχωρίσει τις τρεις κλάσεις και τοποθετεί σωστά τα δεδομένα κυρίως της κλάσης μεγάλο αντικείμενο. Σε αντίθεση στο τέλος αυτής καταφέρνει να ξεχωρίσει τις τρεις κλάσεις με μεγαλύτερη ακρίβεια σε αυτή του μικρού αντικειμένου.

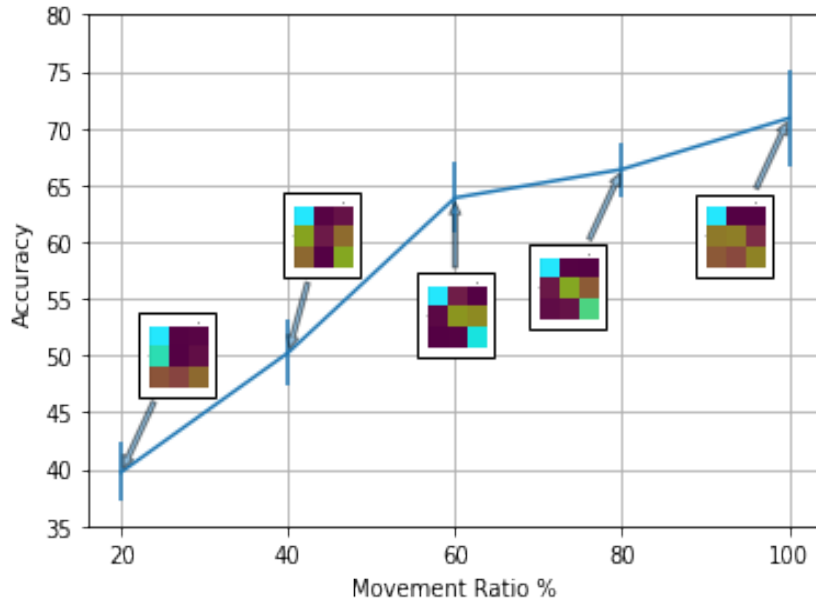
Gradient Boosting

Πίνακας 6

Gradient Boosting					
	20%	40%	60%	80%	100%
AVERAGE	35.62	50.125	60.25	66.87	71.37
STD	5.68	5.48	6.15	9.21	9.37
SE	2	1.94	2.17	3.25	3.31

Όπως παρατηρούμε τα αποτελέσματα για την αρχή της κινήσεις (20%-40%) είναι 35% και 50% αντίστοιχα. Όσον αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 60% και φτάνει στο 71% στο τέλος της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 5 ως 9 στο τέλος αυτής. Αυτό σημαίνει ότι ο αλγόριθμος δεν κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης καθώς οι αποκλίσεις των τιμών από την διαδικασία του K-fold είχαν μεγάλη απόκλιση. Παρόλο που τα αποτελέσματα δείχνουν καλύτερα από τον SVM βλέπουμε ότι στους μέσους ορού υπάρχουν μεγαλύτερες αποκλίσεις. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου

και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 29: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(Gradient Boosting).

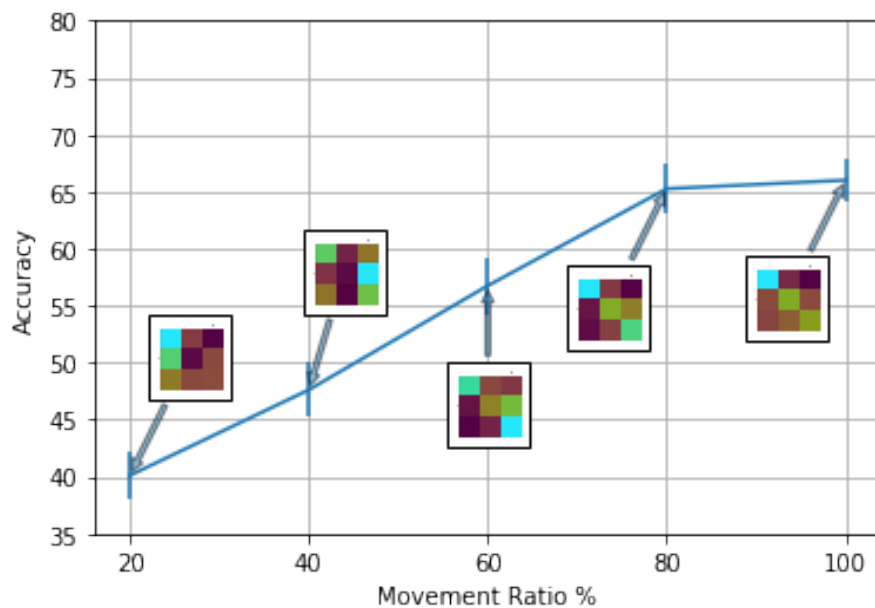
Όπως βλέπουμε ο αλγόριθμος δεν καταφέρνει να ξεχωρίσει τις τρεις κλάσεις και τοποθετεί σωστά τα δεδομένα κυρίως στη κλάση μικρό αντικείμενο στην αρχή της κίνησης ενώ προς το τέλος αυτής καταφέρνει να ξεχωρίσει και την κλάση μεγάλο αντικείμενο.

Decision Tree

Πίνακας 7

Decision Tree					
	20%	40%	60%	80%	100%
AVERAGE	40.12	47.62	56.75	65.25	66
STD	5.96	6.75	7	6.25	5.37
SE	2.1	2.38	2.47	2.2	1.89

Όπως παρατηρούμε τα αποτελέσματα για την αρχή της κίνησης (20%-40%) είναι 40% και 47% αντίστοιχα. Όσον αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 56% και φτάνει στο 66% στο τέλος της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 5 ως 7 στο τέλος αυτής. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης καθώς η αποκλίσεις των τιμών από την διαδικασία του K-fold δεν είχαν μεγάλη απόκλιση. Μάλιστα παρόλο που τα αποτελέσματα δείχνουν χειρότερα από τον Gradient Boosting βλέπουμε ότι στους μέσους ορού υπάρχουν μικρότερες αποκλίσεις. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 30:: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με τρεις κλάσεις(Decision Tree).

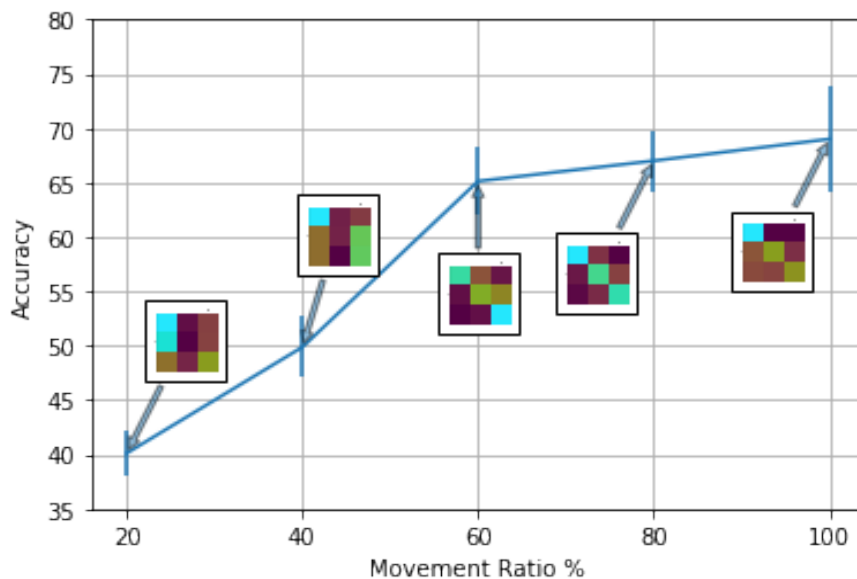
Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις τρεις κλάσεις προς το τέλος της κίνησης.

Random Forest

Πίνακας 8

Random Forest					
	20%	40%	60%	80%	100%
AVERAGE	40.12	49.87	65.12	67	69
STD	6.03	7.88	8.85	7.81	13.64
SE	2.13	2.78	3.13	2.76	4.82

Όπως παρατηρούμε τα αποτελέσματα για την αρχή της κίνησης (20%-40%) είναι 40% και 49% αντίστοιχα. Όσον αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 65% και φτάνει στο 69% στο τέλος της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 6 ως 13 στο τέλος αυτής. Αυτό σημαίνει ότι ο αλγόριθμος δεν κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Μάλιστα τα αποτελέσματα είναι χειρότερα από τον Gradient Boosting και τον Decision Tree καθώς στους μέσους ορού υπάρχουν πολύ μεγάλες αποκλίσεις. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 31: Διάγραμμα απόδοσης αλγορίθμου σε διδιάστατα δεδομένα με τρεις κλάσεις(Random Forest)

Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις τρεις κλάσεις προς το τέλος της κίνησης. Σε αντίθεση με την αρχή οπου κατατάσσει σωστά μόνο τα δεδομένα της κλάσης μικρό αντικείμενο.

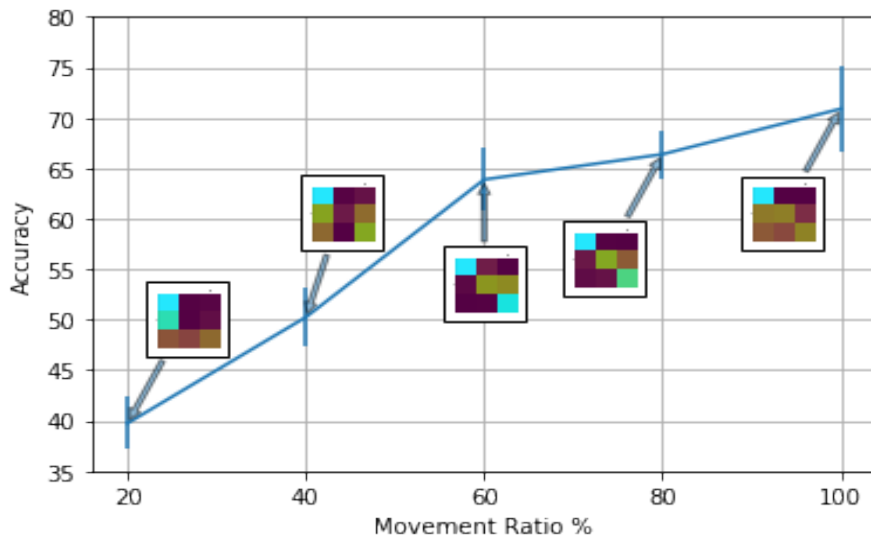
Extra Trees

Πίνακας 9

Extra Trees					
	20%	40%	60%	80%	100%
AVERAGE	39.75	50.25	63.87	66.37	70.87
STD	7.32	7.97	8.85	6.65	11.76
SE	2.58	2.82	3.13	2.35	4.15

Όπως παρατηρούμε τα αποτελέσματα για την αρχή της κίνησης (20%-40%) είναι 39% και 50% αντίστοιχα. Όσον αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 65% και φτάνει στο 70% στο τέλος της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 6 ως 11 στο τέλος αυτής. Αυτό σημαίνει ότι ο αλγόριθμος δεν κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Στο παρακάτω διάγραμμα

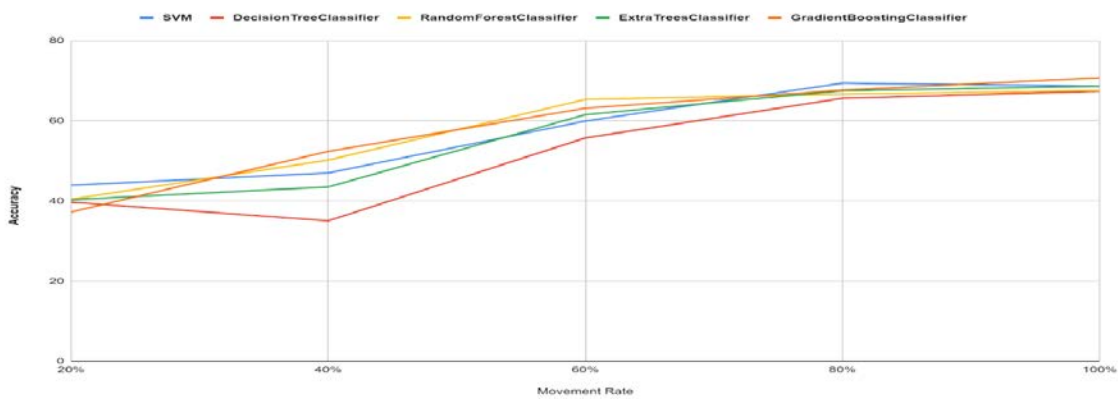
βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 32: Διάγραμμα απόδοσης αλγορίθμου σε διδιάστατα δεδομένα με τρεις κλάσεις(Extra Tree).

Όπως βλέπουμε ο αλγόριθμος δεν καταφέρνει να ξεχωρίσει τις τρεις κλάσεις και τοποθετεί τα δεδομένα σωστά κυρίως στην κλάση μικρό αντικείμενο.

Συγκεντρωτικά αποτελέσματα



Εικόνα 33: Συγκεντρωτικό διάγραμμα απόδοσης αλγορίθμων σε διδιάστατα δεδομένα με τρεις κλάσεις

Όπως μπορούμε να παρατηρήσουμε όλοι οι αλγόριθμοι έχουν ποσοστό ακρίβειας που κυμαίνεται από 35% για την αρχή της κίνησης έως και 70% για το τέλος αυτής. Όμως κοιτώντας την κατανομή των δεδομένων σε κάθε στιγμιότυπο της κινήσεως βλέπουμε ότι τα δεδομένα κατατάσσονται στις κλάσεις μικρό και μεγάλο αντικείμενο παραβλέποντας

την κλάση μεσαίο αντικείμενο. Επίσης όλοι οι αλγόριθμοι έχουν αρκετά υψηλό σφάλμα στο μέσο όρο των τιμών της ακρίβειας σε κάθε κομμάτι της κινήσεις.

3.3 Δισδιάστατο σύνολο δεδομένων με δυο κλάσεις

Στους παρακάτω πίνακες βλέπουμε την απόδοσή των αλγορίθμων για τα δισδιάστατα δεδομένα με δυο κλάσεις για όλους τους συνδυασμούς των τριών κλάσεων.

Μικρό – Μεσαίο

Πίνακας 10

SVM					
	20%	40%	60%	80%	100%
AVERAG E	57.62	63	67.62	71.25	72.12
STD	5.62	8.12	12.15	16.96	18.81
SE	1.99	2.87	4.29	5.99	6.65

Πίνακας 11

Gradient Boosting					
	20%	40%	60%	80%	100%
AVERAG E	57.25	57.5	64.87	71.37	73.75
STD	9.80	11.64	19.48	20.38	17.77
SE	3.46	4.11	6.88	7.2	6.28

Πίνακας 12

Decision Tree					
	20%	40%	60%	80%	100%
AVERAG E	57.25	56.62	64.62	72.25	76.12
STD	7.75	10.25	19.77	19.63	17.33
SE	2.74	3.62	6.99	6.94	6.12

Πίνακας 13

Random Forest					
	20%	40%	60%	80%	100%
AVERAG E	56.5	56.62	65.12	71.87	75.87
STD	7.72	11.75	18.77	20.79	18.69
SE	2.73	4.15	6.63	7.35	6.61

Πίνακας 14

Extra Trees					
	20%	40%	60%	80%	100%
AVERAG E	57.25	57.62	66	72.87	76.25
STD	9.22	11.42	18.91	19.79	17.86
SE	3.26	4.03	6.68	6.99	6.31

Μεσαίο – Μεγάλο

Πίνακας 15

SVM					
	20%	40%	60%	80%	100%
AVERAG E	48.5	50.87	55.25	57.12	45
STD	4.3	8.21	9.86	9.77	12.75
SE	1.52	2.9	3.48	3.45	4.5

Πίνακας 16

Gradient Boosting					
	20%	40%	60%	80%	100%
AVERAG E	47.5	50.62	59.37	59.5	51.12
STD	4.75	5.01	11.99	13.32	12.83
SE	1.67	1.77	4.24	4.7	4.53

Πίνακας 17

Decision Tree					
	20%	40%	60%	80%	100%
AVERAG E	47.5	54.25	56.37	54.62	56
STD	4.5	2.71	9.78	8.89	4.4
SE	1.59	0.95	3.45	3.14	1.55

Πίνακας 18

Random Forest					
	20%	40%	60%	80%	100%
AVERAG E	49.37	53.87	57.37	57.5	58.37
STD	3.06	3.87	12.03	14.66	8.61
SE	1.08	1.36	4.25	5.18	3.04

Πίνακας 19

Extra Trees					
	20%	40%	60%	80%	100%
AVERAG E	44.5	50.62	55.37	53	41.5
STD	6.45	7.30	10.97	13.38	12.15
SE	2.28	2.58	3.88	4.73	4.29

Μικρό – Μεγάλο

Πίνακας 20

SVM					
	20%	40%	60%	80%	100%
AVERAG E	51.25	51.25	55.125	61.5	66.375
STD	5.89	7.88	9.62	6.16	9.53
SE	2.08	2.78	3.4	2.17	3.36

Πίνακας 21

Gradient Boosting					
	20%	40%	60%	80%	100%
AVERAG E	51.125	55	68.75	84.75	93.5
STD	8.23	6.45	7.66	8.89	8.36
SE	2.91	2.28	2.71	3.14	2.95

Πίνακας 22

Decision Tree					
	20%	40%	60%	80%	100%
AVERAG E	58.375	57.125	70.75	76.625	73.875
STD	11.26	7.66	5.33	6.69	12.87
SE	3.98	2.7	1.88	2.36	4.55

Πίνακας 23

Random Forest					
	20%	40%	60%	80%	100%
AVERAG E	54.12	57.62	69.25	86.37	95.37
STD	8.77	9.02	7.12	9.92	8.34
SE	3.1	3.19	2.51	3.5	2.95

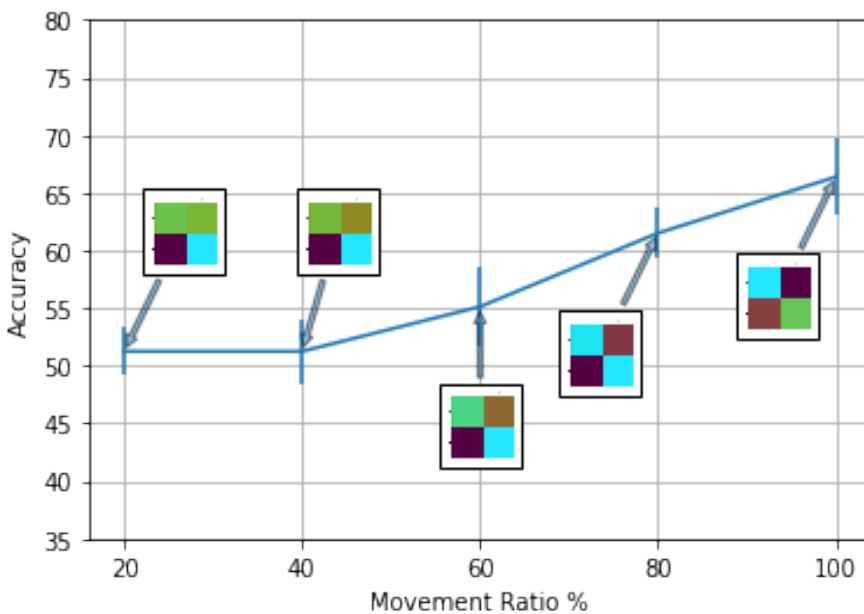
Πίνακας 24

Extra Trees					
	20%	40%	60%	80%	100%
AVERAG E	51.87	56.62	73.75	88	93.12
STD	10.72	9.17	15.06	10.78	8.18
SE	3.79	3.24	5.32	3.81	2.89

Όπως βλέπουμε ο συνδυασμός των κλάσεων μικρού και μεγάλου αντικειμένου έχει την μεγαλύτερη απόδοση. Παρακάτω βλέπουμε πιο αναλυτικά τα γραφήματα απόδοσης των αλγορίθμων σε αυτές τις κλάσεις καθώς και τα γραφήματα με το σφάλμα του μέσου όρου.

SVM

Όπως παρατηρούμε τα αποτελέσματα(πιν.20) για την αρχή της κίνησης (20%-60%) είναι 51% και 55% αντίστοιχα. Όσο αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 61% και φτάνει στο 66% στο τέλος της κίνησης. Οι τιμές αυτές προέρχονται από τον μέσο ορό των αποτελεσμάτων του αλγορίθμου σε κάθε κομμάτι της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 5 ως 9. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



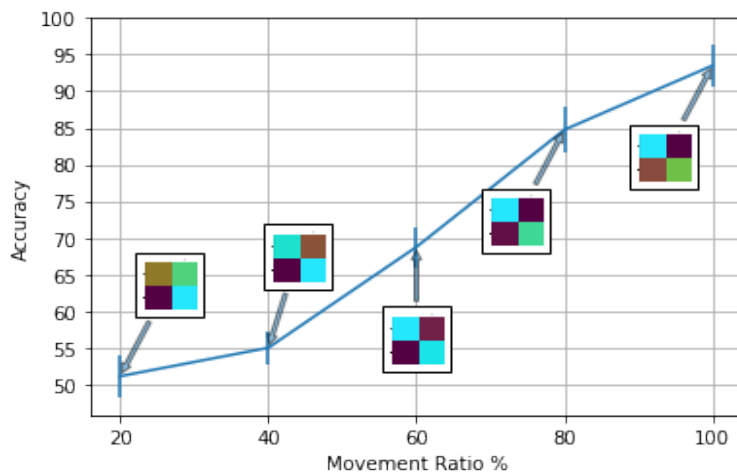
Εικόνα 34: Διάγραμμα απόδοσης αλγορίθμου σε διδιάστατα δεδομένα με δυο κλάσεις(SVM)

Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις δυο κλάσεις αρκετά καλά.

Gradient Boosting

Όπως παρατηρούμε τα αποτελέσματα(πιν.21) για την αρχή της κίνησης (20%-40%) είναι 51% και 55% αντίστοιχα. Όσο αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 68% και φτάνει στο 93% στο τέλος της κίνησης. Οι τιμές αυτές προέρχονται από τον μέσο ορό των αποτελεσμάτων του αλγορίθμου σε κάθε κομμάτι της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 6 ως 9. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Στο παρακάτω διάγραμμα

βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



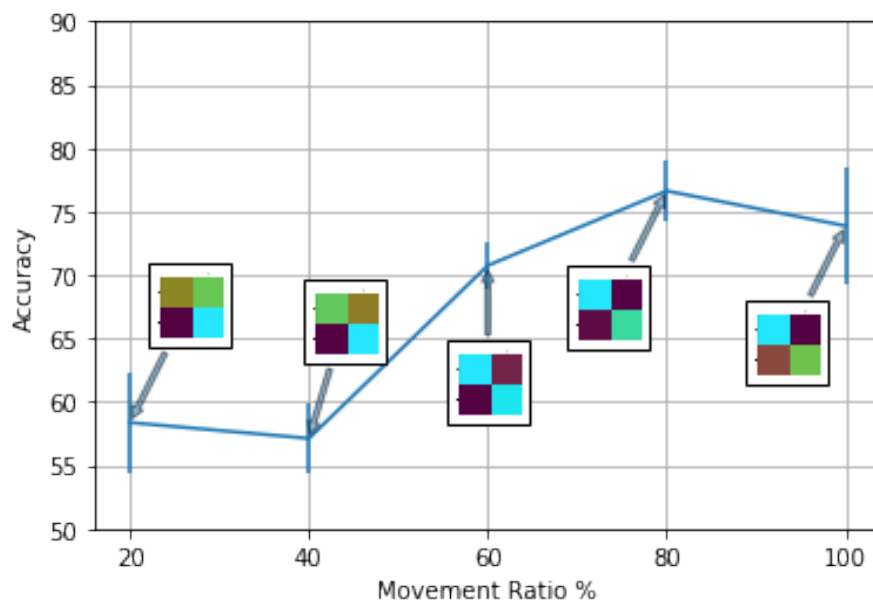
Εικόνα 35: Διάγραμμα απόδοσης αλγορίθμου σε διδιάστατα δεδομένα με δυο κλάσει(Gradient Boosting).

Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις δυο κλάσεις αρκετά καλά.

Decision Tree

Όπως παρατηρούμε τα αποτελέσματα(πιν.22) για την αρχή της κίνησης (20%-40%) είναι 58% και 57% αντίστοιχα. Όσο αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 70% και φτάνει στο 74% στο τέλος της κίνησης. Οι τιμές αυτές προέρχονται από τον μέσο ορό των αποτελεσμάτων του αλγορίθμου σε κάθε κομμάτι της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 6 ως 12. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Στο παρακάτω διάγραμμα

βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.

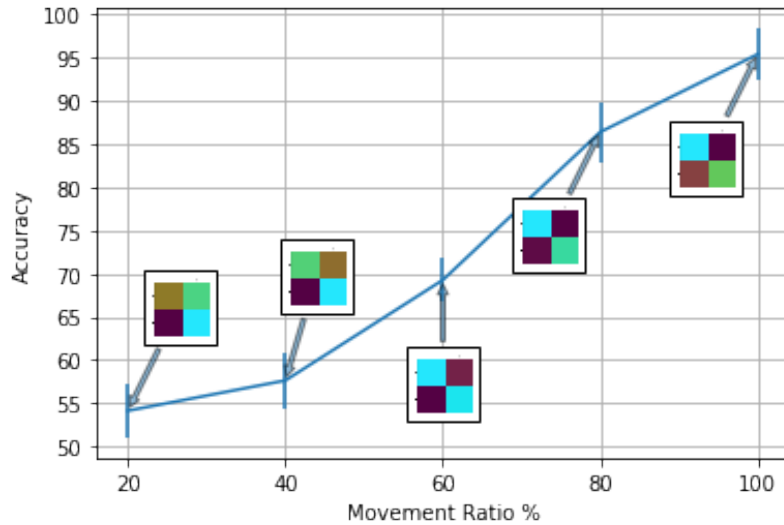


Εικόνα 36: Διάγραμμα απόδοσης αλγορίθμου σε διδιάστατα δεδομένα με δυο κλάσεις(Decision Tree).

Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις δυο κλάσεις αρκετά καλά.

Random Forest

Όπως παρατηρούμε τα αποτελέσματα (πιν.23)για την αρχή της κίνησης (20%-40%) είναι 54% και 57% αντίστοιχα. Όσο αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 69% και φτάνει στο 95% στο τέλος της κίνησης. Οι τιμές αυτές προέρχονται από τον μέσο ορό των αποτελεσμάτων του αλγορίθμου σε κάθε κομμάτι της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 7 ως 9. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Στο παρακάτω διάγραμμα βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



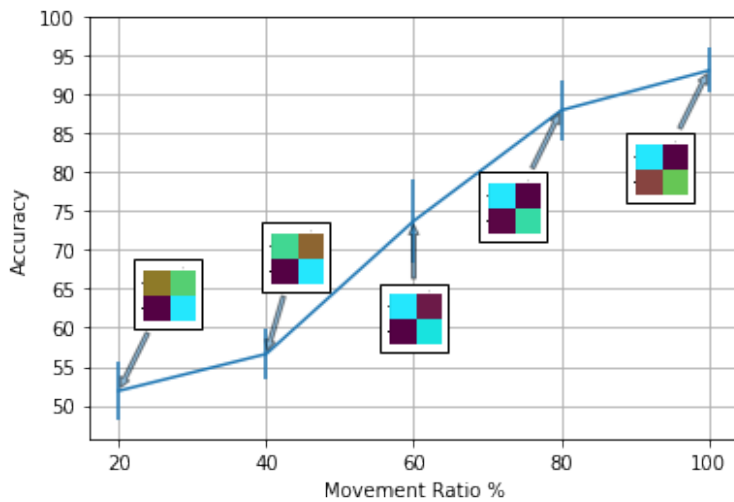
Εικόνα 37: Διάγραμμα απόδοσης αλγορίθμου σε δισδιάστατα δεδομένα με δυο κλάσεις(Random Forest).

Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις δυο κλάσεις αρκετά καλά.

Extra Tree

Όπως παρατηρούμε τα αποτελέσματα(πιν.24) για την αρχή της κίνησης (20%-40%) είναι 33% και 42% αντίστοιχα. Όσο αφορά την μέση της κίνησης η απόδοση του αλγορίθμου ανεβαίνει στο 46% και φτάνει στο 52% στο τέλος της κίνησης. Οι τιμές αυτές προέρχονται από τον μέσο ορό των αποτελεσμάτων του αλγορίθμου σε κάθε κομμάτι της κίνησης. Οι τιμές αυτές έχουν τυπική απόκλιση από 9 ως 15. Αυτό σημαίνει ότι ο αλγόριθμος κατάφερε να κάνει μια καλή γενίκευση των δεδομένων εκπαίδευσης. Στο παρακάτω διάγραμμα

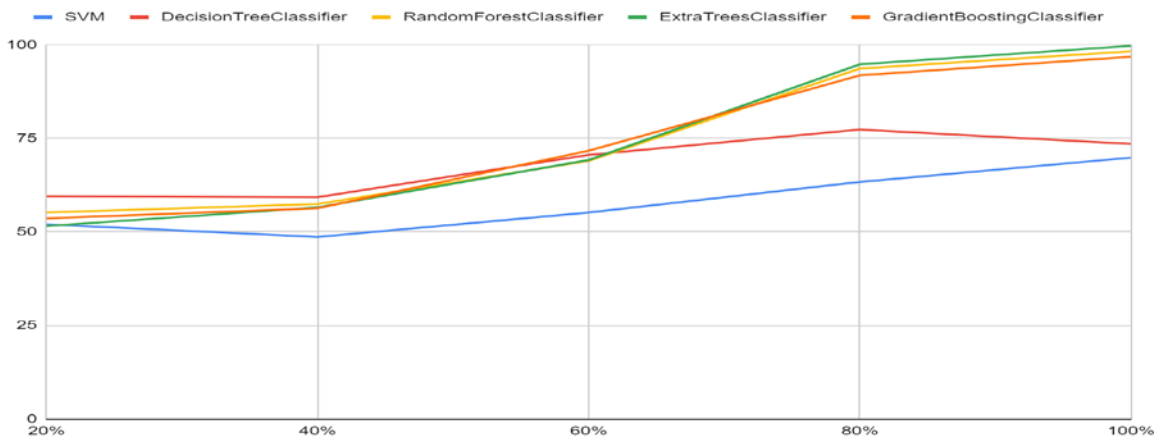
βλέπουμε εκτός από την απόδοση του αλγορίθμου και την κατανομή των δεδομένων από τον αλγόριθμο στις διάφορες κλάσεις.



Εικόνα 38:: Διάγραμμα απόδοσης αλγορίθμου σε διδιάστατα δεδομένα με δυο κλάσεις (Extra Tree).

Όπως βλέπουμε ο αλγόριθμος καταφέρνει να ξεχωρίσει τις δυο κλάσεις αρκετά καλά.

Συγκεντρωτικά αποτελέσματα



Εικόνα 39: Συγκεντρωτικό διάγραμμα απόδοσης αλγορίθμων σε διδιάστατα δεδομένα με δυο κλάσεις.

Όπως μπορούμε να παρατηρήσουμε όλοι οι αλγόριθμοι έχουν ποσοστό ακρίβειας που κυμαίνεται από 50% για την αρχή της κίνησης έως και 94% για το τέλος αυτής. Επίσης όλοι οι αλγόριθμοι έχουν υψηλότερο σφάλμα στον μέσο ορό των τιμών της ακρίβειας. Αυτό προκύπτει από τα χαμηλά ποσοστά ακρίβειας κάθε αλγορίθμου στην προσπάθεια να προβλέψει τις κινήσεις ενός συγκεκριμένου ατόμου.

4 Συμπεράσματα

Για να πετύχουμε τον σκοπό της διπλωματικής θέλουμε να έχουμε την πρόβλεψη όσο πιο νωρίς γίνεται ώστε να προλάβει κάποιο ρομπότ να αντιδράσει. Επομένως δεν μας αφορά η ακρίβεια του αλγορίθμου στο τέλος της κάθε κίνησης αλλά θέλουμε να πετύχουμε ένα καλό ποσοστό ακρίβειας όσο το πιο κοντά στην αρχή αυτής γίνεται. Στον παρακάτω πίνακα βλέπουμε τα αποτελέσματα των αλγορίθμων για όλα τα σύνολα δεδομένων στο 80% της κίνησης.

Πίνακας 25

Accuracy	3D	2D/3Class	2D/2Class
Algorithms			
SVM	44%	69%	63%
Decision Tree	-	65%	77%
Random Forest	-	66%	93%
Extra Tree	-	67%	94%
Gradient Boosting	-	67%	91%

Διαπιστώνουμε ότι το καλύτερο σύνολο δεδομένων για το πείραμά μας είναι το δισδιάστατο με τις δύο μόνο κλάσεις και ο καλύτερος αλγόριθμος με βάση την μέση απόδοση είναι ο Extra Tree με μικρή διαφορά από τους Random Forest και Gradient Boosting. Όμως το πιο γενικό μοντέλο το οποίο πετυχαίνει μια καλή απόδοση στο σύνολο των δεδομένων είναι αυτό του Gradient Boosting αφού έχει και το μικρότερο εύρος στο σφάλμα του μέσου όρου.

Το 80% της κίνησης είναι αρκετά νωρίς ώστε να αντιδράσει ένα ρομπότ;

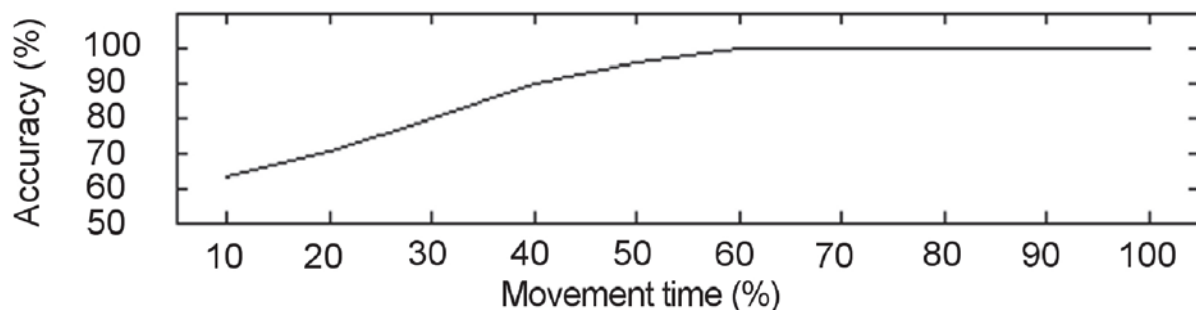
Γνωρίζοντας ότι κάθε κίνηση αποτελείται κατά μέσο ορό από 40 frames όπως βλέπουμε στο ιστόγραμμα(εικ.17) και ότι η καταγραφή έγινε στα 60 hz μια πρόβλεψη στο 80% της κίνησης δίνει σε ένα ρομποτικό βραχίονα για παράδειγμα περίπου 100 ms για να αντιδράσει. Εάν θέλουμε να κάνουμε μια πρόβλεψη πιο νωρίς θα χάσουμε λίγο σε ακρίβεια. Στον πίνακα 8 βλέπουμε τα τελικά αποτελέσματα του καλύτερου αλγορίθμου στο 60% και στο 80% της κίνησης για το δισδιάστατο σύνολο δεδομένων με τις δυο κλάσεις(μικρό – μεγάλο).

Πίνακας 26

Algorithms	Movement Ratio	Movement Ratio
	60%	80%
Gradient Boosting	71%	91%

Όπως παρατηρούμε ο αλγόριθμος έχει μια απόδοση 71% στα δεδομένα μας. Όμως και στο 60 % της κίνησης να γίνει η προβλέψει το παράθυρο αντίδρασης είναι μόλις 200ms. Επομένως το κατά πόσο το μοντέλο που κατασκευάσαμε είναι ακριβές εξαρτάται άμεσα από την ταχύτητα απόκρισης που θα έχει το ρομπότ το οποίο θα τον χρησιμοποιεί.

Στη ερευνά Predicting Object Size from Hand Kinematics[42] τα αποτελέσματα ήταν σαφώς καλύτερα στο τέλος της κίνησης φτάνοντας σχεδόν στο 100% όπως μπορούμε να δούμε στην παρακάτω εικόνα(εικ.40).



Εικόνα 40: Αποτελέσματα Predicting Object Size from Hand Kinematics

Ακόμη μπορούμε να παρατηρήσουμε ότι τα αποτελέσματα από τα αρχικά στάδια της κίνησής είναι κοντά στο 70%. Αυτό θα έδινε αρκετά νωρίς και με μεγάλη ακρίβεια μια πρόβλεψη σε ένα ρομπότ επομένως και αρκετό χρόνο αντίδρασης.

Όμως ο τρόπος συλλογής των δεδομένων αυτής της έρευνας δεν είναι τόσο αντιπροσωπευτικός για πραγματικές συνθήκες εργασίας καθώς για να συλλέγουν τα δεδομένα χρειάστηκε να τοποθετήσουν ειδικά σημάδια στους συμμετέχοντες ώστε να καταφέρουν να εντοπιστούν μέσα από τα βίντεο και η καταγραφή τους έγινε από δέκα κάμερες. Αντίθετα στο δικό μας πείραμα υπήρχε μια κάμερα όπως συνηθίζεται να τοποθετείτε σε ένα ρομπότ και η εύρεση του ανθρώπου στο βίντεο γίνεται αυτόματα. Λαμβάνοντάς υπόψιν τα παραπάνω μπορούμε να πούμε ότι η ερευνά μας μπορεί να έχει από μια πρώτη ανάγνωση χειρότερα αποτελέσματα από μια παρόμοια αλλά είναι πιο αντιπροσωπευτική όσον αφορά τις συνθήκες που επικρατούν σε μια συνεργασία ανθρώπου ρομπότ.

5 Βιβλιογραφία

1. Vollmer, A. L., & Schillingmann, L. (2018). On studying human teaching behavior with robots: a review. *Review of Philosophy and Psychology*, 9(4), 863-903.
2. Liu, R., & Zhang, X. (2017). Systems of natural-language-facilitated human-robot cooperation: A review. *arXiv preprint arXiv:1701.08269*.
3. Pandey, A. K., & Alami, R. (2014). Towards human-level semantics understanding of human-centered object manipulation tasks for hri: Reasoning about effect, ability, effort and perspective taking. *International Journal of Social Robotics*, 6(4), 593-620.
4. Akgun, B., & Thomaz, A. (2016). Simultaneously learning actions and goals from demonstration. *Autonomous Robots*, 40(2), 211-227.
5. Vollmer, A. L., Mühlig, M., Steil, J. J., Pitsch, K., Fritsch, J., Rohlfing, K. J., & Wrede, B. (2014). Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning. *PloS one*, 9(3), e91349.
6. Baraglia, J., Cakmak, M., Nagai, Y., Rao, R., & Asada, M. (2016, March). Initiative in robot assistance during collaborative task execution. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)* (pp. 67-74). IEEE.
7. Kyrarini, M., Haseeb, M. A., Ristić-Durrant, D., & Gräser, A. (2019). Robot learning of industrial assembly task via human demonstrations. *Autonomous Robots*, 43(1), 239-257.
8. Kennington, C., & Schlangen, D. (2017). A simple generative model of incremental reference resolution for situated dialogue. *Computer Speech & Language*, 41, 43-67.
9. Paul, R., Arkin, J., Roy, N., & Howard, T. M. (2017, January). Grounding Abstract Spatial Concepts for Language Interaction with Robots. In *IJCAI* (pp. 4929-4933).
10. Chai, J. Y., She, L., Fang, R., Ottarson, S., Littlely, C., Liu, C., & Hanson, K. (2014, March). Collaborative effort towards common ground in situated human-robot

- dialogue. In 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 33-40). IEEE.
11. Admoni, H., Weng, T., & Scassellati, B. (2016, May). Modeling communicative behaviors for object references in human-robot interaction. In 2016 IEEE International Conference on Robotics and Automation (ICRA) (pp. 3352-3359). IEEE.
 12. Mehlmann, G., Häring, M., Janowski, K., Baur, T., Gebhard, P., & André, E. (2014, November). Exploring a model of gaze for grounding in multimodal HRI. In Proceedings of the 16th International Conference on Multimodal Interaction (pp. 247-254).
 13. Chadalavada, R. T., Andreasson, H., Krug, R., & Lilienthal, A. J. (2015, September). That's on my mind! robot to human intention communication through on-board projection on shared floor space. In 2015 European Conference on Mobile Robots (ECMR) (pp. 1-6). IEEE.
 14. Pereira, A., Carter, E. J., Leite, I., Mars, J., & Lehman, J. F. (2017). Augmented reality dialog interface for multimodal teleoperation. In 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) (pp. 764-771). IEEE.
 15. Zanchettin, A. M., Ceriani, N. M., Rocco, P., Ding, H., & Matthias, B. (2015). Safety in human-robot collaborative manufacturing environments: Metrics and control. *IEEE Transactions on Automation Science and Engineering*, 13(2), 882-893.
 16. Morato, C., Kaipa, K. N., Zhao, B., & Gupta, S. K. (2014). Toward safe human robot collaboration by using multiple kinects based real-time human tracking. *Journal of Computing and Information Science in Engineering*, 14(1).
 17. Wilson, R. A., & Keil, F. C. (Eds.). (2001). *The MIT encyclopedia of the cognitive sciences*. MIT press.
 18. Edgar, T. W., & Manz, D. O. (2017). *Research methods for cyber security*. Syngress.
 19. Εικόνα 1 από <https://medium.com/@alphaxsalt/machine-learning-basics-be1d1be0eff1>

20. Εικόνα 2 από <https://medium.com/marketing-and-entrepreneurship/10-companies-using-machine-learning-in-cool-ways-887c25f913c3>
21. “Support Vector Machines notes”, R.Berwick, MIT University Understanding Support Vector Machine algorithm from examples”, Sunil Ray, Sep. 2017
22. Gandhi, R. (2018). Introduction to machine learning algorithms: Linear regression. Toward Data Science.
23. Patel, S. (2017). Chapter 2: SVM (support vector machine)—theory. Machine learning, 101.
24. Yiu, T. (2019). Understanding random forest. Towardsdatascience. com. June, 12.
25. Kam, H. T. (1995). Random decision forests/Ho Tin Kam. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC.—1995.
26. Breiman, L. (2015). Random forests leo breiman and adele cutler. Random Forests- Classification Description.
27. Ceballos, F. (2019). An Intuitive Explanation of Random Forest and Extra Trees Classifiers. Medium. July, 17.
28. Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. Machine learning, 63(1), 3-42.
29. Bhandari, N. (2018). How does extratreesclassifier reduce the risk of overfitting?. Medium, Oct, 22.
30. D’Souza, J. (2018). A quick guide to boosting in ml.
31. Grover, P. (2017). Gradient Boosting from scratch. Retrieved from Medium.
32. Singh, H. (2018). Understanding gradient boosting machines. Towards Data Science, 3.
33. Brownlee, J. (2016). A gentle introduction to xgboost for applied machine learning. Machine Learning Mastery. Available online: <http://machinelearningmastery.com/gentle-introduction-xgboost-appliedmachine-learning/>(accessed on 2 March 2018).
34. Drakos, G. (2020, February 05). Importance of cross-validation. Retrieved April 07, 2021, from <https://gdcoder.com/importance-of-cross-validation/>

35. Sanjay, M. (2018). Why And How To Cross Validate A Model?. [online] Medium.
36. Ansuini, C., Cavallo, A., Koul, A., Jacono, M., Yang, Y., & Becchio, C. (2015). Predicting object size from hand kinematics: a temporal perspective. *PLoS One*, 10(3), e0120432.
37. Kulić, D., Ott, C., Lee, D., Ishikawa, J., & Nakamura, Y. (2012). Incremental learning of full body motion primitives and their sequencing through human motion observation. *The International Journal of Robotics Research*, 31(3), 330-345.
38. Koppula, H. S., & Saxena, A. (2015). Anticipating human activities using object affordances for reactive robotic response. *IEEE transactions on pattern analysis and machine intelligence*, 38(1), 14-29.
39. Berret, B., Chiovetto, E., Nori, F., & Pozzo, T. (2011). Evidence for composite cost functions in arm movement planning: an inverse optimal control approach. *PLoS Comput Biol*, 7(10), e1002183.
40. Mainprice, J., Hayne, R., & Berenson, D. (2016). Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces. *IEEE Transactions on Robotics*, 32(4), 897-908.
41. Fragkiadaki, K., Levine, S., Felsen, P., & Malik, J. (2015). Recurrent network models for human dynamics. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4346-4354).
42. Jain, A., Zamir, A. R., Savarese, S., & Saxena, A. (2016). Structural-rnn: Deep learning on spatio-temporal graphs. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5308-5317).
43. Ghosh, P., Song, J., Aksan, E., & Hilliges, O. (2017, October). Learning human motion models for long-term predictions. In *2017 International Conference on 3D Vision (3DV)* (pp. 458-466). IEEE.
44. Martinez, J., Black, M. J., & Romero, J. (2017). On human motion prediction using recurrent neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2891-2900).

45. Pavllo, D., Grangier, D., & Auli, M. (2018). Quaternet: A quaternion-based recurrent model for human motion. arXiv preprint arXiv:1805.06485.
46. Pavllo, D., Feichtenhofer, C., Auli, M., & Grangier, D. (2019). Modeling human motion with quaternion-based neural networks. *International Journal of Computer Vision*, 1-18.
47. Todorov, E., & Li, W. (2005, June). A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proceedings of the 2005, American Control Conference, 2005.* (pp. 300-306). IEEE.
48. Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2019). OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence*, 43(1), 172-186.
49. Εικόνα 3 από <https://medium.com/@omaraymanomar/support-vector-machines-c4f4987845f9>
50. Εικόνα 4,5,6 από <https://www.linkedin.com/pulse/introdu%C3%A7%C3%A3o-ao-svm-como-classificador-rodriigo-araujo-?articleId=6643312353838419968>
51. Εικόνα 7 από <https://medium.com/swlh/random-forest-and-its-implementation-71824ced454f>
52. Εικόνα 8 από https://www.researchgate.net/figure/A-simple-example-of-visualizing-gradient-boosting_fig5_326379229
53. Εικόνα 9 από <https://gdcoder.com/importance-of-cross-validation/>