



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

**Εκτίμηση ύπαρξης άγχους κατά την εκτέλεση διαγνωστικού ελέγχου
Covid-19 με χρήση Μηχανικής Μάθησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γρηγόριος Γ. Μπίτσικας

Επιβλέπων Καθηγητής:

Ευάγγελος Β. Χριστοφόρου

Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2024



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

**Εκτίμηση ύπαρξης άγχους κατά την εκτέλεση διαγνωστικού ελέγχου
Covid-19 με χρήση Μηχανικής Μάθησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γρηγόριος Μπίτσικας

Επιβλέπων Καθηγητής:

Ευάγγελος Β. Χριστοφόρου

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 14^η Μαΐου 2024.

Ευάγγελος Χριστοφόρου
Καθηγητής
Ε.Μ.Π.

Εμμανουήλ Χουρδάκης
Επίκουρος
Καθηγητής
Ε.Μ.Π.

Γεώργιος Παναγόπουλος
Επίκουρος Καθηγητής
Ε.Μ.Π.

Αθήνα Ιούνιος 2024

.....
Γρηγόριος Γ. Μπίτσικας

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

Copyright © Γρηγόριος Μπίτσικας, 2024 .

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Η πανδημία Covid-19 αποτελεί ένα γεγονός το οποίο επηρέασε και συνεχίζει να επηρεάζει πολλές πτυχές τόσο της εγχώριας αλλά και της παγκόσμιας κοινωνίας. Η πρωτόγνωρη αλλά και προκλητική φύση του ιού SARS-CoV-2 και η αντίδραση των κοινωνιών και των κρατών προκάλεσε την αντίδραση πολλών ερευνητών, επιστημόνων και πολιτικών κυβερνήσεων να αναζητήσουν λύσεις, να αναλύσουν τις επιπτώσεις και να αντιδράσουν αποτελεσματικά στις προκλήσεις που προκύπτουν.

Στόχος της παρούσας διπλωματικής εργασίας είναι η εξαγωγή απόφασης για την ύπαρξη ή όχι άγχους σε ένα άτομο κατά την διάρκεια εκτέλεσης rapid test για την πιθανότητα νόσησης από Covid-19 με χρήση τεχνικών μηχανικής μάθησης. Το σύνολο των μετρήσεων με τις οποίες έγινε η εκπαίδευση του μοντέλου μηχανικής μάθησης και η εξαγωγή αποτελεσμάτων προέρχεται από έναν αισθητήρα συνδεδεμένο με μικροελεγκτή που μετρά την Γαλβανική Απόκριση Δέρματος. Με τον τρόπο αυτό μπορεί να καταγραφεί το άγχος που υπάρχει στο άτομο που εξετάζεται με rapid test. Κατόπιν χρησιμοποιήθηκε ο αλγόριθμος μηχανικής μάθησης Support Vector Machine (SVM) για την εξαγωγή αποτελεσμάτων.

Λέξεις-κλειδιά: Μηχανική Μάθηση, Γαλβανική απόκριση δέρματος, άγχος, Covid-19, Ταχύς διαγνωστικός έλεγχος, μετρήσεις, Support Vector Machine

Abstract

The Covid-19 pandemic is an event that has affected and continues to affect many aspects of both domestic and global society. The unprecedented and also challenging nature of the SARS-CoV-2 virus and the reaction of societies and states caused the reaction of many researchers, scientists and political governments to seek solutions, analyze the effects and react effectively to the challenges that arise.

The aim of this thesis is to derive a decision on the existence or not of anxiety in a person during the execution of a rapid test for the possibility of being infected by Covid-19 using machine learning techniques. The set of measurements used to train the machine learning model and extract results comes from a sensor connected to a microcontroller that measures the Galvanic Skin Response. In this way, the anxiety present in the person being examined with a rapid test can be recorded. Then Support Vector Machine (SVM) machine learning algorithm was used to extract results.

Keywords: *Machine Learning, Galvanic Skin Response, stress, Covid-19, Rapid test, Measurements, Support Vector Machine*

Πρόλογος – Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα Καθηγητή μου κ. Ευάγγελο Χριστοφόρου, αλλά και τον επιβλέποντα, ερευνητή Αντώνη Γεωργά για την ευκαιρία που μου προσέφεραν να εκπονήσω την διπλωματική μου εργασία πάνω σε ένα θέμα που παρουσιάζει μεγάλο ενδιαφέρον. Επιπλέον τους οφείλω ευχαριστίες για την στήριξη, την κατανόηση και συνεχή καθοδήγηση που μου προσέφεραν κατά την διάρκεια της εκπόνησης της διπλωματικής εργασίας και εκτιμώ το ενδιαφέρον και την εμπιστοσύνη που μου έδειξαν.

Θα ήθελα επίσης να ευχαριστήσω τον Καθηγητή μου κ. Νεκτάριο Κοζύρη για την εμπιστοσύνη και την εκτίμηση που έδειξε στο πρόσωπό μου καθώς και για τις συμβουλές και την καθοδήγηση που μου παρείχε.

Επιπλέον ευχαριστώ τους φίλους μου Κ.Μ. , Ν.Σ. , Γ.Τ. , Σ.Λ. , Γ.Ν. , Δ.Κ. , Η.Τ. , Α.Π. , Γ.Μ. , Β.Π. , με τους οποίους συμπορεύτηκα και μοιράστηκα την ακαδημαϊκή προπτυχιακή πορεία μου στην Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου, μαζί με αυτούς θα ήθελα να ευχαριστήσω επίσης τον φίλο μου Κ.Α. με ο οποίος επίσης με συνόδευσε στην πανεπιστημιακή μου σταδιοδρομία και τους υπόλοιπους φίλους και γνωστούς που με βοήθησαν κατά την διάρκεια των προπτυχιακών σπουδών μου.

Ευχαριστώ επίσης τα αδέρφια μου που μου συμπαραστάθηκαν κατά την διάρκεια των σπουδών μου και τους γονείς μου για την στήριξη που μου προσέφεραν. Ευχαριστώ και τα μέλη της ευρύτερης οικογένειάς μου που με στήριξαν στην πορεία μου.

Την παρούσα διπλωματική εργασία θα ήθελα να την αφιερώσω σε όλους εκείνους τους αξιοθαύμαστους φοιτητές και φοιτήτριες που με ανεξάντλητο ζήλο και υπομονή αψηφούν τις πολυάριθμες προκλήσεις που παρουσιάζονται στον δρόμο τους και παραμένουν σταθεροί μέχρι τέλους στην επιδίωξη των στόχων και των ονείρων τους.

Πίνακας περιεχομένων

Περίληψη	v
Abstract	vi
Πρόλογος – Ευχαριστίες	viii
Πίνακας περιεχομένων	x
Ευρετήριο Σχημάτων	xiii
Ευρετήριο Πινάκων.....	Error! Bookmark not defined.
1. Εισαγωγή.....	1
1.1. Σκοπός και δομή της εργασίας.....	1
1.2. Πανδημία Covid-19 και μέθοδοι για αντιμετώπιση και πρόληψη της νόσου.....	2
1.3. Το άγχος στον άνθρωπο και τρόποι για μέτρησής του	6
2. Γαλβανική Απόκριση Δέρματος (GSR)	11
2.1. Ορισμός Γαλβανικής Απόκρισης Δέρματος (GSR)	11
2.2. Εφαρμογές του GSR	12
2.3. Συλλογή δεδομένων Γαλβανικής Απόκρισης Δέρματος	13
2.4. Ανάλυση δεδομένων GSR.....	15
3. Μηχανική Μάθηση και Εφαρμογές της στην Βιοϊατρική.....	19
4. Εφαρμογή της Μηχανικής Μάθησης στα δεδομένα Γαλβανικής Απόκρισης Δέρματος	37
ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ	45
5. Προεπεξεργασία.....	45
6. Εκπαίδευση και Κατηγοριοποίηση δειγμάτων από το μοντέλο SVM	49
7. Αξιολόγηση Μοντέλων	63
8. Συμπεράσματα και Προτάσεις για Μελλοντική Έρευνα	71
ΒΙΒΛΙΟΓΡΑΦΙΑ	74

Ευρετήριο Σχημάτων

Σχήμα 1: Οι πολιτικές για διαγνωστικούς ελέγχους για covid-19 που ακολουθούνται από τις διάφορες χώρες του κόσμου. [3]	4
Σχήμα 2: Σημεία προσάρτησης των ηλεκτροδίων για καταγραφή σήματος GSR από το χέρι. Αριστερά για καταγραφή από τα δάχτυλα και δεξιά για την καταγραφή από την παλάμη. [9].....	14
Σχήμα 3: Σημεία προσάρτησης των ηλεκτροδίων για καταγραφή σήματος GSR από το πέλμα του ποδιού. [10]	15
Σχήμα 4: Latency (tLatency), Rise Time (tRise), Peak Amplitude, Recovery Time (tRecovery) σε σήμα GSR [13]	17
Σχήμα 5: Κατηγοριοποίηση δειγμάτων με την χρήση SVM φαίνονται τα υπερεπίπεδα (hyperplanes), τα διανύσματα υποστήριξης (supporting vectors. Με γκρι χρώμα έχει χρωματιστεί το περιθώριο (margin) μεταξύ των δύο κλάσεων. [31]	27
Σχήμα 6: Κατηγοριοποίηση δειγμάτων από SVM με χρήση πολυωνυμικού πυρήνα. [32]	31
Σχήμα 7: Κατηγοριοποίηση δεδομένων SVM με χρήση RBF πυρήνα. [33]	33
Σχήμα 8: Κατηγοριοποίηση δεδομένων με χρήση SVM και πυρήνα σιγμοειδούς (sigmoid kernel) [34]	35
Σχήμα 9: Δισδιάστατη απεικόνιση του μοντέλου PAD. Είναι ορατοί οι χαρακτηρισμοί των συναισθημάτων ανάλογα με την θέση τους σε σχέση με τους άξονες. [42].....	40
Σχήμα 10: Τρισδιάστατη απεικόνιση του μοντέλου PAD με την θέση όπου τοποθετούνται τα διάφορα συναισθήματα σε σχέση με τους άξονες. [43]	41
Σχήμα 11: Αρχική μορφή δεδομένων στο λογιστικό φύλλο.	46
Σχήμα 12: Τελική μορφή δεδομένων.	47
Σχήμα 13: Αποτέλεσμα της τεχνικής ομαδοποίησης σε τμήμα δείγματος εξεταζόμενων με διαγνωστικό έλεγχο για Covid-19. Με κίτρινο χρώμα φαίνονται τα	

άτομα τα οποία παρουσίασαν αυξημένα επίπεδα άγχους. Με μωβ αυτά που δεν εμφάνισαν αξιοσημείωτη μεταβολή στα δεδομένα GSR.	50
Σχήμα 14: Αποτελέσματα κατηγοριοποίησης με αλγόριθμο SVM. Με κύκλο αναπαρίστανται τα δεδομένα του συνόλου εκπαίδευσης ενώ με το σύμβολο «x» τα δεδομένα του συνόλου δοκιμής. Με μπλε χρώμα αναπαρίστανται τα δεδομένα που κατηγοριοποιήθηκαν ως μη αγχωμένα ενώ με κόκκινο οι εξεταζόμενοι που κατηγοριοποιήθηκαν ως αγχωμένοι.....	51
Σχήμα 15: Κατηγοριοποίηση SVM με χρήση απλού γραμμικού πυρήνα.	54
Σχήμα 16: Κατηγοριοποίηση με χρήση του μοντέλου LinearSVM της γλώσσας προγραμματισμού Python.....	54
Σχήμα 17: Κατηγοριοποίηση με χρήση πυρήνα Συνάρτησης Ακτινικής Βάσης (RBF kernel) και τιμή γάμμα (gamma) ίση με 2.0.....	55
Σχήμα 18: Κατηγοριοποίηση SVM με χρήση πολυωνυμικού πυρήνα βαθμού 5.56	
Σχήμα 19: Κατηγοριοποίηση SVM με χρήση πολυωνυμικού πυρήνα βαθμού 10.	56
Σχήμα 20: Κατηγοριοποίηση SVM με χρήση σιγμοειδούς πυρήνα (Sigmoid Kernel)	57
Σχήμα 21: Κατηγοριοποίηση NuSVM με χρήση πυρήνα συνάρτησης ακτινικής βάσης και τιμή παραμέτρου nu ίση με 0.02.	59
Σχήμα 22: Κατηγοριοποίηση NuSVM με χρήση πυρήνα συνάρτησης ακτινικής βάσης και τιμή παραμέτρου nu ίση με 0.2.	59
Σχήμα 23: Κατηγοριοποίηση NuSVM με χρήση πυρήνα σιγμοειδούς και τιμή nu ίση με 0.01	60
Σχήμα 24: Κατηγοριοποίηση NuSVM με χρήση πυρήνα σιγμοειδούς και τιμή nu ίση με 0.8	60
Σχήμα 25: Κατηγοριοποίηση με χρήση NuSVM σε συνδυασμό με πολυωνυμικό πυρήνα και τιμή παραμέτρου nu ίση με 0.03.	62
Σχήμα 26: : Κατηγοριοποίηση με χρήση NuSVM σε συνδυασμό με πολυωνυμικό πυρήνα και τιμή παραμέτρου nu ίση με 0.3.	62

1. Εισαγωγή

1.1. Σκοπός και δομή της εργασίας

Η συγκεκριμένη διπλωματική εργασία μελετά την χρήση μηχανικής μάθησης και ιδιαίτερα του αλγορίθμου Support Vector Machine (SVM) για να αποφασιστεί αν κάποιο άτομο βρίσκεται σε κατάσταση συναισθηματικής και ψυχικής ηρεμίας ή σε κατάσταση αυξημένου άγχους κατά την εκτέλεση διαγνωστικού ελέγχου Covid-19. Συχνά, σε άτομα που παίρνουν μέρος σε διαδικασίες μαζικών διαγνωστικών ελέγχων για εντοπισμό κρουσμάτων παρουσιάζονται αυξημένα επίπεδα άγχους, καθώς τα αποτελέσματα μπορεί να επηρεάσουν σημαντικούς τομείς της ζωής του εξεταζόμενου (επαγγελματική ζωή κλπ).

Σε αυτήν την διπλωματική εργασία θα μελετηθεί σε θεωρητικό και θα διερευνηθεί σε πειραματικό επίπεδο η χρήση της μηχανικής μάθησης προκειμένου να διερευνηθεί η ύπαρξη ή όχι άγχους σε υποκείμενο που υπόκειται σε διαγνωστικό έλεγχο για covid-19. Πιο συγκεκριμένα η απόφαση για την ύπαρξη ή όχι άγχους θα λαμβάνεται μέσω της χρήσης του αλγορίθμου SVM.

Στο θεωρητικό μέρος αρχικά περιγράφεται η γαλβανική απόκριση δέρματος (GSR) και το πώς αυτή συνδέεται με την ύπαρξη άγχους και άλλων έντονων συναισθημάτων που προκαλούν έκπληξη, φόβο και γενικότερα συναισθηματική διέγερση.

Κατόπιν παρουσιάζονται τρόποι που μπορούν οι αλγόριθμοι μηχανικής μάθησης να συνδυαστούν με μετρήσεις GSR προκειμένου να αποφασιστεί η συναισθηματική κατάσταση του εξεταζόμενου υποκειμένου. Ιδιαίτερη μελέτη γίνεται στον αλγόριθμο SVM ο οποίος και χρησιμοποιείται για την εξαγωγή των συμπερασμάτων.

Το πειραματικό μέρος της παρούσας διπλωματικής εργασίας παρουσιάζει την προεπεξεργασία των δεδομένων, τα διάφορα μοντέλα που χρησιμοποιήθηκαν (με διαφοροποίηση των παραμέτρων για δοκιμές), αξιολόγηση αποτελεσμάτων των πειραμάτων και τέλος παρουσιάζονται τα συμπεράσματα που εξήχθησαν και

γίνονται προτάσεις για μελλοντική έρευνα και επέκταση των εφαρμογών της μηχανικής μάθησης σε παρόμοιες περιπτώσεις.

1.2. Πανδημία Covid-19 και μέθοδοι για αντιμετώπιση και πρόληψη της νόσου.

1.2.1. Πανδημία Covid-19

Όταν στα τέλη Νοεμβρίου του 2019 αναφέρθηκε το πρώτο κρούσμα του νέου κορωνοϊού στην πόλη Wuhan της επαρχίας Hubei στην Κίνα τίποτα δεν προμήνυε τα γεγονότα που θα ακολουθούσαν τους επόμενους μήνες. Σε πολύ μικρό χρονικό διάστημα η νόσος ξεκίνησε να εντοπίζεται σε όλο και περισσότερες χώρες του κόσμου και να θέτει νέες προκλήσεις στους πολίτες και τις κυβερνήσεις.

Παρόλο που το γεγονός της πανδημίας φαίνεται ως πρωτόγνωρο, αναπάντεχο και μοναδικό εντούτοις με μια ματιά στην παγκόσμια ιστορία διαπιστώνουμε ότι αυτό απέχει πολύ από την αλήθεια. Τις τελευταίες δεκαετίες παρατηρούνται όλο και περισσότερες επιδημικές εξάρσεις παρά την ομολογουμένως εντυπωσιακή πρόοδο της ιατρικής επιστήμης. Για παράδειγμα την δεκαετία 1970 -1980 κάθε χρόνο έχαναν την ζωή τους 1.000.000 κάτοικοι της αφρικανικής ηπείρου από ελονοσία [1]. Νωρίτερα τον 20^ο αιώνα παρατηρείται η πανδημία της ισπανικής γρίπης (έναρξη το 1918). Γίνεται λοιπόν εύκολα αντιληπτό ότι η προσφάτως παρελθούσα πανδημία κάθε άλλο παρά ασυνήθιστο γεγονός μπορεί να χαρακτηριστεί.

Ο νέος κωρονοϊός (COVID-19) αποτελεί μια μολυσματική ασθένεια που προκαλείται από τον ιό SARS-CoV-2. Η επίσημη ανακήρυξη της νόσου ως πανδημίας από τον Παγκόσμιο Οργανισμό Υγείας (ΠΟΥ) έγινε στις 11 Μαρτίου του 2020. Μέχρι εκείνη την ημέρα ο νέος κωρονοϊός είχε εξαπλωθεί σε 172 από τις 195 χώρες του κόσμου [2].

Ένα από τα κύρια χαρακτηριστικά γνωρίσματα του ιού είναι η υψηλή μεταδοτικότητα. Λόγω αυτού σε πολύ μικρό χρονικό διάστημα από την ανακάλυψη των πρώτων κρουσμάτων ο ιός είχε διαδοθεί στον περισσότερο

κόσμο και θεωρούνταν επιδημία. Δεν μπορεί βέβαια να αγνοηθεί το γεγονός ότι η μετάδοση του ιού έγινε σε μια στιγμή όπου οι μετακινήσεις των ανθρώπων στον κόσμο έχουν φτάσει σε επίπεδα που δεν έχουν παρατηρηθεί ποτέ ξανά στην ανθρώπινη ιστορία.

1.2.2. Αντιμετώπιση και πρόληψη

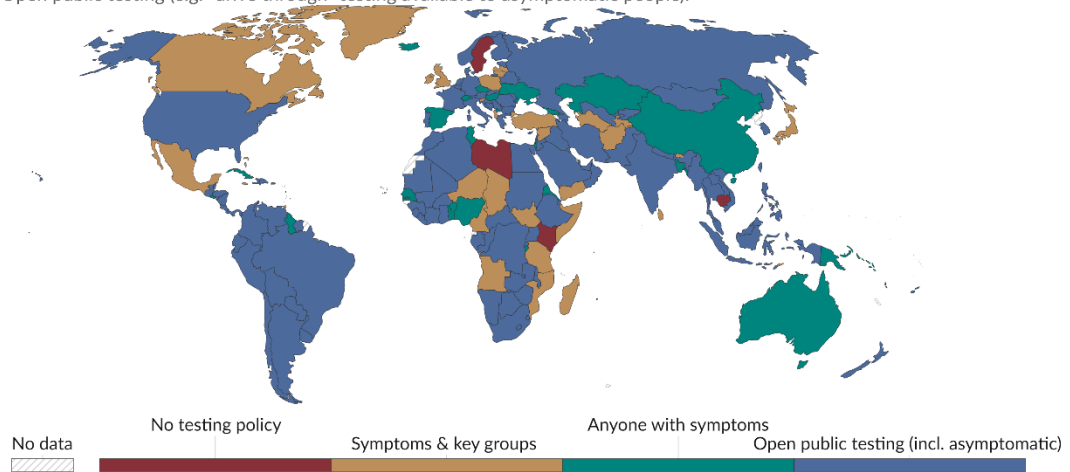
Η υψηλή μεταδοτικότητα του ιού οδήγησε πολύ γρήγορα στην εφαρμογή μέτρων κοινωνικής απομόνωσης των επιβεβαιωμένων κρουσμάτων η οποία έμεινε γνωστή ως «καραντίνα». Παράλληλα χρησιμοποιήθηκαν στρατηγικές εκτεταμένου διαγνωστικού ελέγχου στον γενικό πληθυσμό (rapid test ή PCR test) προκειμένου να ανακαλυφθούν ασυμπτωματικοί νοσούντες και να τεθούν σε περιορισμό (screening) ή διαγνωστικοί έλεγχοι σε άτομα που εμφανίζουν συμπτώματα.

Οι κυβερνήσεις των κρατών ανά τον κόσμο χρησιμοποίησαν διαφορετικές πολιτικές για την διενέργεια διαγνωστικών ελέγχων κατά την διάρκεια της πανδημίας. Στο παρακάτω σχήμα ([Σχήμα 1](#)) βλέπουμε τις τρεις βασικές πολιτικές που ακολουθήθηκαν όσον αφορά την διενέργεια διαγνωστικών ελέγχων μέχρι τον Δεκέμβριο του 2022.

Αν κοιτάξουμε το σχήμα θα διαπιστώσουμε ότι ακολουθούνται τρεις διαφορετικές πολιτικές.

- Παροχή tests σε άτομα με συμπτώματα και σε ομάδες πληθυσμού που βρίσκονται σε υψηλό κίνδυνο (ηλικιωμένοι, χρόνια νοσήματα και άλλα). Αυτές οι χώρες απεικονίζονται με καφέ χρώμα στον χάρτη.
- Παροχή tests μόνο σε άτομα που εμφανίζουν συμπτώματα (πράσινα κράτη στον χάρτη).
- Παροχή tests σε όλον τον γενικό πληθυσμό ανεξαρτήτως του αν εμφανίζουν συμπτώματα ή αν ανήκουν σε ομάδα υψηλού κινδύνου. (Μπλε περιοχές στον χάρτη).

- No testing policy.
- Only those who both (a) have symptoms and also (b) meet specific criteria (e.g. key workers, admitted to hospital, came into contact with a known case, returned from overseas).
- Testing of anyone showing COVID-19 symptoms.
- Open public testing (e.g. "drive through" testing available to asymptomatic people).



Σχήμα 1: Οι πολιτικές για διαγνωστικούς ελέγχους για covid-19 που ακολουθούνται από τις διάφορες χώρες του κόσμου. [3]

Παρατηρώντας τον χάρτη γίνεται αντιληπτό ότι το είδος της στρατηγικής ελέγχου δεν συνδέεται άμεσα ούτε με την γεωγραφική τοποθεσία (βλέπουμε ότι σε ίδιες ηπείρους ακολουθούνται διαφορετικές στρατηγικές από κάθε κράτος) ούτε με το επίπεδο οικονομικής ανάπτυξης (για παράδειγμα η Σουηδία που ανήκει στα κράτη του ανεπτυγμένου κόσμου δεν ακολουθεί καμία στρατηγική για διαγνωστικούς ελέγχους). Τα παραπάνω γεγονότα μας οδηγούν στο συμπέρασμα ότι η επιλογή της εκάστοτε στρατηγικής κατά κύριο λόγο επαφίεται στην κυβέρνηση του εκάστοτε κράτους.

Ο αντίκτυπος των στρατηγικών για τους διαγνωστικούς ελέγχους ποικίλλει ανάλογα με την προσέγγιση που υιοθετούν οι αντίστοιχες χώρες τους. Σε χώρες με ισχυρή υποδομή δοκιμών και προσβασιμότητα, οι πολίτες συχνά βιώνουν μια αίσθηση ασφάλειας, γνωρίζοντας ότι μπορούν εύκολα να υποβληθούν σε εξετάσεις εάν εμφανίσουν συμπτώματα ή έχουν έρθει σε επαφή με ένα μολυσμένο άτομο. Ο έγκαιρος έλεγχος επιτρέπει την έγκαιρη ανίχνευση και την έγκαιρη ιατρική παρέμβαση, μειώνοντας τον κίνδυνο σοβαρής ασθένειας και μετάδοσης εντός των κοινοτήτων. Επιπλέον, οι εκτεταμένες δοκιμές μπορεί να διευκολύνουν την εφαρμογή στοχευμένων μέτρων δημόσιας υγείας, όπως τοπικά lockdown ή πρωτόκολλα καραντίνας, τα οποία μπορούν να βοηθήσουν στον περιορισμό των

εστιών και στην προστασία των ευάλωτων πληθυσμών. Ωστόσο, σε περιοχές όπου οι δοκιμές είναι περιορισμένες ή όχι άμεσα διαθέσιμες, οι πολίτες μπορεί να αντιμετωπίσουν αβεβαιότητα, άγχος και εμπόδια στην πρόσβαση σε βασικές υπηρεσίες υγειονομικής περίθαλψης. Σε τέτοιες περιπτώσεις, η αδυναμία έγκαιρης εξέτασης μπορεί να οδηγήσει σε καθυστερημένη διάγνωση, αυξημένη μετάδοση και αυξημένη πίεση στα συστήματα υγειονομικής περίθαλψης. Συνολικά, η αποτελεσματικότητα και η προσβασιμότητα των διαγνωστικών ελέγχων για COVID-19 διαμορφώνουν σε βάθος τις εμπειρίες και τα συναισθήματα των πολιτών κατά τη διάρκεια της συνεχιζόμενης πανδημίας.

Επιπρόσθετα, ο αντίκτυπος των πολιτικών δοκιμών COVID-19 μπορεί επίσης να επεκταθεί πέρα από την υγειονομική περίθαλψη σε οικονομικούς και κοινωνικούς τομείς, επηρεάζοντας βαθιά την καθημερινή ζωή των πολιτών. Σε χώρες με ολοκληρωμένα προγράμματα δοκιμών, τα άτομα μπορεί να αισθάνονται μεγαλύτερη αυτοπεποίθηση συμμετέχοντας σε οικονομικές δραστηριότητες όπως εργασία, ταξίδια και κοινωνικές αλληλεπιδράσεις, γνωρίζοντας ότι οι εκτεταμένοι έλεγχοι συμβάλλουν σε ένα ασφαλέστερο περιβάλλον. Αυτό μπορεί να ενισχύσει την εμπιστοσύνη των καταναλωτών, να τονώσει την οικονομική ανάκαμψη και να ανακουφίσει ορισμένα από τα κοινωνικοοικονομικά βάρη που επιβάλλονται από την πανδημία. Επιπλέον, οι αποτελεσματικές πολιτικές ελέγχων συχνά συμβαδίζουν με διαφανείς προσπάθειες επικοινωνίας και δημόσιας εκπαίδευσης, ενισχύοντας την εμπιστοσύνη στις κυβερνητικές απαντήσεις και προάγοντας την τήρηση προληπτικών μέτρων. Ωστόσο, σε περιοχές όπου η ικανότητα ελέγχων είναι ανεπαρκής ή άνισα κατανομημένη, οι πολίτες μπορεί να βιώσουν αυξημένη οικονομική αβεβαιότητα, μειωμένη κινητικότητα και κοινωνικές διαταραχές λόγω παρατεταμένων περιορισμών. Οι ανισότητες στην πρόσβαση σε διαγνωστικούς ελέγχους μπορούν να επιδεινώσουν τις υπάρχουσες ανισότητες, επηρεάζοντας δυσανάλογα τις περιθωριοποιημένες κοινότητες και επιδεινώνοντας τις κοινωνικές εντάσεις. Έτσι, οι συνέπειες των πολιτικών διαγνωστικών ελέγχων COVID-19 υπερβαίνουν την υγειονομική περίθαλψη, διαμορφώνοντας βαθιά τον κοινωνικοοικονομικό ιστό και την ευημερία των πολιτών σε κάθε χώρα.

1.3. Το άγχος στον άνθρωπο και τρόποι για μέτρησής του

Παρακάτω θα παρουσιαστούν αναλυτικότερα ο επιστημονικός ορισμός του άγχους, ποιες μέθοδοι ακολουθούνται προκειμένου αυτό να μετρηθεί και να ποσοτικοποιηθεί και από ποια μέθοδο προέρχονται τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα διπλωματική και γιατί.

1.3.1. Επιστημονικός ορισμός άγχους

Το συναισθηματικό στρες στους ανθρώπους είναι ένα πολύπλευρο φαινόμενο που περιλαμβάνει μια σειρά ψυχολογικών, φυσιολογικών και συμπεριφορικών αντιδράσεων σε δύσκολες ή απειλητικές καταστάσεις. Επιστημονικά, το συναισθηματικό στρες περιλαμβάνει την ενεργοποίηση των συστημάτων απόκρισης στο στρες του σώματος, συμπεριλαμβανομένου του άξονα υποθαλάμου-υπόφυσης-επινεφριδίων (HPA) και του συμπαθητικού νευρικού συστήματος. Όταν τα άτομα αντιμετωπίζουν στρεσογόνους παράγοντες που αντιλαμβάνονται ως συναισθηματικά σημαντικούς ή απαιτητικούς, όπως συγκρούσεις σχέσεων, εργασιακές πιέσεις ή τραυματικά γεγονότα, αυτά τα συστήματα ενεργοποιούνται, οδηγώντας στην απελευθέρωση ορμονών του στρες όπως η κορτιζόλη και η αδρεναλίνη. Αυτές οι ορμονικές αλλαγές προετοιμάζουν το σώμα για μια ταχεία απόκριση, κινητοποιώντας τους ενεργειακούς πόρους και αυξάνοντας την εγρήγορση. Ωστόσο, το παρατεταμένο ή έντονο συναισθηματικό στρες μπορεί να έχει αρνητικές επιπτώσεις τόσο στην ψυχική όσο και στη σωματική υγεία, συμβάλλοντας σε καταστάσεις όπως αγχώδεις διαταραχές, κατάθλιψη, καρδιαγγειακές παθήσεις και εξασθενημένη ανοσοποιητική λειτουργία. Η κατανόηση του επιστημονικού ορισμού του συναισθηματικού στρες περιλαμβάνει την εξέταση της περίπλοκης αλληλεπίδρασής του με βιολογικούς, ψυχολογικούς και περιβαλλοντικούς παράγοντες, καθώς και τη διερεύνηση στρατηγικών για τον μετριασμό των αρνητικών του επιπτώσεων στη συνολική ευημερία του ατόμου.

Όσον αφορά τις μετρήσεις βιομετρικών στοιχείων που σχετίζονται ιδιαίτερα και αξιόπιστα με το επίπεδο άγχους το άρθρο “Detecting Stress During Real-World Driving Tasks Using Physiological Sensors” των Jennifer A. Healey και Rosalind W. Picard [4] που παρουσιάζει αποτελέσματα πειραμάτων για την μέτρηση άγχους σε οδηγούς συμπέρανε ότι για την ευρεία πλειοψηφία των ατόμων που μελετήθηκαν οι μετρήσεις καρδιακού ρυθμού και αγωγιμότητας του δέρματος μπορούν να συνδεθούν με τα επίπεδα άγχους που βιώνουν. Τα δεδομένα πάνω στα οποία εφαρμόστηκαν οι στρατηγικές της μηχανικής μάθησης στην παρούσα διπλωματική προέρχονται από τιμές μετρήσεων της αγωγιμότητας του δέρματος λόγω της αξιοπιστίας που παρέχει αυτό το βιομετρικό χαρακτηριστικό για τα επίπεδα άγχους .

1.3.2. Μέτρηση και ανίχνευση άγχους μέσω της παρατήρησης του ατόμου

Η εργασία των de Santos Sierra et al. [5] παρουσιάζει μια μέθοδο ανίχνευσης άγχους η οποία βασίζεται στην χρήση καρδιακών παλμών (HR) και GSR συνδυάζοντας Ασαφή λογική και περιπτωσιολογικό συλλογισμό.

Η μελέτη των Andren and Funk [6] ακολουθεί μια στρατηγική ανίχνευσης άγχους παρατηρώντας τον ρυθμό και τον τρόπο που ένα άτομο πληκτρολογεί χαρακτήρες σε ένα πληκτρολόγιο.

Η έρευνα των Dingens et al. [7] διερευνά την ανίχνευση του άγχους μέσω αναγνώρισης προσώπου (facial recognition).

Η χρήση της τεχνολογίας αναγνώρισης προσώπου για τον προσδιορισμό της ύπαρξης άγχους έχει κερδίσει έδαφος τόσο σε ερευνητικές όσο και σε πρακτικές εφαρμογές. Αυτή η προσέγγιση αξιοποιεί τις εξελίξεις στην όραση υπολογιστών και τους αλγόριθμους μηχανικής μάθησης για να αναλύει τις εκφράσεις του προσώπου και να ανιχνεύει λεπτές ενδείξεις που σχετίζονται με το στρες, όπως μυϊκή ένταση, μικροεκφράσεις και αλλαγές στο χρώμα ή την υφή του δέρματος. Καταγράφοντας και αναλύοντας δεδομένα προσώπου σε πραγματικό χρόνο, οι ερευνητές και οι επαγγελματίες μπορούν να συμπεράνουν τη συναισθηματική κατάσταση ενός ατόμου,

συμπεριλαμβανομένης της παρουσίας στρες, με αρκετά ικανοποιητικό βαθμό ακρίβειας.

Τα αυτοματοποιημένα συστήματα ανάλυσης προσώπου μπορούν να αξιολογήσουν ένα ευρύ φάσμα χαρακτηριστικών του προσώπου, όπως το αυλάκι των φρυδιών, το σφίξιμο των χειλιών και το στένωση των ματιών, τα οποία είναι ενδεικτικά των αντιδράσεων στο στρες. Επιπλέον, αυτά τα συστήματα μπορούν να παρέχουν συνεχή παρακολούθηση των επιπέδων στρες σε διάφορα πλαίσια, όπως κατά τη διάρκεια συνεντεύξεων, δημόσιας ομιλίας ή κλινικών αξιολογήσεων. Η τεχνολογία αναγνώρισης προσώπου προσφέρει πολλά υποσχόμενες ευκαιρίες για μη επεμβατική και αντικειμενική αξιολόγηση του άγχους. Πρέπει ωστόσο να τονιστεί ότι ενώ οι εκφράσεις του προσώπου οδηγούν στον προσδιορισμό ενός συναισθήματος αδυνατούν να μετρήσουν την ένταση αυτού.

1.3.3. Μέτρηση και ανίχνευση άγχους με χρήση ασαφούς λογικής

Το πλήθος των βιομετρικών δεδομένων που μπορούν να συλλεχθούν σήμερα είναι μεγάλο. Λόγω του ότι αυτά τα δεδομένα είναι αποτελέσματα διεργασιών του ανθρωπίνου σώματος και στις περισσότερες περιπτώσεις όχι απλά τυχαία σήματα γίνεται εύκολα αντιληπτό ότι μπορούν να αποδειχθούν πολύ χρήσιμα στο εγχείρημα αποκρυπτογράφησης των συναισθημάτων ενός ατόμου.

Οι de Santos Sierra et al. στην εργασία τους παρουσιάζουν μια ακόμα μέθοδο ανίχνευσης άγχους που συνδυάζει Ασαφή λογική , περιπτωσιολογικό συλλογισμό και σήματα του ανθρώπινου σώματος. Συχνά παρατηρείται σε τέτοιου είδους μελέτες η στρατηγική της συλλογής πολλών βιομετρικών σημάτων (για παράδειγμα καρδιακοί παλμοί, εκφράσεις προσώπου, γαλβανική απόκριση δέρματος και άλλων) προκειμένου να σχηματιστεί μια πληρέστερη εικόνα για την συναισθηματική κατάσταση του ατόμου.

Παρομοίως μια μελέτη των Jiang and Wang [8] κάνει χρήση της Ασαφούς Λογικής, ώστε να εξαχθούν συμπεράσματα σχετικά με το κατά πόσον το εξεταζόμενο άτομο βρίσκεται σε κατάσταση άγχους. Σε αυτήν την μελέτη

γίνεται χρήση του αλγορίθμου ασαφούς ομαδοποίησης c-means (FCM) για την συνεχή ανίχνευση του άγχους. Υπολογίζεται η ακριβής τιμή τάσης κάθε περιόδου και κατασκευάζεται τελικά μια συνεχής καμπύλη τάσεων. Η συλλογή των δεδομένων έγινε από οδηγούς και σχηματίζονται πολυδιάστατα χαρακτηριστικά διανύσματα μετά από την συλλογή και την εξαγωγή των κατάλληλων χαρακτηριστικών από τις μετρήσεις. Κατόπιν με την χρήση του αλγορίθμου FCM τα διανύσματα αυτά ομαδοποιούνται σε συμπλέγματα (clusters). Τα αποτελέσματα της παραπάνω μελέτης δείχνουν ότι με την ακολουθούμενη μέθοδο είναι δυνατόν να εντοπιστεί το άγχος στις διάφορες φάσεις της οδήγησης και μέσω της καμπύλης των τάσεων μπορεί να παρακολουθηθεί η μεταβολή του άγχους του οδηγού.

1.3.4. Άλλες τεχνικές για μέτρηση του άγχους

Υπάρχει μια πληθώρα βιομετρικών σημάτων τα οποία μπορούν να χρησιμοποιηθούν για να εξαχθούν συμπεράσματα σχετικά με το αν ένα άτομο είναι σε κατάσταση στρες ή όχι. Μερικά από αυτά είναι η μεταβλητότητα των καρδιακών παλμών (HR variability), η διαστολή κόρης (Pupil Dilation - PD) και η κίνηση του ματιού (Eyetracking - ET) που παρουσιάζουν άμεση συσχέτιση με το άγχος. Για παράδειγμα παρατηρείται αυξημένη ταχύτητα στην κίνηση του ματιού κατά την έκθεση σε αγχωτικές καταστάσεις.

Υπάρχουν και μελέτες που κάνουν χρήση αρκετών βιομετρικών σημάτων όπως η διαστολή της κόρης των ματιών (Pupil Dilation - PD), παρακολούθηση ματιών (Eyetracking - ET), και γαλβανικής απόκρισης δέρματος (GSR) με σκοπό να κάνουν μια ακριβέστερη αναγνώριση συναισθημάτων μέσω του συνδυασμού τους.

Ακόμα μια μέθοδος που μπορεί να ακολουθηθεί είναι το ηλεκτροεγκεφαλογράφημα (HEG - EEG). Το HEG συνίσταται στην μέτρηση και αναπαράστασης της ηλεκτρικής δραστηριότητας στο τριχωτό της κεφαλής ενός ατόμου. Μέσω της καταγραφής του HEG μπορεί να διευκρινισθούν οι περιοχές των εγκεφαλικών ημισφαιρίων που ενεργοποιούνται κατά την διάρκεια της εξέτασης. Επιπλέον είναι δυνατό να παρακολουθηθούν η

διέγερση και η απογοήτευση που προκαλεί κάποιο ερέθισμα. Τέλος πρέπει να αναφερθεί ότι το ΗΕΓ αποτελεί ένα από τα ταχύτερα βιομετρικά στοιχεία που χρησιμοποιούνται σήμερα.

Τέλος ακόμα ένας τρόπος συλλογής βιομετρικών δεδομένων είναι οι Ηλεκτρομυογραφικοί αισθητήρες. (ΗΜΓ - EMG). Μέσω αυτών των αισθητήρων γίνεται καταγραφή της ηλεκτρικής ενέργειας που δημιουργείται από κινήσεις του σώματος (πρόσωπο, δάχτυλα και άλλα). Με αυτόν τον τρόπο μπορούν να καταγραφούν οι μυϊκές αποκρίσεις σε πολλά ερεθίσματα.

1.3.5. Γαλβανική απόκριση δέρματος (GSR) για τον καθορισμό του επιπέδου άγχους κατά την διάρκεια διαγνωστικού ελέγχου για COVID-19.

Τα δεδομένα που χρησιμοποιήθηκαν για την συγκεκριμένη διπλωματική εργασία προέρχονται από αισθητήρα που μετρά την Γαλβανική Απόκριση δέρματος (GSR). Η μέτρηση του συγκεκριμένου βιομετρικού σήματος χαρακτηρίζεται από υψηλή ακρίβεια όσον αφορά τον καθορισμό του επιπέδου άγχους του εξεταζόμενου. Επιπλέον η διαδικασία καταγραφής είναι εύκολη, γρήγορη και δεν προκαλεί δυσφορία στους εξεταζόμενους. Παρακάτω, στο [Κεφάλαιο 2](#) θα παρουσιαστεί εκτενέστερα η Γαλβανική Απόκριση Δέρματος και το πώς αυτή χρησιμοποιείται για να μετρηθούν τα επίπεδα άγχους σε ένα άτομο.

2. Γαλβανική Απόκριση Δέρματος (GSR)

2.1. Ορισμός Γαλβανικής Απόκρισης Δέρματος (GSR)

Η γαλβανική απόκριση δέρματος (GSR), επίσης γνωστή ως αγωγιμότητα του δέρματος ή ηλεκτροδερμική δραστηριότητα, είναι ένα φυσιολογικό φαινόμενο που αφορά την ηλεκτρική αγωγιμότητα του δέρματος ως απόκριση σε συναισθηματική διέγερση ή ψυχολογικά ερεθίσματα. Το δέρμα περιέχει ιδρωτοποιούς αδένες που νευρώνονται από το συμπαθητικό νευρικό σύστημα, το οποίο ενεργοποιείται κατά τη συναισθηματική διέγερση ή το στρες. Όταν ένα άτομο βιώνει ένα συναισθηματικό ερέθισμα, όπως ενθουσιασμό, φόβο ή άγχος, η δραστηριότητα των ιδρωτοποιών αδένων αυξάνεται, οδηγώντας σε αύξηση της αγωγιμότητας του δέρματος. Αυτή η αύξηση της αγωγιμότητας ανιχνεύεται από ηλεκτρόδια που τοποθετούνται στην επιφάνεια του δέρματος, συνήθως στα δάχτυλα ή στην παλάμη, και καταγράφεται ως διακυμάνσεις στην ηλεκτρική αντίσταση.

Το GSR χρησιμοποιείται συχνά ως έμμεσος τρόπος μέτρησης της συναισθηματικής διέγερσης ή των επιπέδων στρες καθώς παρέχει φυσιολογικά δεδομένα σε πραγματικό χρόνο χωρίς να απαιτεί λεκτικές ή συνειδητές απαντήσεις από τον συμμετέχοντα. Επιπλέον, το GSR χρησιμοποιείται σε κλινικά περιβάλλοντα για να βοηθηθούν τα άτομα να μάθουν να ρυθμίζουν τις αποκρίσεις στην συναισθηματική πίεση και το άγχος. Συμπερασματικά, η γαλβανική απόκριση του δέρματος αποτελεί ένα πολύτιμο εργαλείο για την κατανόηση των ψυχοφυσιολογικών πτυχών του ανθρώπινου συναισθήματος και των αντιδράσεων στο στρες.

Από τα παραπάνω γίνεται εύκολα αντιληπτό ότι ένα άτομο δεν είναι ποτέ σε θέση να ελέγχει συνειδητά το επίπεδο αγωγιμότητας του δέρματός του. Οπότε η Γαλβανική Απόκριση Δέρματος αποτελεί σημαντικό και αξιόπιστο βιομετρικό δεδομένο που μπορεί να χρησιμοποιηθεί για να διαπιστωθεί αν το εξεταζόμενο άτομο βρίσκεται σε κατάσταση συναισθηματικής διέγερσης.

2.2. Εφαρμογές του GSR

Η γαλβανική απόκριση δέρματος έχει χρησιμοποιηθεί σε πολλές μελέτες που αφορούν την συναισθηματική φόρτιση, την μνήμη, την συγκέντρωση, την προσοχή και την απόφαση. Η Γαλβανική Απόκριση Δέρματος μπορεί επίσης να προβλέψει φυσιολογική ή μη φυσιολογική συμπεριφορά σε υγιή ή ασθενή άτομα.

Ενδεικτικά κάποιες εφαρμογές μετρήσεων και συλλογής δεδομένων που αφορούν το GSR στον σημερινό κόσμο είναι:

- **Θεραπεία με βιοανάδραση:** Τα συστήματα βιοανάδρασης που βασίζονται σε GSR χρησιμοποιούνται σε προγράμματα διαχείρισης του άγχους και παρεμβάσεις ψυχικής υγείας για να βοηθήσουν τα άτομα να μάθουν τεχνικές χαλάρωσης και να βελτιώσουν τις δεξιότητες συναισθηματικής ρύθμισης.
- **Ηλεκτρονικά παιχνίδια και εικονική πραγματικότητα:** Η τεχνολογία GSR ενισχύει τις εμπειρίες παιχνιδιών και εικονικής πραγματικότητας προσαρμόζοντας δυναμικά το περιεχόμενο με βάση τις συναισθηματικές αντιδράσεις των χρηστών, δημιουργώντας έτσι πιο καθηλωτικές και ελκυστικές αλληλεπιδράσεις.
- **Έρευνα Αγοράς και Διαφήμιση:** Το GSR χρησιμοποιείται στην έρευνα αγοράς για τη μέτρηση των αντιδράσεων των καταναλωτών σε προϊόντα, διαφημίσεις και περιεχόμενο πολυμέσων, επιτρέποντας στις εταιρείες να μετρήσουν τα επίπεδα συναισθηματικής δέσμησης και διέγερσης.
- **Εφαρμογές υγειονομικής περίθαλψης:** Σε περιβάλλοντα υγειονομικής περίθαλψης, το GSR χρησιμοποιείται σε ψυχολογικές αξιολογήσεις, θεραπείες αγχώδους διαταραχής και θεραπεία βιοανάδρασης για καταστάσεις όπως το μετατραυματικό σύνδρομο (PTSD) και οι φοβίες, βοηθώντας στη διάγνωση και τη διαχείριση διαταραχών ψυχικής υγείας.

2.3. Συλλογή δεδομένων Γαλβανικής Απόκρισης Δέρματος

Προκειμένου να συλλέξουμε δεδομένα GSR χρειαζόμαστε έναν αισθητήρα που θα μας οδηγήσει στο να μετρήσουμε την αγωγιμότητα του δέρματος. Αυτό συμβαίνει διότι το GSR είναι άμεσα συνδεδεμένο με την δραστηριότητα των ιδρωτοποιών αδένων. Συγκεκριμένα όσο μεγαλώνει η διέγερση του ιδρωτοποιού αδένου αυξάνεται η ποσότητα ιδρώτα στην επιφάνεια του δέρματος οπότε μειώνεται η τιμή αντίστασης του δέρματος άρα αυξάνεται και η αντίστοιχη αγωγιμότητα με αποτέλεσμα την αύξηση της Γαλβανικής Απόκρισης Δέρματος.

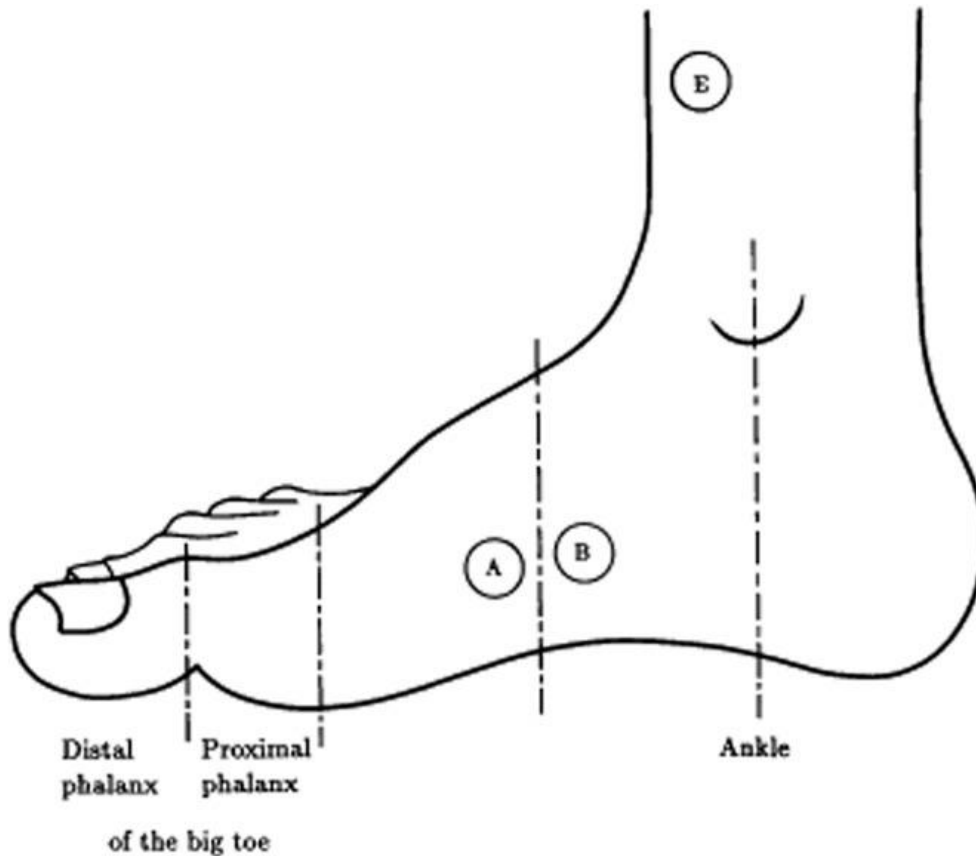
Για την μέτρηση του σήματος GSR χρησιμοποιούμε ως επί το πλείστον αισθητήρες. Η θέση αυτών των αισθητήρων είναι τα δάχτυλα των χεριών ή των ποδιών. Η αλήθεια είναι ότι ιδρωτοποιοί αδένες υπάρχουν σε ολόκληρο το ανθρώπινο σώμα όμως στις παλάμες των χεριών και τα δάχτυλα των ποδιών και των χεριών είναι πιο εύκολο να ανιχνευθούν με μεγαλύτερη ακρίβεια και αξιοπιστία αυξομειώσεις στην εφίδρωση οπότε και χρησιμοποιούνται για την καταγραφή του σήματος GSR.

- **Δάχτυλα χεριού:** Η τεχνική αυτή χρησιμοποιείται όταν δεν αναμένονται ιδιαίτερες κινήσεις από τα χέρια του εξεταζόμενου. Στις περισσότερες περιπτώσεις ο αισθητήρας προσαρτάται στο μη κυρίαρχο χέρι. Με αυτόν τον τρόπο ο εξεταζόμενος μπορεί να κάνει χρήση του άλλου χεριού του προκειμένου να φέρει εις πέρας μια ενέργεια. Τα ηλεκτρόδια του αισθητήρα τοποθετούνται συνήθως στον μέσο και τον δείκτη του χεριού. ([Σχήμα 2](#))
- **Παλάμη:** Τα ηλεκτρόδια εξοπλίζονται με ειδικές αυτοκόλλητες επιφάνειες προκειμένου να προσαρτηθούν σταθερά στην παλάμη του εξεταζόμενου. Αυτή η τακτική ακολουθείται σε περιπτώσεις όπου προβλέπεται αυξημένη χρήση των χεριών του ατόμου κατά την διαδικασία της μέτρησης. Στο [Σχήμα 2](#) τα σημεία για την προσάρτηση των ηλεκτροδίων στην παλάμη φαίνονται στην αριστερή φωτογραφία.
- **Πέλμα:** Η καταγραφή από το πέλμα προτιμάται σε περιπτώσεις όπου προβλέπεται αρκετά αυξημένη κίνηση των χεριών του εξεταζόμενου.

Τα ηλεκτρόδια τοποθετούνται στις θέσεις Α, Β όπως φαίνεται στο [Σχήμα 3](#).



Σχήμα 2: Σημεία προσάρτησης των ηλεκτροδίων για καταγραφή σήματος GSR από το χέρι. Αριστερά για καταγραφή από τα δάχτυλα και δεξιά για την καταγραφή από την παλάμη. [9]



Σχήμα 3: Σημεία προσάρτησης των ηλεκτροδίων για καταγραφή σήματος GSR από το πέλμα του ποδιού. [10]

2.4. Ανάλυση δεδομένων GSR

Για να γίνει η σωστή συλλογή GSR μετρήσεων εφαρμόζεται μια τάση στο δέρμα και στη συνέχεια μετρείται η αγωγιμότητα του δέρματος. [11]

Το συλλεγόμενο σήμα GSR αποτελείται από δύο κύρια στοιχεία: Skin Conductance Level (SCL) και Skin Conductance Response (SCR).

- **Skin Conductance Level (SCL):** Το SCL μεταβάλλεται ανάλογα με την ξηρότητα, την ενυδάτωση και την αυτόνομη ρύθμιση του δέρματος. Η μεταβολή του είναι πολύ αργή και αλλάζει ελαφρώς σε χρονική κλίμακα από δεκάδες δευτερόλεπτα έως λεπτά. Από τα παραπάνω γίνεται αντιληπτό ότι το SCR δεν είναι κατατοπιστικός δείκτης για το επίπεδο της συναισθηματικής διέγερσης του εξεταζόμενου.

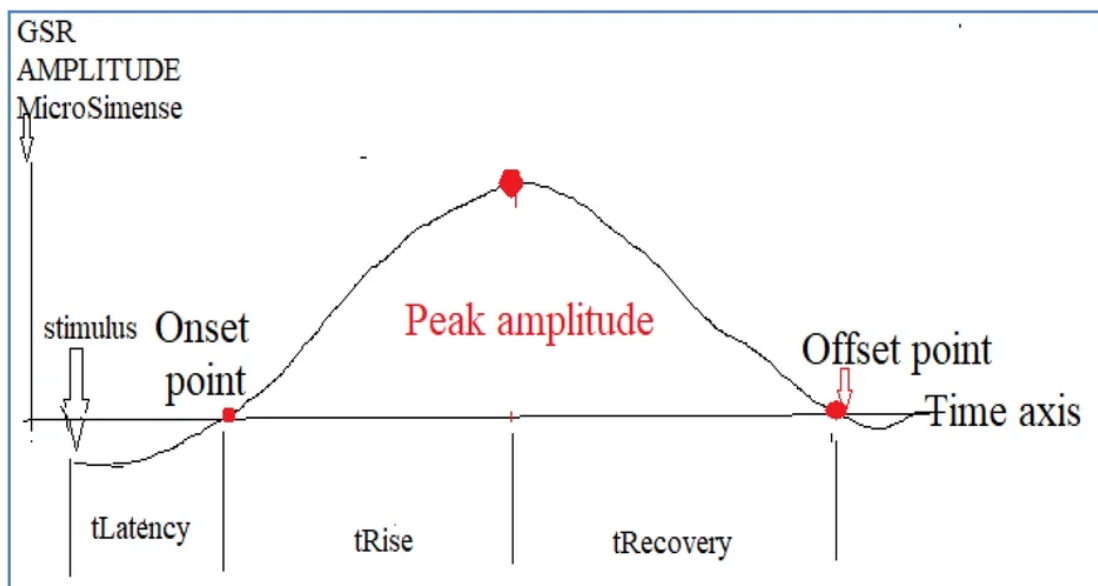
- **Skin Conductance Response (SCR):** Το SCR μεταβάλλεται ανάλογα με τον αριθμό των ιδρωτοποιών αδένων που ενεργοποιούνται. Ως εκ τούτου αυξάνεται όταν αυξάνεται και η συναισθηματική διέγερση του εξεταζόμενου. Οπότε αποτελεί έναν αξιόπιστο δείκτη του επιπέδου συναισθηματικής διέγερσης. [12]

Το SCR ονομάζεται επίσης και κορυφή “peak” της ηλεκτροδερμικής δραστηριότητας, επίσης αναφέρεται και ως “GSR peak” και παρατηρείται ως μια απότομη αύξηση της τιμής του σήματος. Αν το SCR παρατηρηθεί ως μια απόκριση σε ένα ερέθισμα τυπικά σε χρονικό διάστημα από 1 έως 5 δευτερόλεπτα τότε χαρακτηρίζεται ως “event-related SCR” (ER - SCRs). Αν εμφανίζεται χωρίς να υπάρχει κάποια παρατηρήσιμη αιτία τότε ονομάζεται “Non Specific SCR” (NS-SCR). Τα NS-SCRs συμβαίνουν αυθόρμητα στο σώμα 1-3 φορές ανά λεπτό και δεν σχετίζονται με κάποιο ερέθισμα.

Κατά την ανάλυση δεδομένων GSR το ενδιαφέρον εστιάζεται στις SCR που σχετίζονται με γεγονότα καθώς συσχετίζονται άμεσα με μετρήσεις διέγερσης των εξεταζόμενων. Τα μεμονωμένα SCR μπορούν να χαρακτηριστούν από τα παρακάτω χαρακτηριστικά όπως φαίνεται και στο [Σχήμα 4](#).

- **Αφάνεια (Latency):** Χρονική διάρκεια από την έναρξη του ερεθίσματος έως την έναρξη της φασικής έκρηξης. Συνήθως οι ER – SCRs εμφανίζονται 1-5 δευτερόλεπτα μετά την έναρξη του ερεθίσματος. Η έναρξη τοποθετείται την χρονική στιγμή όπου η καμπύλη GSR υπερβαίνει ένα όριο ελαχίστου πλάτους (0.01 μ S ή 0.05 μ S αντίστοιχα). Οποιοσδήποτε μεταβολές στο GSR παρατηρούνται πριν από αυτήν την περίοδο χαρακτηρίζονται ως NS – SCR και δεν σχετίζονται με το πείραμα.
- **Πλάτος Κορυφής (Peak Amplitude):** Η διαφορά μεταξύ πλάτους έναρξης και πλάτους κορυφής.
- **Χρόνος Ανύψωσης (Rise Time):** Χρονική διάρκεια από την έναρξη του ερεθίσματος έως την κορυφή.
- **Χρόνος Αποθεραπείας (Recovery Time):** Χρονική διάρκεια από την κορυφή έως την 100% ανάκτηση των αρχικών επιπέδων GSR. Η ανάκτηση

είναι συνήθως αρκετά πιο ομαλή από την ανύψωση οπότε και οι χρόνοι αποθεραπείας κατά κανόνα είναι πιο μεγάλοι από τους χρόνους ανύψωσης.



Σχήμα 4: Latency ($t_{Latency}$), Rise Time (t_{Rise}), Peak Amplitude, Recovery Time ($t_{Recovery}$) σε σήμα GSR [13]

Στην παρούσα διπλωματική για την κατηγοριοποίηση των εξεταζόμενων σε αγχωμένους ή μη χρησιμοποιήθηκε το πλάτος κορυφής και η διαφορά πλάτους κορυφής πλην την μέση τιμή της μέτρησης GSR για κάθε εξεταζόμενο. Η συγκεκριμένη επιλογή έγινε διότι μετά από πειραματισμούς με μια πληθώρα χαρακτηριστικών της μέτρησης GSR προέκυψε το συμπέρασμα ότι ο αλγόριθμος SVM που έχει χρησιμοποιηθεί στα πλαίσια της παρούσης διπλωματικής επιτύγχανε καλύτερα αποτελέσματα όσον αφορά την εκπαίδευσή του και την κατηγοριοποίηση νέων δειγμάτων ως αγχωμένα ή μη.

3. Μηχανική Μάθηση και Εφαρμογές της στην Βιοϊατρική

3.1. Γενική παρουσίαση της μηχανικής μάθησης

3.1.1. Επιστημονικός ορισμός μηχανικής μάθησης

Η μηχανική μάθηση (ML) είναι ένα πεδίο μελέτης στην τεχνητή νοημοσύνη που ασχολείται με την ανάπτυξη και τη μελέτη στατιστικών αλγορίθμων που μπορούν να μάθουν από δεδομένα και να γενικεύσουν σε μη ορατά δεδομένα και έτσι να εκτελέσουν εργασίες χωρίς ρητές οδηγίες. Πρόσφατα, τα τεχνητά νευρωνικά δίκτυα μπόρεσαν να ξεπεράσουν πολλές προηγούμενες προσεγγίσεις σε απόδοση. [14], [15], [16]

Ο βασικός σκοπός της μηχανικής μάθησης είναι να επιτρέψει στους υπολογιστές να γίνουν καλύτεροι σε μια συγκεκριμένη εργασία μέσω της έκθεσής τους με την πάροδο του χρόνου σε όλο και περισσότερα δεδομένα.

3.1.2. Εφαρμογές Μηχανικής Μάθησης

Η μηχανική μάθηση έχει βρει εφαρμογές σε διάφορους τομείς, φέρνοντας επανάσταση στους κλάδους και οδηγώντας την καινοτομία. Στην υγειονομική περίθαλψη, διευκολύνει τη διάγνωση ασθενειών, τα εξατομικευμένα σχέδια θεραπείας και την ανακάλυψη φαρμάκων μέσω της ανάλυσης ιατρικών εικόνων, γονιδιωματικών δεδομένων και αρχείων ασθενών. Στα χρηματοοικονομικά, οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται για τον εντοπισμό απάτης, την αξιολόγηση κινδύνου και τις αλγοριθμικές συναλλαγές, επιτρέποντας ακριβέστερες προβλέψεις και ενημερωμένη λήψη αποφάσεων.

Επιπλέον, στο μάρκετινγκ και το ηλεκτρονικό εμπόριο, τα συστήματα προτάσεων που υποστηρίζονται από αλγόριθμους μηχανικής μάθησης ενισχύουν την εμπειρία των χρηστών παρέχοντας εξατομικευμένες προτάσεις προϊόντων με βάση τη συμπεριφορά και τις προτιμήσεις του παρελθόντος. Επιπλέον, στα αυτόνομα οχήματα και τη ρομποτική, η μηχανική μάθηση

διαδραματίζει κρίσιμο ρόλο στη διευκόλυνση της πλοήγησης, της αναγνώρισης αντικειμένων και της λήψης αποφάσεων σε πραγματικό χρόνο, συμβάλλοντας στην πρόοδο των αυτοοδηγούμενων αυτοκινήτων. Αυτές οι εφαρμογές απεικονίζουν την ολοένα αυξανόμενη επίδραση της μηχανικής μάθησης στην αναμόρφωση των βιομηχανιών και στη βελτίωση της αποδοτικότητας, της παραγωγικότητας και της ποιότητας ζωής.

Η ευελιξία της μηχανικής μάθησης εκτείνεται πέρα από τις παραδοσιακές βιομηχανίες, με τις δυνατότητές της να αντιμετωπίζει πολύπλοκες κοινωνικές προκλήσεις και να αναδιαμορφώνει το μέλλον. Στην περιβαλλοντική επιστήμη, οι τεχνικές μηχανικής μάθησης εφαρμόζονται για την ανάλυση συνόλων δεδομένων μεγάλης κλίμακας, όπως δορυφορικές εικόνες και κλιματικά μοντέλα, για την παρακολούθηση της αποψίλωσης των δασών, την πρόβλεψη φυσικών καταστροφών και τη διαχείριση των πόρων με πιο βιώσιμο τρόπο [17], [18]. Στην εκπαίδευση, οι πλατφόρμες προσαρμοστικής μάθησης αξιοποιούν αλγόριθμους μηχανικής μάθησης για την εξατομίκευση των μαθησιακών εμπειριών, τον εντοπισμό των μεμονωμένων αναγκών των μαθητών και τη βελτιστοποίηση της παράδοσης εκπαιδευτικού περιεχομένου [19]. Επιπλέον, στην ασφάλεια του κυβερνοχώρου, οι αλγόριθμοι μηχανικής μάθησης εντοπίζουν και μετριάζουν τις απειλές αναλύοντας μοτίβα κυκλοφορίας δικτύου, εντοπίζοντας ανωμαλίες και ενισχύοντας την άμυνα του συστήματος σε πραγματικό χρόνο [20]. Καθώς η τεχνολογία συνεχίζει να εξελίσσεται, η μηχανική μάθηση αναμένεται να οδηγήσει σε καινοτομίες σε τομείς όπως η υγειονομική περίθαλψη, ο μετριασμός της κλιματικής αλλαγής και η κοινωνική δικαιοσύνη, ανοίγοντας το δρόμο για ένα πιο δίκαιο μέλλον.

3.1.3. Βιοϊατρικές εφαρμογές μηχανικής μάθησης

Η εξάπλωση της μηχανικής μάθησης ασφαλώς δεν μπορεί να μην επηρεάσει τον τομέα της βιοϊατρικής. Η μηχανική μάθηση έχει προκαλέσει επανάσταση στον τομέα της βιοϊατρικής προσφέροντας καινοτόμες λύσεις σε μια ποικιλία από εφαρμογές.

Για παράδειγμα η μηχανική μάθηση μπορεί να συνδυαστεί με την χρήση βιοαισθητήρων που μετρούν την συγκέντρωση Covid-19 στο σάλιο. Με την

βοήθεια των αλγορίθμων της μηχανικής μάθησης είναι δυνατόν κάθε μέτρηση να κατηγοριοποιηθεί ως ακριβής ή λανθασμένη. Η κατηγοριοποίηση είναι δυνατή αφού οι χρησιμοποιούμενοι αλγόριθμοι έχουν εκπαιδευτεί με ένα σύνολο δεδομένων ικανοποιητικού μεγέθους όπου κάθε μέτρηση έχει χαρακτηριστεί ως ακριβής ή λανθασμένη. [21]

Επιπρόσθετα η μηχανική μάθηση βρίσκει εφαρμογή στην αναγνώριση, ανάλυση και διάγνωση ασθενειών. Οι αλγόριθμοι της μηχανικής μάθησης είναι πολύ αποδοτικοί στην ανάλυση ιατρικών δεδομένων και συντελούν με αυτόν τον τρόπο στην έγκαιρη ανίχνευση και ταχεία διάγνωση. Για παράδειγμα μπορούν να χρησιμοποιηθούν για να προβλέψουν ασθένειες εκπαιδευόμενοι με συμπτώματα, αποτελέσματα εξετάσεων και δεδομένα ιατρικών απεικονίσεων [22]. Η παραπάνω δυνατότητα βρίσκει εφαρμογή στην έγκαιρη ανίχνευση καρκίνου από μαστογραφίες.

Επιπλέον η μηχανική μάθηση (ML) έχει επηρεάσει σημαντικά τον τομέα της ανακάλυψης και της κατασκευής φαρμάκων. Έχει μεταμορφώσει τις παραδοσιακές διαδικασίες, οδηγώντας σε καινοτόμες λύσεις. Ένας βασικός τομέας είναι ο *de novo* σχεδιασμός φαρμάκων, όπου οι αλγόριθμοι ML προβλέπουν μοριακές δομές και ιδιότητες για τη δημιουργία νέων φαρμάκων. Επίσης, βοηθά στην ανακάλυψη νέων φαρμάκων, αναλύοντας χημικές βιβλιοθήκες για τον εντοπισμό πιθανών ενώσεων για συγκεκριμένους βιολογικούς στόχους. Η ανάλυση ποσοτικής σχέσης δομής-δραστηριότητας (QSAR) συσχετίζει τα μοριακά χαρακτηριστικά με τη βιολογική δραστηριότητα, βοηθώντας στην πρόβλεψη της αποτελεσματικότητας μιας ένωσης. Η μηχανική μάθηση διαδραματίζει επίσης κρίσιμο ρόλο στην ψηφιακή παθολογία και στη βελτιστοποίηση της παραγωγής. Παρά τις προκλήσεις, η συνεχιζόμενη έρευνα συνεχίζει να αξιοποιεί τις δυνατότητες της μηχανικής μάθησης, οδηγώντας τις προόδους στη φαρμακευτική επιστήμη και τη φροντίδα των ασθενών. [23], [24], [25]

Ακόμα, η μηχανική μάθηση μπορεί να αποδειχθεί πολύ χρήσιμη στον τομέα της πραγματοποίησης κλινικών δοκιμών. Κατά την προετοιμασία για μια κλινική δοκιμή με χρήση μηχανικής μάθησης μπορούν να σχεδιαστούν πιο

αποδοτικά πρωτόκολλα, τακτικές και στρατηγικές για την σχεδιαζόμενη έρευνα. Επιπλέον μπορεί να βοηθήσει στην πιο αποδοτική επιλογή συμμετεχόντων υπολογίζοντας πολλούς διαφορετικούς παράγοντες ενδιαφέροντος. Ακόμα, η μηχανική μάθηση ανιχνεύει τυχόν ανωμαλίες στα δεδομένα και αυτοματοποιεί την διαδικασία συλλογής στοιχείων ενδιαφέροντος στην έρευνα. Μπορεί επίσης να φανεί χρήσιμη στο να συμπληρωθούν ελλείψεις σε δεδομένα βελτιώνοντας έτσι την ακεραιότητα των δεδομένων. Βέβαια, ενόσω η επιρροή της μηχανικής μάθησης στον τομέα των κλινικών δοκιμών ολοένα και αυξάνεται, είναι απαραίτητες και οι αντίστοιχες μελέτες [26], [27] προκειμένου να διασαφηνιστεί η πραγματική επίδραση της μηχανικής μάθησης στον συγκεκριμένο τομέα και να ξεπεραστούν οι νέες προκλήσεις που ανακύπτουν. Σε κάθε περίπτωση, η εξέλιξη της μηχανικής μάθησης μπορεί να ωφελήσει ασθενείς και ερευνητές συντελώντας στην γενικότερη πρόοδο της ιατρικής επιστήμης.

Άλλη μια εφαρμογή της μηχανικής μάθησης συνίσταται στην πρόβλεψη ξεσπάσματος επιδημιών. Κατά την διάρκεια της παγκόσμιας πανδημίας COVID-19, οι αρχές και οι ερευνητές στράφηκαν σε διάφορα μοντέλα για να προβλέψουν τις τάσεις της επιδημίας και να αναβαθμίσουν τα μέτρα ελέγχου. Ενώ τα τυπικά επιδημιολογικά και στατιστικά μοντέλα έχουν τραβήξει την προσοχή, η ακρίβειά τους για μακροπρόθεσμες προβλέψεις παραμένει περιορισμένη λόγω της αβεβαιότητας και της έλλειψης δεδομένων. Ως εναλλακτική λύση, έχουν προκύψει μοντέλα μηχανικής μάθησης (ML). Συγκεκριμένα, το πολυεπίπεδο perceptron (multi-layered perceptron - MLP) και το προσαρμοστικό σύστημα ασαφούς συμπερασμάτων που βασίζεται σε δίκτυο (adaptive network-based fuzzy inference system - ANFIS) έχουν δείξει πολλά υποσχόμενα αποτελέσματα στην πρόβλεψη εστιών COVID-19. Δεδομένης της περίπλοκης φύσης του ιού και της ποικίλης συμπεριφοράς του μεταξύ των διαφόρων ομάδων, η μηχανική μάθηση προσφέρει ένα αποτελεσματικό εργαλείο για τη μοντελοποίηση και την κατανόηση της δυναμικής της επιδημίας [28]. Αυτά τα μοντέλα μπορούν να συμπληρώσουν παραδοσιακές προσεγγίσεις όπως τα μοντέλα ευπαθών-μολυσμένων-

ανάκτησης (susceptible–infected–recovered - SIR) και ευπαθών-εκτεθειμένων-λοιμωδών-αφαιρούμενων (susceptible-exposed-infectious-removed - SEIR), ενισχύοντας την ικανότητά μας να προβλέπουμε και να διαχειριζόμαστε επιδημίες. [29]

Περαιτέρω εφαρμογή της μηχανικής μάθησης στον τομέα της υγείας είναι ορατή στον τομέα των «Εξυπνων Ηλεκτρονικών Μητρώων Υγείας» (Smart Electronic Health Records - EHR). Τα Ηλεκτρονικά Μητρώα Υγείας (EHR) περιέχουν πληθώρα πληροφοριών σχετικά με το ιστορικό υγείας των ασθενών, συμπεριλαμβανομένων τόσο δομημένων όσο και μη δομημένων δεδομένων. Ενώ οι παραδοσιακές προσεγγίσεις έχουν επικεντρωθεί στην εξαγωγή πολύτιμων γνώσεων από δομημένα δεδομένα, όπως κωδικούς ασθενειών, εργαστηριακά αποτελέσματα και θεραπείες, το να βασίζεται κανείς σε αποκλειστικά δομημένα δεδομένα μπορεί να μην συλλαμβάνει πλήρως τις πληροφορίες των ασθενών και μπορεί περιστασιακά να περιλαμβάνει λανθασμένα αρχεία.

Τα τελευταία χρόνια, οι τεχνικές μηχανικής μάθησης και βαθιάς μάθησης (Deep Learning - DL) έχουν αποκτήσει εξέχουσα θέση. Οι ερευνητές έχουν εξερευνήσει χρησιμοποιώντας πολυτροπικά δεδομένα (multimodal data), τα οποία συνδυάζουν τόσο δομημένες όσο και μη δομημένες πληροφορίες από EHR, ως είσοδο για συμβατικά μοντέλα μηχανικής μάθησης ή DL. Αυτά τα μοντέλα στοχεύουν στην περάτωση συγκεκριμένων διαδικασιών που σχετίζονται με EHR. Ένα παράδειγμα μιας τέτοιας διαδικασίας είναι η υποβοήθηση των ιατρών προκειμένου να αποφασίσουν την τελική διάγνωση [30]

3.2. Αλγόριθμοι και τεχνικές Μηχανικής Μάθησης

Παρακάτω θα αναλυθούν μια σειρά από τεχνικές και αλγορίθμους που χρησιμοποιούνται στην μηχανική μάθηση. Αρχικά θα παρουσιαστούν οι δυο βασικές στρατηγικές μάθησης των μοντέλων που χρησιμοποιούνται στην μηχανική μάθηση σήμερα δηλαδή η εποπτευόμενη μάθηση (supervised learning) και η μη εποπτευόμενη μάθηση (unsupervised learning). Κατόπιν θα παρουσιαστεί ο αλγόριθμος Support Vector Machine που χρησιμοποιήθηκε στην παρούσα διπλωματική για την κατηγοριοποίηση των δειγμάτων.

3.2.1. Εποπτευόμενη Μάθηση (Supervised Learning)

Η εποπτευόμενη μάθηση περιλαμβάνει την εκπαίδευση μιας μηχανής χρησιμοποιώντας δεδομένα στα οποία έχει ανατεθεί κάποιος χαρακτηρισμός. Σε αυτήν την προσέγγιση, κάθε στοιχείο στο σύνολο δεδομένων εκπαίδευσης συνδυάζεται με μια αντίστοιχη σωστή απάντηση ή ταξινόμηση. Ο αλγόριθμος μαθαίνει να προβλέπει την έξοδο για τα νέα δεδομένα εισόδου με βάση τα παραδείγματα που έχει δει.

Οι συνήθεις περιπτώσεις χρήσης περιλαμβάνουν ταξινόμηση (ανάθεση δεδομένων εισόδου σε προκαθορισμένες κατηγορίες) και παλινδρόμηση (πρόβλεψη συνεχών τιμών). Για παράδειγμα, όταν εκπαιδεύεται ένα μηχάνημα να αναγνωρίζει φρούτα, παρέχονται ετικέτες μήλων, μπανανών και πορτοκαλιών και το μηχάνημα μαθαίνει τη σχέση μεταξύ των χαρακτηριστικών εισαγωγής (όπως σχήμα, χρώμα και υφή) και των αντίστοιχων ετικετών φρούτων. Κατόπιν όταν του παρέχεται ένα καινούριο στοιχείο από το σύνολο των δεδομένων (φρούτα στο παρόν παράδειγμα), μπορεί με βάση την γνώση που απέκτησε κατά την εκπαίδευση να κατατάξει σε κάποια κατηγορία το νέο στοιχείο.

3.2.2. Μη εποπτευόμενη μάθηση (Unsupervised Learning)

Σε αντίθεση με προηγουμένως, η μη εποπτευόμενη μάθηση (Unsupervised Learning) λειτουργεί με δεδομένα χωρίς χαρακτηρισμό. Σε αυτήν την

περίπτωση, τα δεδομένα εισόδου στερούνται αντιστοίχιση με χαρακτηρισμό εξόδου. Ο αλγόριθμος διερευνά τα δεδομένα για να ανακαλύψει μοτίβα, δομές ή σχέσεις χωρίς καμία προκαθορισμένη καθοδήγηση. Οι περιπτώσεις χρήσης για μάθηση χωρίς επίβλεψη περιλαμβάνουν ομαδοποίηση (clustering - ομαδοποίηση παρόμοιων σημείων δεδομένων μαζί), μείωση διαστάσεων και ανίχνευση ανωμαλιών. Για παράδειγμα, κατά την ανάλυση μιας συλλογής εικόνων χωρίς ετικέτα, το μηχάνημα εντοπίζει κοινά σημεία μεταξύ τους χωρίς να γνωρίζει συγκεκριμένες κατηγορίες (π.χ. μήλα, μπανάνες).

Συνοπτικά, τόσο η εποπτευόμενη όσο και η μη εποπτευόμενη μάθηση διαδραματίζουν κρίσιμους ρόλους στην προώθηση των εφαρμογών μηχανικής μάθησης, καθεμία με τα μοναδικά πλεονεκτήματα και τις εφαρμογές της.

Στην παρούσα διπλωματική χρησιμοποιήθηκαν και οι δύο τακτικές. Στην περίπτωση μη επιβλεπόμενης μάθησης, τα δεδομένα των μετρήσεων GSR δόθηκαν στον αλγόριθμο ομαδοποίησης και κατόπιν αυτός τα κατέταξε σε αγχωμένους και μη. Κατόπιν, το σύνολο των ομαδοποιημένων δεδομένων χρησιμοποιήθηκε για εκπαίδευση του αλγορίθμου SVM.

Όσον αφορά την επιβλεπόμενη μάθηση, χρησιμοποιήθηκαν δείγματα που είχαν ήδη χαρακτηριστεί ως αγχωμένα ή μη και ακολουθήθηκε με αυτά η διαδικασία εκπαίδευσης του αλγορίθμου.

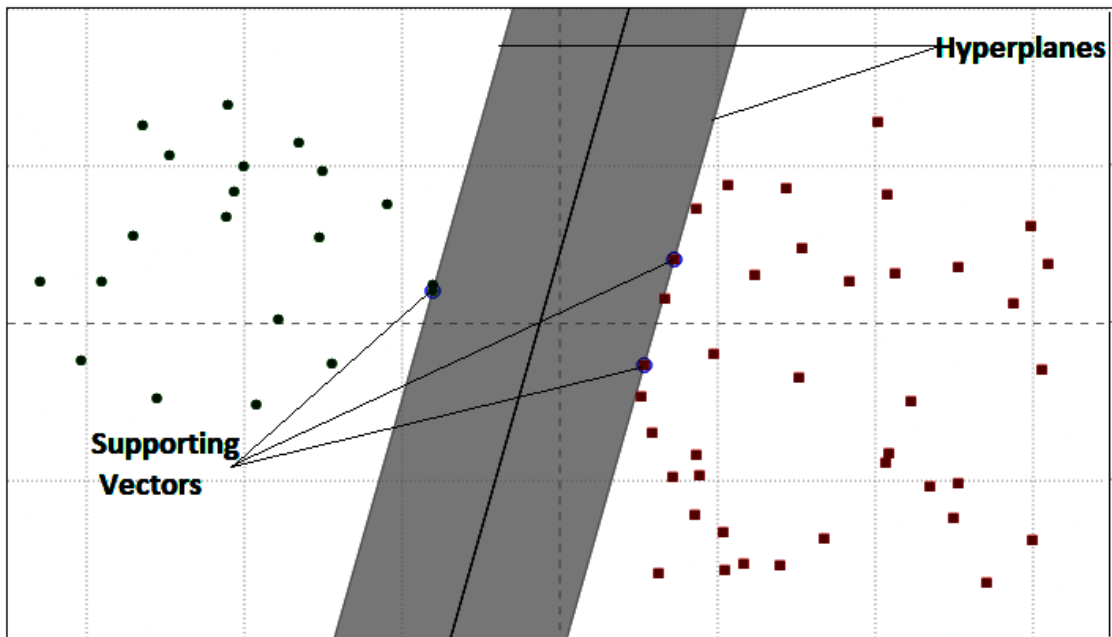
3.3. Αλγόριθμος Support Vector Machine (SVM)

Στη συνέχεια θα παρουσιαστεί ο αλγόριθμος Support Vector Machine που χρησιμοποιήθηκε στα πλαίσια της παρούσας διπλωματικής για την κατηγοριοποίηση των δειγμάτων.

3.3.1. SVM γενικά

Ο αλγόριθμος Support Vector Machine (SVM) είναι μια ισχυρή και ευέλικτη εποπτευόμενη μέθοδος μάθησης που χρησιμοποιείται για περιπτώσεις ταξινόμησης και παλινδρόμησης. Στον πυρήνα του, το SVM στοχεύει να βρει το βέλτιστο υπερεπίπεδο (hyperplane – επίπεδο σε χώρους άνω των δύο διαστάσεων) που διαχωρίζει καλύτερα τα σημεία δεδομένων που ανήκουν σε διαφορετικές κλάσεις σε ένα χώρο χαρακτηριστικών πολλών διαστάσεων. Αυτό το υπερεπίπεδο καθορίζεται μεγιστοποιώντας το περιθώριο (margin), το οποίο είναι η απόσταση μεταξύ του επιπέδου και των πλησιέστερων σημείων δεδομένων (data points), γνωστά ως διανύσματα υποστήριξης (support vectors). Μεγιστοποιώντας το περιθώριο, το SVM όχι μόνο επιτυγχάνει καλή απόδοση γενίκευσης, αλλά επίσης εμφανίζει ανθεκτικότητα σε ακραίες τιμές.

Σε περιπτώσεις όπου τα δεδομένα δεν είναι γραμμικά διαχωρίσιμα, το SVM χρησιμοποιεί το χαρακτηριστικό του πυρήνα (kernel) για να αντιστοιχίσει τα δεδομένα εισόδου σε έναν χώρο υψηλότερης διάστασης, όπου είναι δυνατός ο γραμμικός διαχωρισμός. Αυτό επιτρέπει στο SVM να χειρίζεται αποτελεσματικά τα μη γραμμικά όρια αποφάσεων. Το SVM έχει κερδίσει δημοτικότητα λόγω της ικανότητάς του να χειρίζεται δεδομένα πολλών διαστάσεων, της αποτελεσματικότητάς του τόσο σε γραμμικές όσο και σε μη γραμμικές εργασίες ταξινόμησης και στη σταθερή θεωρητική βάση του. Βρίσκει εφαρμογές σε διάφορους τομείς, όπως ταξινόμηση εικόνων, ταξινόμηση κειμένου, βιοπληροφορική και χρηματοοικονομικές προβλέψεις, αποτελώντας με αυτόν τον τρόπο ένα ευέλικτο εργαλείο στον τομέα της μηχανικής μάθησης.



Σχήμα 5: Κατηγοριοποίηση δειγμάτων με την χρήση SVM φαίνονται τα υπερεπίπεδα (hyperplanes), τα διανύσματα υποστήριξης (supporting vectors). Με γκρι χρώμα έχει χρωματιστεί το περιθώριο (margin) μεταξύ των δύο κλάσεων. [31]

3.3.2. Πυρήνας (kernel) στο SVM

Στο SVM, ένας πυρήνας είναι ένα κρίσιμο στοιχείο που επιτρέπει στον αλγόριθμο να λειτουργεί αποτελεσματικά σε χώρους υψηλών διαστάσεων και να χειρίζεται μη γραμμικές σχέσεις στα δεδομένα.

Ουσιαστικά, μια συνάρτηση πυρήνα υπολογίζει το εσωτερικό γινόμενο μεταξύ ζευγών σημείων δεδομένων στον χώρο εισόδου, αντιστοιχίζοντας αποτελεσματικά τον αρχικό χώρο χαρακτηριστικών σε έναν χώρο υψηλότερης διάστασης όπου ο γραμμικός διαχωρισμός μπορεί να είναι πιο εφικτός (για παράδειγμα μεταφέροντας τα δεδομένα από δύο σε τρεις διαστάσεις). Αυτός ο μετασχηματισμός επιτρέπει στο SVM να κατασκευάσει σύνθετα όρια αποφάσεων που μπορούν να ταξινομήσουν με ακρίβεια σημεία δεδομένων που δεν είναι γραμμικά διαχωρισμένα στον αρχικό χώρο χαρακτηριστικών. Οι κοινές συναρτήσεις πυρήνα που χρησιμοποιούνται στο SVM περιλαμβάνουν τον γραμμικό πυρήνα (linear kernel), τον πολυωνυμικό πυρήνα (polynomial kernel), τον πυρήνα της συνάρτησης ακτινικής βάσης (RBF) και τον πυρήνα

του σιγμοειδούς (sigmoid kernel). Κάθε πυρήνας έχει τα δικά του χαρακτηριστικά και είναι κατάλληλος για διαφορετικούς τύπους δεδομένων.

Η επιλογή της συνάρτησης πυρήνα συχνά εξαρτάται από τα συγκεκριμένα χαρακτηριστικά του συνόλου δεδομένων και την επιθυμητή πολυπλοκότητα του ορίου απόφασης. Συμπερασματικά, ο πυρήνας παίζει κρίσιμο ρόλο στην ενίσχυση της ευελιξίας και της απόδοσης του SVM, επιτρέποντάς του να μοντελοποιεί αποτελεσματικά μη γραμμικές σχέσεις στα δεδομένα.

Παρακάτω θα παρουσιαστούν συνοπτικά τα είδη πυρήνων που χρησιμοποιούνται στο SVM.

- **Γραμμικός πυρήνας (linear kernel)**

Ο γραμμικός πυρήνας στο SVM είναι μια από τις πιο απλές συναρτήσεις πυρήνα που χρησιμοποιούνται για εργασίες ταξινόμησης. Λειτουργεί με υπολογισμό του εσωτερικού γινομένου μεταξύ ζευγών σημείων δεδομένων στον αρχικό χώρο χαρακτηριστικών, χωρίς μετασχηματισμό.

Ουσιαστικά, ο γραμμικός πυρήνας ορίζει ένα όριο απόφασης που είναι ένα υπερεπίπεδο στον αρχικό χώρο χαρακτηριστικών, με στόχο να διαχωρίσει σημεία δεδομένων διαφορετικών κλάσεων με ένα γραμμικό όριο απόφασης. Αυτό το όριο απόφασης χαρακτηρίζεται από το ότι είναι μια ευθεία γραμμή στον δισδιάστατο χώρο, ένα επίπεδο στον τρισδιάστατο χώρο και ένα υπερεπίπεδο σε χώρους υψηλότερων διαστάσεων.

Παρά την απλότητά του, ο γραμμικός πυρήνας είναι αποτελεσματικός για γραμμικά διαχωρίσιμα σύνολα δεδομένων όπου οι κλάσεις μπορούν να διαχωριστούν καθαρά από μια ευθεία γραμμή ή υπερεπίπεδο. Είναι υπολογιστικά αποδοτικό και εύκολο στην ερμηνεία, καθιστώντας το κατάλληλο για μεγάλης κλίμακας σύνολα δεδομένων και σενάρια όπου η ευκολία κατανόησης του μοντέλου είναι σημαντική.

Ωστόσο, ο γραμμικός πυρήνας μπορεί να μην έχει καλή απόδοση για σύνολα δεδομένων με πολύπλοκες μη γραμμικές σχέσεις, καθώς δεν μπορεί να συλλάβει τα περίπλοκα μοτίβα που υπάρχουν στα δεδομένα. Συνολικά, ο γραμμικός πυρήνας χρησιμεύει ως θεμελιώδες δομικό στοιχείο στο SVM, παρέχοντας μια βασική προσέγγιση για εργασίες ταξινόμησης και χρησιμεύει ως σημείο αναφοράς για τη σύγκριση της απόδοσης πιο περίπλοκων λειτουργιών του πυρήνα.

Παραπάνω στο [Σχήμα 5](#) φαίνεται το πώς διαχωρίζει τα δεδομένα ένας γραμμικός πυρήνας

- **Πολυωνυμικός πυρήνας (polynomial kernel)**

Ο πολυωνυμικός πυρήνας στο SVM είναι μια ευέλικτη συνάρτηση πυρήνα που χρησιμοποιείται για το χειρισμό μη γραμμικών σχέσεων στα δεδομένα αντιστοιχίζοντάς τα σε έναν χώρο υψηλότερης διάστασης. Λειτουργεί υπολογίζοντας το εσωτερικό γινόμενο ζευγών σημείων δεδομένων στον αρχικό χώρο χαρακτηριστικών και στη συνέχεια εφαρμόζοντας μια πολυωνυμική συνάρτηση σε αυτό το εσωτερικό γινόμενο. Ο πολυωνυμικός πυρήνας εισάγει τη μη γραμμικότητα αυξάνοντας το αποτέλεσμα του εσωτερικού γινομένου σε μια καθορισμένη ισχύ, που καθορίζεται από την παράμετρο βαθμού. Αυτός ο μετασχηματισμός επιτρέπει αποτελεσματικά στο SVM να συλλαμβάνει σύνθετα όρια αποφάσεων που είναι πολυωνυμικής φύσης, όπως καμπύλες ή επιφάνειες υψηλότερων βαθμών.

Η παράμετρος βαθμού (degree) ελέγχει την ευελιξία του ορίου απόφασης: οι υψηλότεροι βαθμοί επιτρέπουν πιο σύνθετα όρια απόφασης, αλλά μπορεί επίσης να οδηγήσουν σε υπερπροσαρμογή (overfitting) εάν δεν συντονιστούν προσεκτικά.

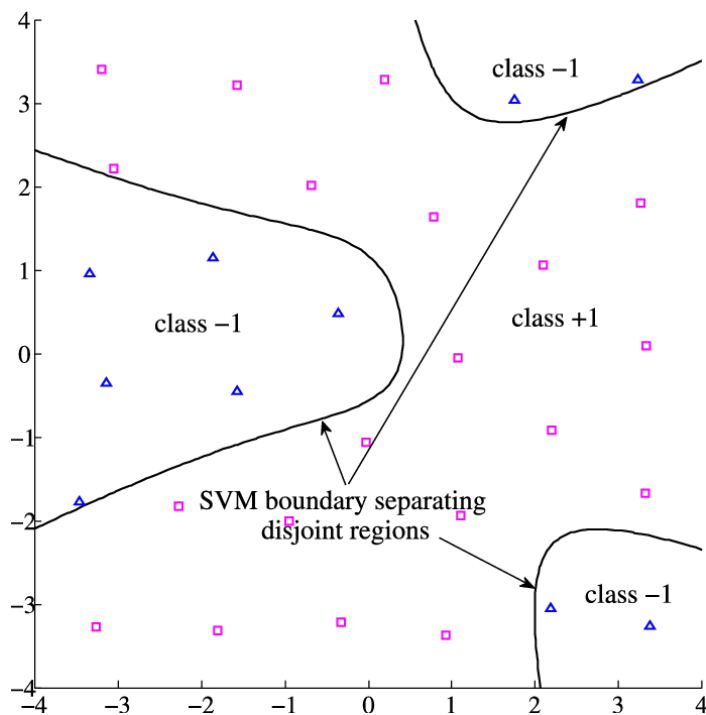
Ο πολυωνυμικός πυρήνας είναι ιδιαίτερα χρήσιμος για σύνολα δεδομένων όπου οι κλάσεις εμφανίζουν πολυωνυμικές σχέσεις ή όπου ο γραμμικός διαχωρισμός δεν επαρκεί για την ακριβή ταξινόμηση των

δεδομένων. Ωστόσο, η επιλογή μιας κατάλληλης παραμέτρου βαθμού μπορεί να είναι δύσκολη και μπορεί να απαιτεί πειραματισμό ή διασταυρούμενη επικύρωση για τη βελτιστοποίηση της απόδοσης. Από τα παραπάνω γίνεται εύκολα αντιληπτό ότι ο πολυωνυμικός πυρήνας προσφέρει μια ευέλικτη και ισχυρή προσέγγιση για το χειρισμό μη γραμμικών σχέσεων στο SVM, καθιστώντας τον κατάλληλο για ένα ευρύ φάσμα εργασιών ταξινόμησης.

Ο κίνδυνος υπερπροσαρμογής (overfitting) που αναφέρθηκε παραπάνω συμβαίνει όταν ο αλγόριθμος μαθαίνει πολύ καλά τα δεδομένα εκπαίδευσης, καταγράφοντας θόρυβο και άσχετα μοτίβα εκτός από τις υποκείμενες σχέσεις. Αυτό οδηγεί σε ένα μοντέλο που αποδίδει εξαιρετικά καλά στα δεδομένα εκπαίδευσης, αλλά αποτυγχάνει να γενικεύσει σε μη ορατά δεδομένα.

Στο SVM, η υπερπροσαρμογή μπορεί να συμβεί όταν η πολυπλοκότητα του μοντέλου είναι πολύ υψηλή σε σχέση με τον όγκο των διαθέσιμων δεδομένων εκπαίδευσης ή όταν η παράμετρος κανονικοποίησης (παράμετρος C) είναι πολύ χαμηλή. Τα σύνθετα όρια αποφάσεων με πολλά διανύσματα υποστήριξης μπορεί να οδηγήσουν σε υπερπροσαρμογή, καθώς το μοντέλο προσπαθεί να προσαρμόσει τον θόρυβο στα δεδομένα και όχι τα υποκείμενα μοτίβα. Για να μετριαστεί η υπερπροσαρμογή στο SVM, μπορούν να χρησιμοποιηθούν τεχνικές όπως η διασταυρούμενη επικύρωση (cross validation) και η επιλογή χαρακτηριστικών (feature selection). Η διασταυρούμενη επικύρωση βοηθά στην αξιολόγηση της απόδοσης του μοντέλου σε μη ορατά δεδομένα και η επιλογή χαρακτηριστικών εστιάζει στον εντοπισμό και τη διατήρηση μόνο των πιο σχετικών χαρακτηριστικών για ταξινόμηση.

Η εξισορρόπηση της πολυπλοκότητας του μοντέλου με την απόδοση γενίκευσης είναι ζωτικής σημασίας για την αποφυγή υπερπροσαρμογής και για τη διασφάλιση της καλής απόδοσης του μοντέλου SVM σε νέα, άορατα δεδομένα.



Σχήμα 6: Κατηγοριοποίηση δειγμάτων από SVM με χρήση πολυωνυμικού πυρήνα. [32]

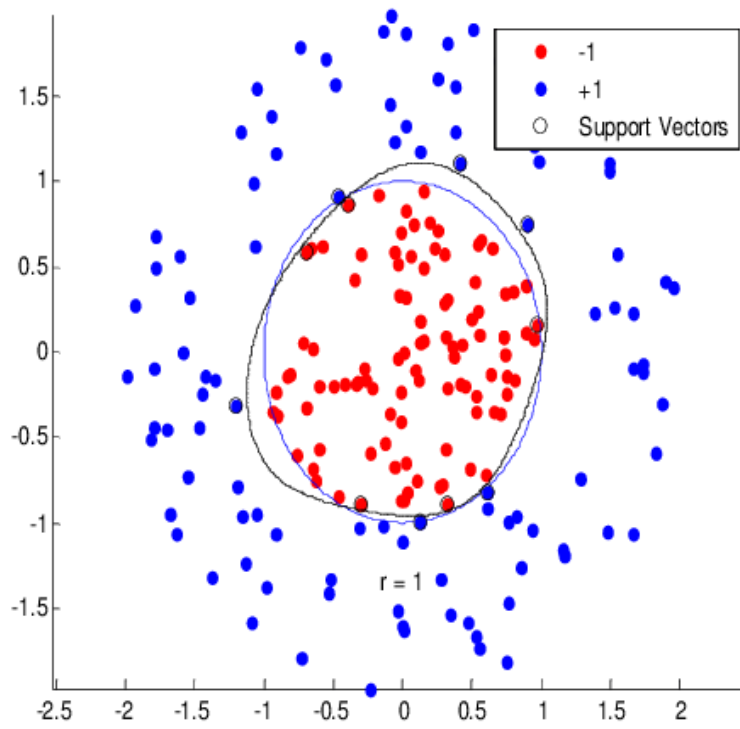
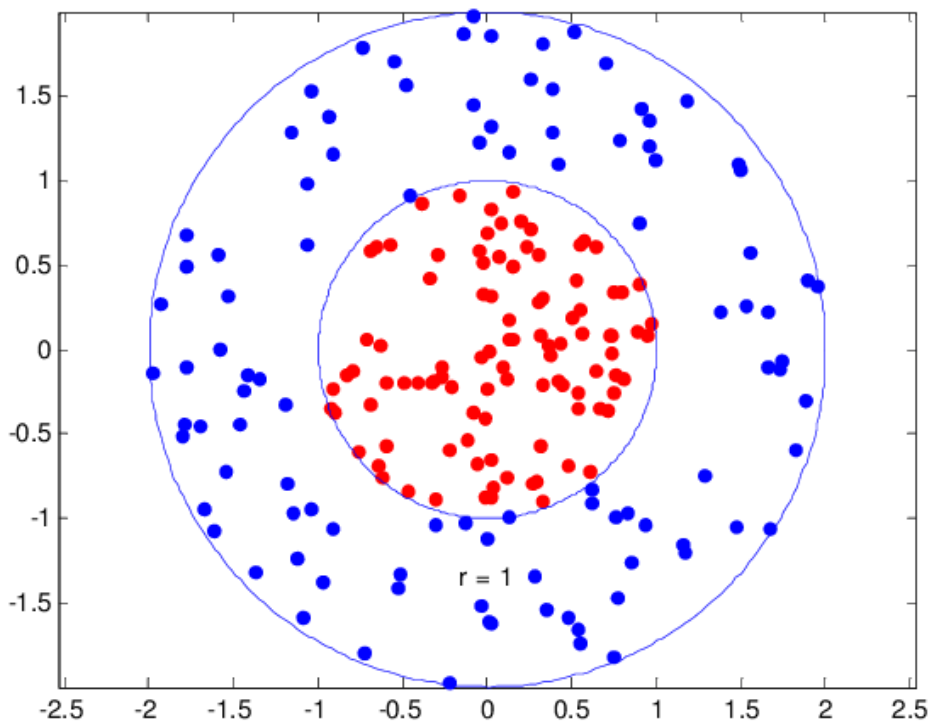
- **Πυρήνας της συνάρτησης ακτινικής βάσης (RBF)**

Ο πυρήνας συνάρτησης ακτινικής βάσης (Radial Basis Function, RBF) στο SVM είναι μια ισχυρή και ευρέως χρησιμοποιούμενη συνάρτηση πυρήνα για το χειρισμό μη γραμμικών σχέσεων στα δεδομένα. Λειτουργεί μετατρέποντας τα δεδομένα εισόδου σε χώρο υψηλών διαστάσεων χρησιμοποιώντας μια συνάρτηση ακτινικής βάσης Gauss. Ο πυρήνας RBF υπολογίζει την ομοιότητα μεταξύ ζευγών σημείων δεδομένων με βάση την Ευκλείδεια απόστασή τους στον αρχικό χώρο χαρακτηριστικών. Αυτή η απόσταση στη συνέχεια χαρτογραφείται σε ένα χώρο υψηλότερης διάστασης, όπου η ομοιότητα μεταξύ των σημείων καθορίζεται από την εγγύτητά τους μεταξύ τους.

Ο πυρήνας RBF χαρακτηρίζεται από δύο παραμέτρους: γάμμα (γ) και C. Το γάμμα καθορίζει το πλάτος της γκαουσιανής συνάρτησης,

ελέγχοντας την επίδραση μεμονωμένων δειγμάτων εκπαίδευσης στο όριο απόφασης. Μια μικρότερη τιμή γάμμα οδηγεί σε μια ευρύτερη γκαουσιανή συνάρτηση και ένα πιο ομαλό όριο απόφασης, ενώ μια μεγαλύτερη τιμή γάμμα οδηγεί σε μια στενότερη γκαουσιανή συνάρτηση και ένα πιο περίπλοκο όριο απόφασης. Η παράμετρος C ελέγχει την αντιστάθμιση μεταξύ της μεγιστοποίησης του περιθωρίου και της ελαχιστοποίησης του σφάλματος ταξινόμησης.

Ο πυρήνας RBF είναι εξαιρετικά ευέλικτος και ικανός να καταγράφει σύνθετες μη γραμμικές σχέσεις στα δεδομένα, καθιστώντας τον κατάλληλο για ένα ευρύ φάσμα εργασιών ταξινόμησης. Ωστόσο, η επιλογή των κατάλληλων τιμών για το γάμμα και το C είναι ζωτικής σημασίας για την αποφυγή υπερπροσαρμογής και τη διασφάλιση της βέλτιστης απόδοσης του μοντέλου SVM.



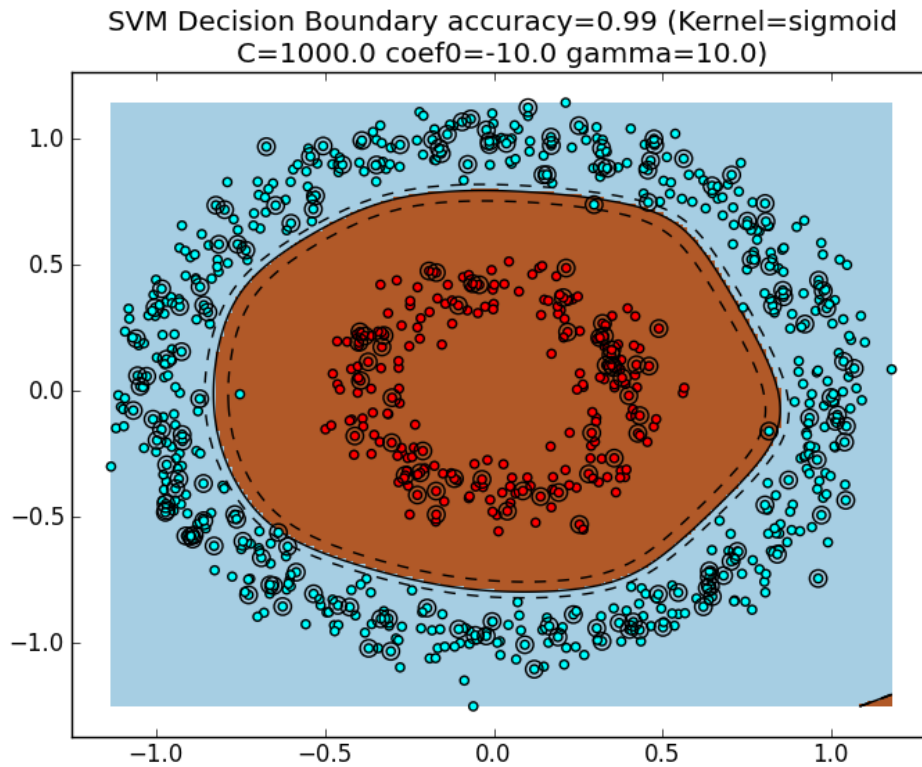
Σχήμα 7: Κατηγοριοποίηση δεδομένων SVM με χρήση RBF πυρήνα. [33]

- **Πυρήνας Σιγμοειδούς (Sigmoid Kernel)**

Ο πυρήνας σιγμοειδούς στο SVM είναι μια μη γραμμική συνάρτηση πυρήνα που χρησιμοποιείται για την αντιστοίχιση δεδομένων σε χώρο υψηλότερης διάστασης. Λειτουργεί με την εφαρμογή μιας σιγμοειδούς συνάρτησης στο βαθμωτό γινόμενο των ζευγών σημείων δεδομένων στον αρχικό χώρο χαρακτηριστικών. Η σιγμοειδής συνάρτηση μετατρέπει την είσοδο σε μια περιοχή μεταξύ 0 και 1, παράγοντας ένα ομαλό όριο απόφασης που μπορεί να συλλάβει μη γραμμικές σχέσεις στα δεδομένα.

Ο πυρήνας του σιγμοειδούς χαρακτηρίζεται από δύο παραμέτρους: γάμμα (γ) και συντελεστή (r). Το γάμμα ελέγχει το πλάτος της σιγμοειδούς συνάρτησης, προσδιορίζοντας την επίδραση μεμονωμένων δειγμάτων εκπαίδευσης στο όριο απόφασης. Μια μικρότερη τιμή γάμμα οδηγεί σε μια ευρύτερη σιγμοειδής συνάρτηση και ένα πιο ομαλό όριο απόφασης, ενώ μια μεγαλύτερη τιμή γάμμα οδηγεί σε μια στενότερη σιγμοειδή συνάρτηση και ένα πιο περίπλοκο όριο απόφασης. Ο συντελεστής (r) κλιμακώνει το βαθμωτό γινόμενο των σημείων δεδομένων πριν από την εφαρμογή της σιγμοειδούς συνάρτησης, επιτρέποντας τη λεπτομερή ρύθμιση της συμπεριφοράς του πυρήνα.

Ο πυρήνας σιγμοειδούς είναι χρήσιμος για σύνολα δεδομένων με μη γκαουσιανά χαρακτηριστικά ή όπου η σχέση μεταξύ χαρακτηριστικών και κλάσεων είναι μη γραμμική. Ωστόσο, η σωστή ρύθμιση παραμέτρων είναι απαραίτητη για την αποφυγή υπερπροσαρμογής και τη διασφάλιση της βέλτιστης απόδοσης του μοντέλου SVM. Στο [Σχήμα 8](#) φαίνεται η χρήση αυτού του πυρήνα για την κατηγοριοποίηση δεδομένων.



Σχήμα 8: Κατηγοριοποίηση δεδομένων με χρήση SVM και πυρήνα σιγμοειδούς (sigmoid kernel) [34]

Από το παραπάνω σχήμα γίνεται εμφανής η ικανότητα του σιγμοειδούς πυρήνα να διαχωρίζει με αποτελεσματικό τρόπο στοιχεία τα οποία δεν είναι γραμμικά διαχωρίσιμα.

Συμπερασματικά ο αλγόριθμος SVM διαθέτει πολλές μεθόδους για διαχωρισμό δεδομένων. Η επιλογή πυρήνα διαδραματίζει κεντρικό ρόλο στους αλγόριθμους SVM επιτρέποντας τη μοντελοποίηση μη γραμμικών σχέσεων στα δεδομένα. Είτε πρόκειται για την απλότητα του γραμμικού πυρήνα, την ευελιξία του πολυωνυμικού πυρήνα, την προσαρμοστικότητα του πυρήνα RBF ή την ομαλότητα του σιγμοειδούς πυρήνα, κάθε πυρήνας προσφέρει ξεχωριστά πλεονεκτήματα στην καταγραφή σύνθετων μοτίβων και στη βελτίωση της απόδοσης ταξινόμησης. Η επιλογή του πυρήνα εξαρτάται από τα συγκεκριμένα χαρακτηριστικά του συνόλου δεδομένων και την επιθυμητή πολυπλοκότητα του ορίου απόφασης. Ο σωστός συντονισμός παραμέτρων είναι απαραίτητος για την αποφυγή υπερπροσαρμογής και τη διασφάλιση της βέλτιστης απόδοσης του μοντέλου. Γενικά, οι πυρήνες στο SVM παρέχουν ένα ισχυρό πλαίσιο για την αντιμετώπιση ενός εκτεταμένου φάσματος εργασιών ταξινόμησης με διάφορους βαθμούς πολυπλοκότητας.

4. Εφαρμογή της Μηχανικής Μάθησης στα δεδομένα Γαλβανικής Απόκρισης Δέρματος

Στο παρόν κεφάλαιο θα παρουσιαστεί η ανάλυση της ήδη υπάρχουσας έρευνας πάνω στην οποία βασίστηκε η παρούσα διπλωματική εργασία. Κατόπιν θα παρουσιαστεί σε γενικές γραμμές η μέθοδος που ακολουθήθηκε στο πειραματικό κομμάτι.

4.1. Συνοπτική εικόνα υπάρχουσας έρευνας

Για την εκπόνηση της παρούσας διπλωματικής εργασίας χρησιμοποιήθηκε ως πρότυπο η εργασία με τίτλο «EMOTION RECOGNITION VIA GALVANIC SKIN RESPONSE: COMPARISON OF MACHINE LEARNING ALGORITHMS AND FEATURE EXTRACTION METHODS» των Deger Ayata, Yusuf Yasalan και Mustafa Kamasak. [35]

Στην παραπάνω εργασία επιχειρείται να αναγνωριστούν συναισθήματα μέσω δεδομένων καταγραφής μετρήσεων GSR. Για την κατηγοριοποίηση των δειγμάτων σε ομάδες συναισθημάτων εξήχθησαν αρκετά χαρακτηριστικά από τις μετρήσεις GSR. Κατόπιν, αυτά τα χαρακτηριστικά χρησιμοποιήθηκαν για την δημιουργία διανυσμάτων γνωρισμάτων (feature vectors).

Στη συνέχεια, με χρήση αλγορίθμων μηχανικής μάθησης τα διανύσματα αυτά χρησιμοποιήθηκαν για τον υπολογισμό του επιπέδου συναισθηματικής διέγερσης (Arousal) που σε συνδυασμό με τον χαρακτηρισμό του επαγόμενου συναισθήματος ως θετικού ή αρνητικού (Valence) οδήγησαν στην απόδοση χαρακτηρισμού στα συναισθήματα των εξεταζόμενων ατόμων.

Τέλος, τα δεδομένα που προέκυψαν από την εργασία συγκρίθηκαν με τα ήδη γνωστά συναισθήματα όπως είχαν περιγραφεί από τους συμμετέχοντες. Η κατηγοριοποίηση από τους αλγορίθμους της μηχανικής

μάθησης παρήγαγε αποτελέσματα που συμφωνούσαν σε πολύ ικανοποιητικό βαθμό με τους προηγούμενους χαρακτηρισμούς των τιμών.

4.2. Μοντέλο PAD

4.2.1. Ορισμός μοντέλου

Το μοντέλο Ευχαρίστησης – Διέγερσης – Κυριαρχίας (Pleasure Arousal Dominance, PAD) αναπτύχθηκε από τους Albert Mehrabian και James A. Russell, παρέχει ένα ολοκληρωμένο πλαίσιο για την περιγραφή και τη μέτρηση των συναισθηματικών καταστάσεων. Αυτό το μοντέλο χρησιμοποιεί τρεις αριθμητικές διαστάσεις για να αναπαραστήσει όλα τα συναισθήματα [36]:

- Ευχαρίστηση (Pleasure - P): Αυτή η διάσταση μετρά πόσο ευχάριστα ή δυσάρεστα αισθάνεται κανείς για κάτι. Για παράδειγμα, τόσο ο θυμός όσο και ο φόβος ανήκουν στην πλευρά της δυσαρέσκειας, ενώ η χαρά αντιπροσωπεύει ένα ευχάριστο συναίσθημα.
- Διέγερση (Arousal - A): Η Κλίμακα Διέγερσης-Μη διέγερσης μετράει πόσο ενεργητικά ή ανιαρά αισθάνεται κάποιος. Πρόκειται περισσότερο για το επίπεδο ενεργοποίησης. Για παράδειγμα, η θλίψη και η κατάθλιψη μπορεί να είναι έντονα συναισθήματα χαμηλής διέγερσης, ενώ η οργή έχει υψηλότερη κατάσταση διέγερσης.
- Κυριαρχία (Dominance - D): Η Κλίμακα Κυριαρχίας-Υποταγής αντικατοπτρίζει το κατά πόσον το άτομο αισθάνεται να έχει ή όχι τον έλεγχο σε μια κατάσταση (κυριαρχία - υποταγή). Για παράδειγμα ενώ τόσο ο φόβος όσο και ο θυμός είναι δυσάρεστα συναισθήματα, ο θυμός είναι κυρίαρχος, ενώ ο φόβος είναι υποτακτικός.

Το μοντέλο PAD χρησιμοποιείται σε δύο εκδοχές. Μια δισδιάστατη, όπου τα συναισθήματα κατηγοριοποιούνται μόνο μέσω ευχαρίστησης (Valence) και διέγερσης (Arousal), και μια τρισδιάστατη όπου εκτός από τα δύο παραπάνω χαρακτηριστικά εμφανίζεται και η Κυριαρχία (Dominance). Στην πρότυπη εργασία χρησιμοποιήθηκε η δισδιάστατη εκδοχή του μοντέλου.

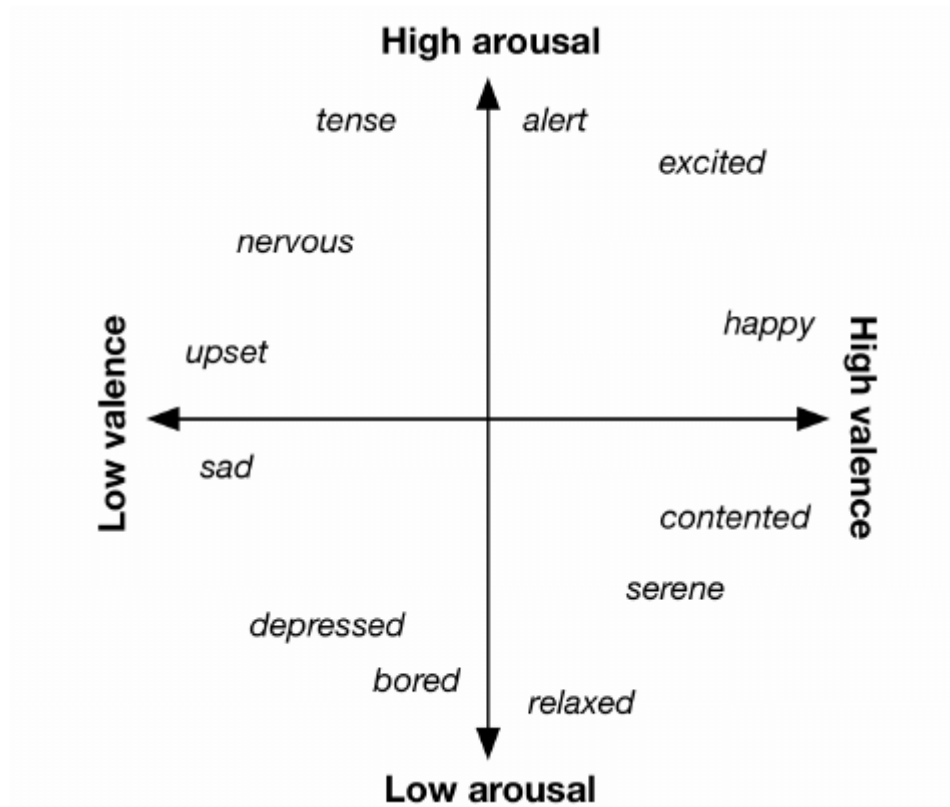
4.2.2. Εφαρμογές μοντέλου PAD

Το μοντέλο PAD έχει βρει διάφορες εφαρμογές στον σημερινό κόσμο.

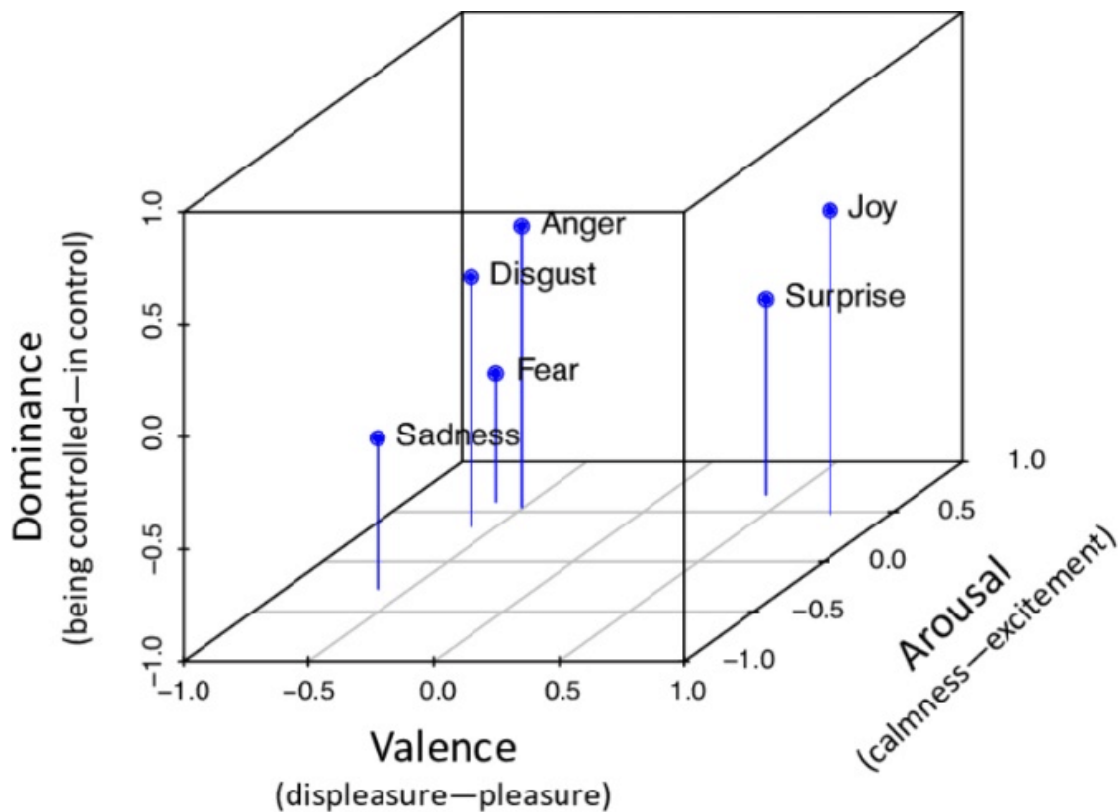
- Περιβαλλοντική Ψυχολογία: Χρησιμοποιήθηκε αρχικά για να διερευνήσει πώς τα φυσικά περιβάλλοντα επηρεάζουν τους ανθρώπους μέσω της συναισθηματικής τους επίδρασης.
- Συναισθηματικά επεισόδια (Emotional Episodes): Ο James A. Russell το χρησιμοποίησε για να αναπτύξει μια θεωρία για τα σχετικά σύντομα συναισθηματικά φορτισμένα γεγονότα [37].
- Φυσιολογική Θεωρία του Συναισθήματος (Physiological Theory of Emotion): Ο Peter Lang και οι συνεργάτες του την επέκτειναν για να προτείνουν μια φυσιολογική θεωρία του συναισθήματος [38].
- Μη λεκτική επικοινωνία: Έχει χρησιμοποιηθεί για τη μελέτη της γλώσσας του σώματος και των μη λεκτικών ενδείξεων στην ψυχολογία [39].
- Μελέτη καταναλωτών και προώθηση προϊόντων: Εφαρμόζεται για την κατανόηση της συμπεριφοράς και της συναισθηματικής ανταπόκρισης των καταναλωτών [40].
- Εικονική Πραγματικότητα: Χρησιμοποιούνται για την κατασκευή κινούμενων χαρακτήρων που εκφράζουν συναισθήματα [41].

Στον σημερινό δυναμικό κόσμο, το μοντέλο συναισθηματικής κατάστασης PAD συνεχίζει να παίζει καθοριστικό ρόλο σε διάφορους τομείς. Από τις στρατηγικές μάρκετινγκ μέχρι τον εικονικό σχεδιασμό χαρακτήρων, η επιρροή του είναι διευρυμένη. Αναλύοντας τα συναισθήματα σε

ευχαρίστηση, διέγερση και κυριαρχία, αποκτούμε γνώσεις για την ανθρώπινη συμπεριφορά, την επικοινωνία και τη λήψη αποφάσεων. Καθώς περιηγούμαστε σε αυτό το περίπλοκο τοπίο, το μοντέλο PAD χρησιμεύει ως πυξίδα, καθοδηγώντας μας προς μια βαθύτερη κατανόηση των δικών μας συναισθημάτων και των συναισθημάτων των άλλων. Έτσι, είτε αναλύουμε τις αντιδράσεις των καταναλωτών, δημιουργούμε κινούμενους χαρακτήρες ή εξερευνούμε φυσιολογικές αποκρίσεις, το μοντέλο PAD παραμένει ένα πολύτιμο εργαλείο στην προσπάθειά μας να αποκωδικοποιήσουμε την περίπλοκη ποικιλία των ανθρώπινων συναισθημάτων και αλληλεπιδράσεων.



Σχήμα 9: Δισδιάστατη απεικόνιση του μοντέλου PAD. Είναι ορατοί οι χαρακτηρισμοί των συναισθημάτων ανάλογα με την θέση τους σε σχέση με τους άξονες. [42]



Σχήμα 10: Τρισδιάστατη απεικόνιση του μοντέλου PAD με την θέση όπου τοποθετούνται τα διάφορα συναισθήματα σε σχέση με τους άξονες. [43]

Η σύγκριση μεταξύ δισδιάστατων και τρισδιάστατων εκδόσεων του μοντέλου PAD (Pleasure-Arousal-Dominance) προσφέρει πληροφορίες για τη διαφοροποιημένη αναπαράσταση συναισθημάτων. Ενώ το δισδιάστατο μοντέλο απεικονίζει συναισθήματα κατά μήκος των αξόνων ευχαρίστησης-διέγερσης (Valence - Arousal), η προσθήκη κυριαρχίας ως τρίτης διάστασης στο τρισδιάστατο μοντέλο παρέχει μια πιο ολοκληρωμένη κατανόηση των συναισθηματικών εμπειριών. Το δισδιάστατο μοντέλο προσφέρει απλότητα και ευκολία ερμηνείας και είναι κατάλληλο για βασικές αναλύσεις και απεικονίσεις. Αντίθετα, το τρισδιάστατο μοντέλο εισάγει βάθος και πολυπλοκότητα, επιτρέποντας μια πιο πλούσια απεικόνιση των συναισθημάτων λαμβάνοντας υπόψη την επιρροή της κυριαρχίας στις συναισθηματικές αντιδράσεις. Ενώ και οι δύο εκδόσεις του μοντέλου PAD είναι πολύτιμες για τη μελέτη συναισθηματικών εμπειριών, το τρισδιάστατο μοντέλο προσφέρει μια πιο πλήρη αναπαράσταση, επιτρέποντας στους ερευνητές να διερευνήσουν την περίπλοκη αλληλεπίδραση μεταξύ ευχαρίστησης, διέγερσης και κυριαρχίας στη διαμόρφωση συναισθηματικών αντιδράσεων και συμπεριφορών.

4.3. Μεθοδολογία πρότυπης έρευνας

Στο παρόν υποκεφάλαιο θα παρουσιαστεί η διαδικασία που ακολουθήθηκε στην εργασία που χρησιμοποιήθηκε ως πρότυπο της παρούσας διπλωματικής εργασίας.

Τα δεδομένα από τα οποία έγινε η εξαγωγή στοιχείων και η εφαρμογή των αλγορίθμων μηχανικής μάθησης προέρχονται από το σύνολο δεδομένων Dear [44]. Αυτό το σύνολο δεδομένων αποτελείται από δεδομένα ηλεκτροεγκεφαλογραφήματος (EEG) και άλλων βιομετρικών δεδομένων από 32 συμμετέχοντες οι οποίοι παρακολούθησαν 40 βίντεο διάρκειας ενός λεπτού το κάθε ένα. Αυτά είχαν ως σκοπό να προκαλέσουν συναισθηματική διέγερση στα άτομα που τα παρακολούθησαν.

Οι συμμετέχοντες βαθμολόγησαν κάθε βίντεο όσον αφορά την διέγερση, την ευχαρίστηση, την αρέσκεια / δυσαρέσκεια, την κυριαρχία και την οικειότητα.

Από το παραπάνω σύνολο δεδομένων μόνο οι μετρήσεις που αφορούν GSR χρησιμοποιήθηκαν για την εξαγωγή στοιχείων. Αρχικά, το σήμα GSR υποβλήθηκε σε διαδικασίες μείωσης του θορύβου και δειγματοληψίας. Κατόπιν, πραγματοποιήθηκε επεξεργασία του σήματος GSR στο πεδίο του χρόνου και μετά από αυτήν την επεξεργασία προέκυψαν χαρακτηριστικά που αφορούν το σήμα όπως ο αριθμητικός μέσος, μέγιστη και ελάχιστη τιμή, τυπική απόκλιση, διασπορά, συντελεστής λοξότητας (skewness coefficient), συντελεστής κύρτωση (kurtosis coefficient), εντροπία και μέση ενέργεια.

Στη συνέχεια, τα δεδομένα αυτά χρησιμοποιήθηκαν για την εκπαίδευση των αλγορίθμων μηχανικής μάθησης αλλά και ως σύνολο νέων άορατων δειγμάτων για δοκιμές ακρίβειας των μοντέλων. Χρησιμοποιήθηκαν οι εξής αλγόριθμοι: Δέντρο Απόφασης (Decision Tree – DT), Τυχαίο Δάσος (Random Forest - RF), k-πλησιέστεροι γείτονες (k-Nearest Neighbors - kNN) και το Support Vector Machine (SVM).

Τέλος, τα μοντέλα αξιολογήθηκαν βάση του κατά πόσον συμφωνούν με τις τιμές διέγερσης και ευχαρίστησης αλλά και τα συναισθήματα που δήλωσαν ότι αισθάνονταν οι ίδιοι οι συμμετέχοντες. Για παράδειγμα ο αλγόριθμος kNN πέτυχε ποσοστό ακρίβειας 58,12% στο επίπεδο διέγερσης και 60.54% στο επίπεδο ευχαρίστησης. Το SVM είχε αρκετά υψηλό ποσοστό με 71,4% στο επίπεδο διέγερσης και 70.54% στο επίπεδο ευχαρίστησης.

Στην παρούσα εργασία αξιοποιήθηκε για την ανάλυση σήματος και η τεχνική του Empirical Mode Decomposition (EMD) με την οποία επετεύχθησαν ακόμα καλύτερα αποτελέσματα.

Η EMD αποτελεί μια τεχνική επεξεργασίας σήματος που αναλύει ένα σήμα στα υποκείμενα ταλαντευτικά στοιχεία του, γνωστά ως Intrinsic Mode Functions (IMFs). Σε αντίθεση με τις παραδοσιακές μεθόδους, η EMD βασίζεται σε δεδομένα και δεν απαιτεί προηγούμενη γνώση του περιεχομένου συχνότητας του σήματος ή των χαρακτηριστικών των υποκείμενων ταλαντώσεων [45].

Με την χρήση EMD επετεύχθησαν ποσοστά ακρίβειας 81,81% και 89,29% για τα επίπεδα διέγερσης και ευχαρίστησης αντίστοιχα. Τα ποσοστά αυτά είναι σαφώς καλύτερα από τα αντίστοιχα που προέκυψαν χωρίς την χρήση EMD τα οποία είναι 71,53% για το επίπεδο διέγερσης και 71,04% για το επίπεδο ευχαρίστησης.

Συμπερασματικά, η τεχνική EMD αυξάνει την ακρίβεια των προβλέψεων κάτι το οποίο δικαιολογεί απόλυτα κατά τους συγγραφείς της εργασίας την επιπρόσθετη επεξεργασία του σήματος GSR.

ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

5. Προεπεξεργασία

Όπως προαναφέρθηκε, ακολουθήθηκε η μέθοδος της εργασίας με τίτλο «EMOTION RECOGNITION VIA GALVANIC SKIN RESPONSE: COMPARISON OF MACHINE LEARNING ALGORITHMS AND FEATURE EXTRACTION METHODS» των Deger Ayata, Yusuf Yasalan και Mustafa Kamasak. [35]

Προτού γίνει οποιαδήποτε εφαρμογή αλγορίθμου μηχανικής μάθησης έπρεπε να γίνει η σωστή προεπεξεργασία των στοιχείων. Η προεπεξεργασία δεδομένων για αλγόριθμους μηχανικής μάθησης περιλαμβάνει καθαρισμό, μετασχηματισμό και προετοιμασία ακατέργαστων δεδομένων για να διασφαλιστεί η καταλληλότητά τους για ανάλυση. Τεχνικές όπως ο καθαρισμός δεδομένων, και η κανονικοποίηση εφαρμόζονται για τη βελτίωση της ποιότητας των δεδομένων και τη βελτίωση της απόδοσης του μοντέλου.

5.1. Έλεγχος δεδομένων και επιλογή στοιχείων για περαιτέρω εξέταση

Στην αρχή ελέγχθηκε προσεκτικά το λογιστικό φύλλο με το σύνολο των δεδομένων προκειμένου να ανιχνευθούν τυχόν ελλείψεις και ασυνέπειες στα στοιχεία που έχουν συλλεχθεί από τον αισθητήρα. Κατόπιν, επιλέχθηκαν τα δεδομένα από το λογιστικό φύλλο με τα οποία θα γίνει η εκπαίδευση και η εφαρμογή του μοντέλου. Τα δεδομένα αυτά είναι οι τιμές (σε μS) της αγωγιμότητας του δέρματος.

Στο παρακάτω σχήμα ([Σχήμα 11](#)) φαίνεται η αρχική μορφή των δεδομένων στο λογιστικό φύλλο. Τα δεδομένα συλλέχθηκαν με την βοήθεια αισθητήρα Arduino εξού και οι μετρήσεις σε μονάδες Arduino (units arduino).

		test1					
Αριθμηση	Time (sec)	units arduino	Vx (Volt)	I(μA)	X(KΩ)	Y (μS)	
1	2	585	2,8665	4,267	671,783454	1,48857492	
2	4	585	2,8665	4,267	671,783454	1,48857492	
3	6	586	2,8714	4,2572	674,480879	1,48262172	
4	8	586	2,8714	4,2572	674,480879	1,48262172	
5	10	585	2,8665	4,267	671,783454	1,48857492	
6	12	584	2,8616	4,2768	669,098391	1,4945485	
7	14	582	2,8518	4,2964	663,765013	1,50655726	
8	16	582	2,8518	4,2964	663,765013	1,50655726	
9	18	581	2,8469	4,3062	661,11653	1,51259264	
10	20	580	2,842	4,316	658,480074	1,51864884	
11	22	581	2,8469	4,3062	661,11653	1,51259264	
12	24	580	2,842	4,316	658,480074	1,51864884	
13	26	578	2,8322	4,3356	653,242919	1,53082409	
14	28	576	2,8224	4,3552	648,052902	1,5430839	
15	30	574	2,8126	4,3748	642,90939	1,55542914	
16	32	573	2,8077	4,3846	640,354878	1,56163408	
17	34	568	2,7832	4,4336	627,751714	1,59298649	
18	36	566	2,7734	4,4532	622,788107	1,60568256	
19	38	565	2,7685	4,463	620,322653	1,61206429	
20	40	563	2,7587	4,4826	615,424084	1,62489578	
21	42	563	2,7587	4,4826	615,424084	1,62489578	
22	44	565	2,7685	4,463	620,322653	1,61206429	

Σχήμα 11: Αρχική μορφή δεδομένων στο λογιστικό φύλλο.

Κατόπιν από τις μονάδες Arduino μπόρεσαν να εξαχθούν η τιμή σε Volt για την τάση, η τιμή σε μA για το ρεύμα, για την αντίσταση σε KΩ και τέλος για την αγωγιμότητα σε μS.

Από τις τιμές της αγωγιμότητας του δέρματος εξήχθησαν τα χαρακτηριστικά : μέση τιμή (mean), χρόνος ανύψωσης (rise time), πλάτος κορυφής (peak amplitude), αφάνεια (latency) και η διαφορά πλάτους κορυφής πλην την μέση τιμή (peak amplitude – mean). Το τελευταίο χαρακτηριστικό εξήχθη μετά από δοκιμές που πραγματοποιήθηκαν με τον αλγόριθμο SVM καθώς διαπιστώθηκε ότι με αυτό επιτυγχάνεται καλύτερος διαχωρισμός των δειγμάτων σε κλάσεις (αγχωμένοι, μη αγχωμένοι).

	Mean	Peak-Mean	Rise Time	Peak Amplitude (μ S)	Latency (sec)
Test 1	1.56	0.06	33	0.14	5
Test 2	2.58	0.37	26	0.88	16
Test 3	2.28	0.71	8	1.06	6

Σχήμα 12: Τελική μορφή δεδομένων.

Τα παραπάνω δεδομένα χρησιμοποιήθηκαν ως σύνολο εκπαίδευσης και ελέγχου επίδοσης του μοντέλου SVM σε συνδυασμό με την τεχνική της ομαδοποίησης (clustering) για την απόδοση ετικέτας (αγχωμένοι, μη αγχωμένοι) στα δείγματα.

Ακόμα μια προσέγγιση που ακολουθήθηκε ήταν η εκπαίδευση του μοντέλου με δείγματα από το σύνολο δεδομένων YAAD [46] και ως δείγμα δοκιμών τα δεδομένα μετρήσεων από τους διαγνωστικούς ελέγχους για covid -19 που παρουσιάστηκαν παραπάνω.

Στο επόμενο κεφάλαιο θα αναλυθούν οι δυο παραπάνω προσεγγίσεις και θα παρουσιαστούν εκτενέστερα τα δεδομένα που προέκυψαν από τα πειράματα.

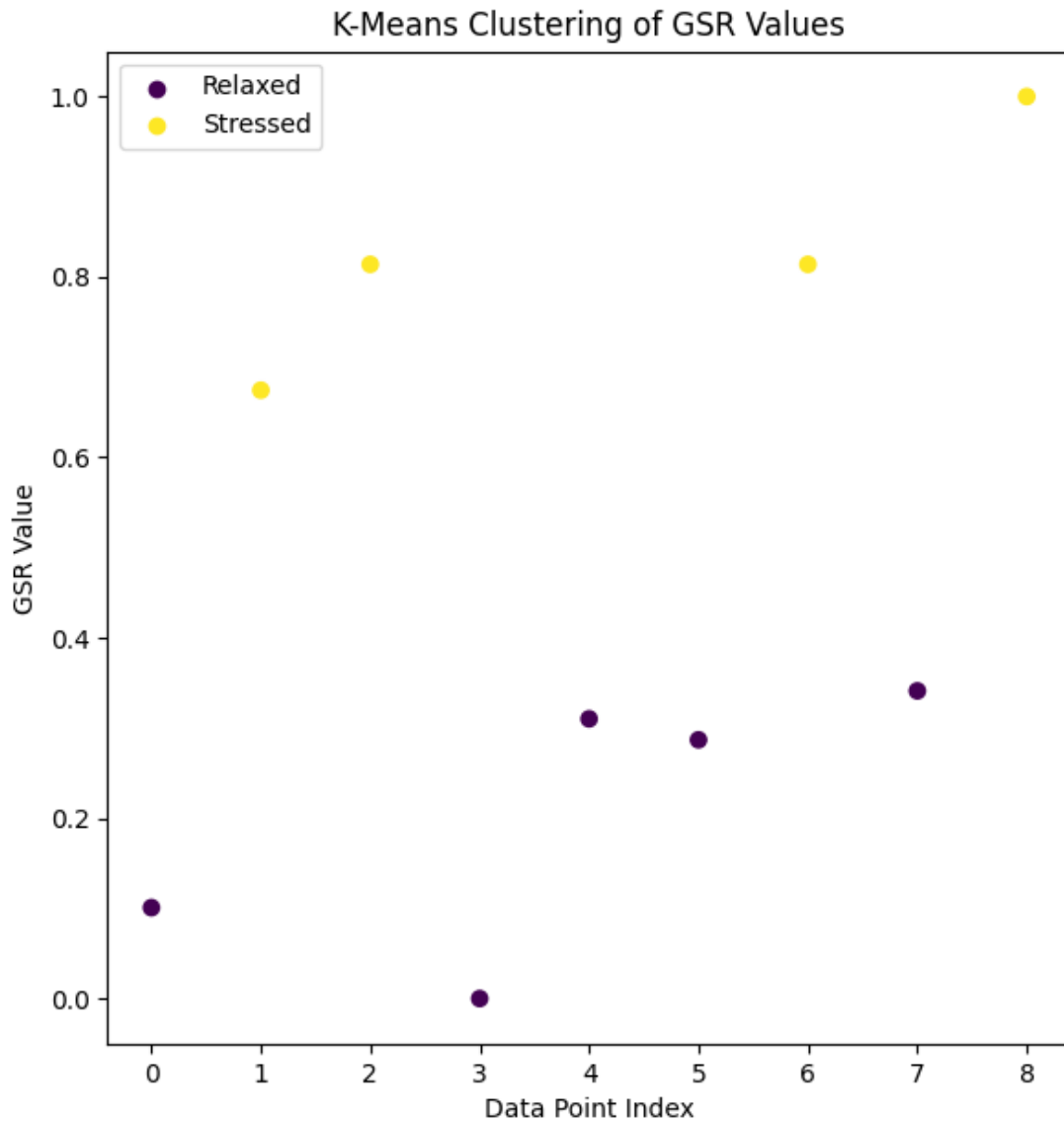
6. Εκπαίδευση και Κατηγοριοποίηση δειγμάτων από το μοντέλο SVM

6.1. Ομαδοποίηση και πρώτη κατάταξη

Ως πρώτη προσέγγιση για το μοντέλο επιλέχθηκε ένα μέρος από το δείγμα των ατόμων που συμμετείχαν στην διαδικασία διαγνωστικού ελέγχου Covid-19. Το μικρό μέγεθος του δείγματος είναι ιδανικό ως πρώτη προσέγγιση για το πώς λειτουργεί ο αλγόριθμος SVM.

Στο [Σχήμα 13](#) φαίνονται τα αποτελέσματα της τεχνικής της ομαδοποίησης (clustering). Με μωβ χρώμα φαίνονται τα άτομα που δεν παρουσίασαν ιδιαίτερη μεταβολή στα επίπεδα του GSR κατά την διάρκεια της διαδικασίας διαγνωστικού ελέγχου. Αντιθέτως με κίτρινο παρουσιάζονται τα άτομα που κατηγοριοποιήθηκαν ως αγχωμένα από τον αλγόριθμο της ομαδοποίησης. Η ομαδοποίηση αναθέτει ετικέτες (αγχωμένοι – stressed, μη αγχωμένοι – relaxed) στα δεδομένα. Κατόπιν αφού τα δεδομένα πλέον έχουν αποκτήσει χαρακτηρισμό μπορεί να εφαρμοστεί σε αυτά ο αλγόριθμος SVM οπότε και έχουμε μια διαδικασία επιβλεπόμενης μάθησης (supervised learning).

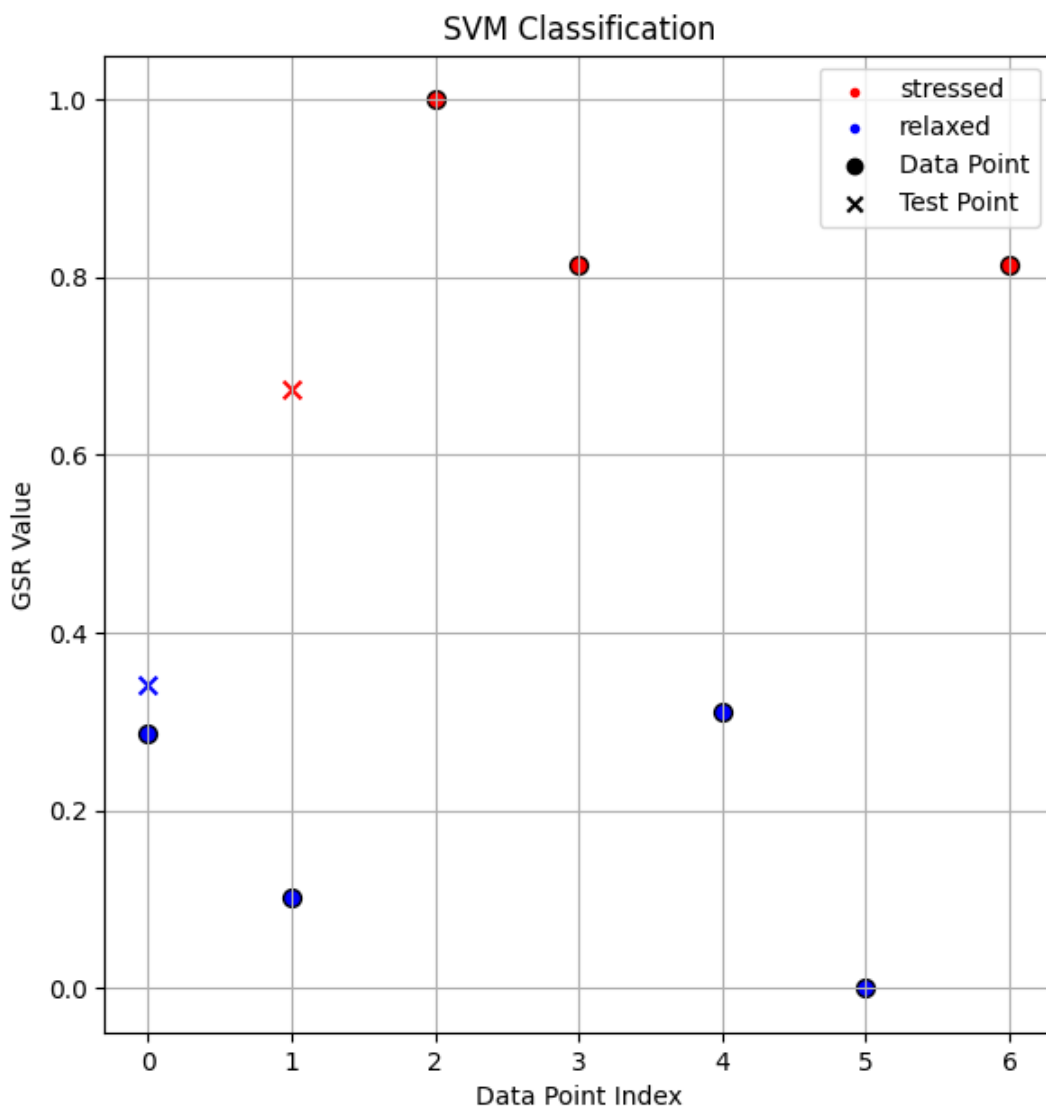
Στον κατακόρυφο άξονα φαίνεται η τιμή του πλάτους κορυφής GSR (peak amplitude) η οποία έχει υποστεί την διαδικασία της κανονικοποίησης πριν αναπαρασταθεί. Στον οριζόντιο άξονα φαίνεται ο δείκτης σειράς για το κάθε δείγμα (1^ο δείγμα, 2^ο δείγμα, και ούτω καθ' εξής).



Σχήμα 13: Αποτέλεσμα της τεχνικής ομαδοποίησης σε τμήμα δείγματος εξεταζόμενων με διαγνωστικό έλεγχο για Covid-19. Με κίτρινο χρώμα φαίνονται τα άτομα τα οποία παρουσίασαν αυξημένα επίπεδα άγχους. Με μωβ αυτά που δεν εμφάνισαν αξιοσημείωτη μεταβολή στα δεδομένα GSR.

Η κανονικοποίηση στη μηχανική μάθηση διασφαλίζει ότι όλα τα χαρακτηριστικά έχουν παρόμοια κλίμακα. Με την κλιμάκωση των χαρακτηριστικών σε ένα τυπικό εύρος, όπως μεταξύ 0 και 1, η κανονικοποίηση εμποδίζει τις μεταβλητές με μεγαλύτερη τιμή να κυριαρχούν στη διαδικασία εκμάθησης. Αυτό βελτιώνει τη σταθερότητα του μοντέλου, τη σύγκλιση και τη δικαιοσύνη στις συγκρίσεις μεταξύ των χαρακτηριστικών.

Στο [Σχήμα 14](#) φαίνεται το αποτέλεσμα της εφαρμογής του αλγορίθμου SVM στα δεδομένα που έχουν πρώτα υποστεί κατηγοριοποίηση. Αρχικά έγινε διαχωρισμός των δεδομένων με τυχαίο τρόπο σε σύνολο εκπαίδευσης (training set) και σύνολο δοκιμής (testing set). Το σύνολο εκπαίδευσης αποτελεί το 80 % του αρχικού συνόλου δεδομένων και αντίστοιχα το σύνολο δοκιμής το 20 % του αρχικού συνόλου δεδομένων.



Σχήμα 14: Αποτελέσματα κατηγοριοποίησης με αλγόριθμο SVM. Με κύκλο αναπαρίστανται τα δεδομένα του συνόλου εκπαίδευσης ενώ με το σύμβολο «x» τα δεδομένα του συνόλου δοκιμής. Με μπλε χρώμα αναπαρίστανται τα δεδομένα που κατηγοριοποιήθηκαν ως μη αγχωμένα ενώ με κόκκινο οι εξεταζόμενοι που κατηγοριοποιήθηκαν ως αγχωμένοι.

Όσον αφορά την ακρίβεια (Accuracy) που επιτυγχάνει το μοντέλο σε αυτήν την περίπτωση αυτή είναι 100 % . Αυτό είναι απόλυτα αναμενόμενο και οφείλεται αφενός στον μικρό αριθμό δεδομένων (μόλις 9 μετρήσεις) και αφετέρου στο γεγονός ότι οι τιμές για το πλάτος κορυφής GSR των δεδομένων είναι εύκολα διαχωρίσιμες. Τα δύο στοιχεία του συνόλου δοκιμών μπορούν να διαχωριστούν με γραμμικό τρόπο.

Σε περίπτωση που ο αριθμός των δεδομένων είναι μεγαλύτερος τότε αυξάνεται η πιθανότητα αυτά να είναι πιο δύσκολο να διαχωριστούν σε δύο κλάσεις με ένα γραμμικό όριο απόφασης. Ως αποτέλεσμα αναμένεται να εμφανιστούν περιπτώσεις τόσο λανθασμένης κατηγοριοποίησης δειγμάτων δοκιμής (misclassification) όσο και δείγματα εκπαίδευσης τα οποία θα αντιστοιχούν για παράδειγμα σε αγχωμένα άτομα αλλά θα βρίσκονται σε περιοχή όπου το SVM θεωρεί ότι ανήκει σε μη αγχωμένους. Αυτό θα έχει ως προφανή συνέπεια η ακρίβεια του μοντέλου να μειωθεί.

Παραπάνω χρησιμοποιήθηκε ο αλγόριθμος ομαδοποίησης k-means από την βιβλιοθήκη scikit-learn. Στην Python, η βιβλιοθήκη scikit-learn παρέχει υλοποίηση του αλγόριθμου ομαδοποίησης k-means μέσω της κλάσης «KMeans». Προσφέρει έναν απλό και αποτελεσματικό τρόπο για την εκτέλεση ομαδοποίησης σε σύνολα δεδομένων. Καθορίζοντας τον αριθμό των ομάδων προς αναγνώριση, οι χρήστες μπορούν να εφαρμόσουν ομαδοποίηση k-means για να χωρίσουν τα δεδομένα σε διακριτές ομάδες με βάση τις ομοιότητές τους. Ο αλγόριθμος εκχωρεί επαναληπτικά σημεία δεδομένων στα πλησιέστερα κέντρα ομάδας και ενημερώνει τα κέντρα με βάση τον μέσο όρο των σημείων που έχουν εκχωρηθεί σε κάθε σύμπλεγμα. Αυτή η διαδικασία συνεχίζεται μέχρι τη σύγκλιση, όπου οι εκχωρήσεις στις ομάδες δεν αλλάζουν πλέον σημαντικά. Επιπλέον, το scikit-learn παρέχει χρήσιμες μεθόδους για την αξιολόγηση των αποτελεσμάτων της ομαδοποίησης και την οπτικοποίηση των αποτελεσμάτων.

Παρά το μικρό πλήθος των δεδομένων η παραπάνω περίπτωση αποτελεί ιδανικό παράδειγμα προκειμένου να φανεί το πώς λειτουργεί η κατηγοριοποίηση με χρήση SVM. Σε επόμενα παραδείγματα θα γίνει εμφανές το πώς συμπεριφέρονται κατά την διάρκεια της κατηγοριοποίησης οι διάφοροι τύποι SVM που χρησιμοποιήθηκαν.

6.2. SVM και σύνολο δεδομένων YAAD

Στο παρόν υποκεφάλαιο θα αναλυθεί η δεύτερη τακτική που ακολουθήθηκε για τον συνδυασμό κατηγοριοποίησης μέσω SVM. Αυτή συνίσταται στην χρήση δεδομένων μετρήσεων GSR από το σύνολο δεδομένων YAAD. Το σύνολο δεδομένων Young Adult's Affective Data (YAAD) [46] είναι μια πολύτιμη πηγή για εργασίες αναγνώρισης συναισθημάτων. Περιλαμβάνει δεδομένα ηλεκτροκαρδιογραφημάτων (ECG) και GSR που συλλέχθηκαν από 25 συμμετέχοντες χρησιμοποιώντας αισθητήρες ECG και GSR shimmer3.

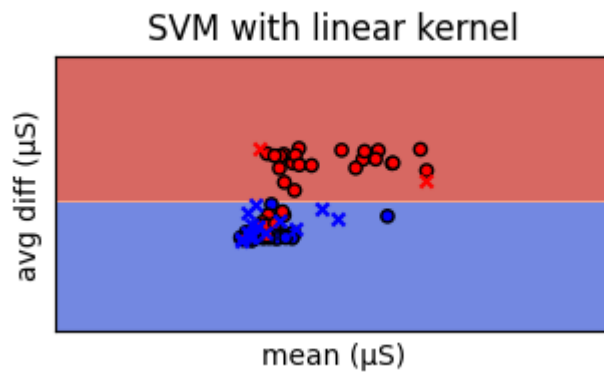
Τα δεδομένα GSR του YAAD χρησιμοποιήθηκαν για την εκπαίδευση του μοντέλου. Κατόπιν για την διαδικασία δοκιμών χρησιμοποιήθηκαν τα δεδομένα από τα άτομα που υποβλήθηκαν σε διαδικασία διαγνωστικού ελέγχου για Covid-19 και δεδομένα από το σύνολο δεδομένων «Collection and validation of psychophysiological data from professional and amateur players: a multimodal esports dataset.» των Smerdon, A., Zhou, B., Lukowicz, P., & Somon, A [47]. Το συγκεκριμένο σύνολο δεδομένων αφορά μια πληθώρα βιομετρικών και μη δεδομένων που έχουν συλλεγεί από 10 παίκτες κατά τη διάρκεια 22 αγώνων του διαδικτυακού παιχνιδιού League of Legends. Από τα δεδομένα του συνόλου αυτού στην παρούσα διπλωματική εργασία μόνο οι μετρήσεις GSR έχουν ληφθεί υπόψιν.

Στη συνέχεια, θα παρουσιαστούν οι δοκιμές με τα διάφορα μοντέλα SVM. Τα χαρακτηριστικά που έχουν χρησιμοποιηθεί για την κατηγοριοποίηση των δειγμάτων είναι η μέση τιμή (mean value) της γραφικής παράστασης και η διαφορά πλάτους κορυφής GSR μείον την μέση τιμή (GSR peak amplitude – mean value). Κατόπιν, στο Κεφάλαιο 3 θα υπάρξει μια πιο προσεκτική ματιά στο πόσο ακριβές είναι το κάθε μοντέλο που κατασκευάστηκε.

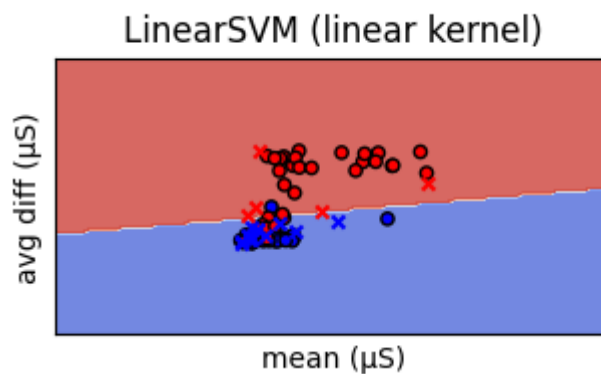
6.2.1. Κατηγοριοποίηση με Γραμμικό SVM

Αρχικά, επιχειρήθηκε τα δείγματα να κατηγοριοποιηθούν με χρήση γραμμικού πυρήνα στον αλγόριθμο SVM. Ο γραμμικός πυρήνας προσπαθεί να διαχωρίσει τα δείγματα δημιουργώντας ένα γραμμικό όριο απόφασης

(linear decision boundary). Τα αποτελέσματα φαίνονται παρακάτω στο [Σχήμα 15](#) και στο [Σχήμα 16](#).



Σχήμα 15: Κατηγοριοποίηση SVM με χρήση απλού γραμμικού πυρήνα.

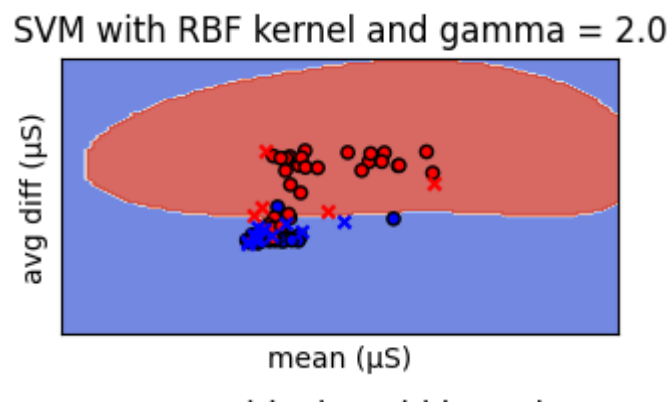


Σχήμα 16: Κατηγοριοποίηση με χρήση του μοντέλου LinearSVM της γλώσσας προγραμματισμού Python.

Στις παραπάνω κατηγοριοποιήσεις χρησιμοποιήθηκαν δύο μοντέλα SVM για γραμμικό διαχωρισμό και χαρακτηρισμό των δειγμάτων. Το πρώτο (SVM) αποτελεί την κλασική υλοποίηση του αλγορίθμου SVM με χρήση γραμμικού πυρήνα. Το δεύτερο μοντέλο όμως (LinearSVM) χρησιμοποιεί διαφορετική υλοποίηση η οποία έχει σχεδιαστεί με αποκλειστικό στόχο την γρηγορότερη εκπαίδευση και άρα παραγωγή αποτελεσμάτων. Συμπερασματικά, το LinearSVM αποτελεί ιδανική επιλογή χρήσης σε περιπτώσεις όπου αφενός είναι επιθυμητός ο γραμμικός διαχωρισμός των δεδομένων σε κλάσεις και αφετέρου το πλήθος τους είναι αρκετά μεγάλο.

6.2.2. Κατηγοριοποίηση με χρήση Πυρήνα Συνάρτησης Ακτινικής Βάσης (RBF kernel)

Ως επόμενη μέθοδος κατηγοριοποίησης επιλέχθηκε η μέθοδος του πυρήνα συνάρτησης ακτινικής βάσης (RBF kernel). Ο πυρήνας RBF είναι εξαιρετικά ευέλικτος και ικανός να καταγράψει σύνθετες μη γραμμικές σχέσεις στα δεδομένα. Για την περίπτωση αυτή επιλέχθηκε τιμή της παραμέτρου γάμμα (gamma) να είναι ίση με 2.0. Μια μικρότερη τιμή γάμμα οδηγεί σε μια ευρύτερη γκαουσιανή συνάρτηση και ένα πιο ομαλό όριο απόφασης, ενώ μια μεγαλύτερη τιμή γάμμα οδηγεί σε μια στενότερη γκαουσιανή συνάρτηση και ένα πιο περίπλοκο όριο απόφασης. Στο [Σχήμα 17](#) βλέπουμε την περίπτωση αυτής της κατηγοριοποίησης.



Σχήμα 17: Κατηγοριοποίηση με χρήση πυρήνα Συνάρτησης Ακτινικής Βάσης (RBF kernel) και τιμή γάμμα (gamma) ίση με 2.0.

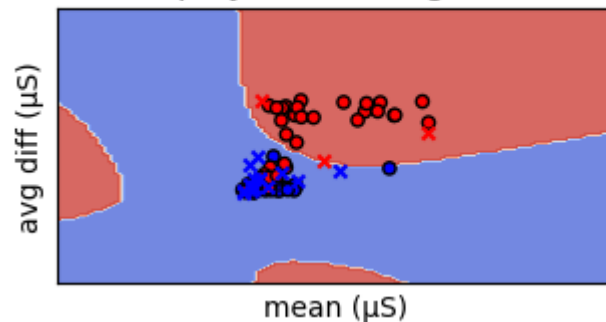
Είναι φανερό ότι πλέον το όριο απόφασης αποκτά τελείως διαφορετική μορφή σε σχέση με τον γραμμικό τρόπο διαχωρισμού. Αυτό οφείλεται στην μετατροπή των δεδομένων εισόδου σε χώρο υψηλότερης διάστασης και χρήσης της εγγύτητάς τους σε αυτόν τον νέο χώρο βάσει της ευκλείδειας απόστασης προκειμένου να αποφασιστεί το σε ποια κλάση θα καταταχθεί το κάθε ένα.

6.2.3. Κατηγοριοποίηση με χρήση πολυωνυμικού πυρήνα

Στη συνέχεια, θα παρουσιαστεί η περίπτωση κατηγοριοποίησης των ίδιων δειγμάτων με την χρήση πολυωνυμικού πυρήνα (polynomial kernel). Ο πολυωνυμικός πυρήνας υπολογίζει το εσωτερικό γινόμενο ζευγών σημείων δεδομένων στον αρχικό χώρο χαρακτηριστικών και στη συνέχεια εφαρμόζει μια πολυωνυμική συνάρτηση σε αυτό το εσωτερικό γινόμενο. Ο πολυωνυμικός πυρήνας μπορεί να κατατάξει αποτελεσματικότερα μη γραμμικά δεδομένα χρησιμοποιώντας την παράμετρο βαθμού (degree). Η παράμετρος αυτή επιτρέπει στο SVM να κατασκευάζει όρια αποφάσεων που είναι πολυωνυμικής μορφής.

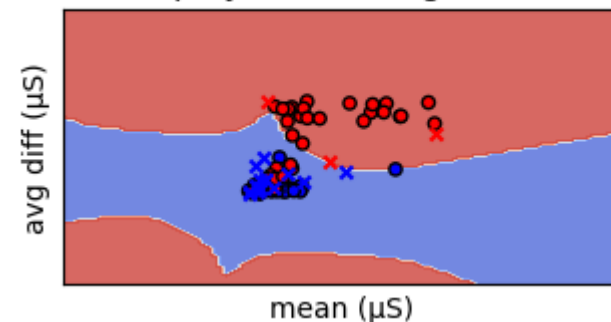
Παρακάτω στο [Σχήμα 18](#) και στο [Σχήμα 19](#) παρουσιάζεται η κατηγοριοποίηση με χρήση του παραπάνω τύπου πυρήνα και με τιμή βαθμού 5 και 10 αντίστοιχα.

SVM with polynomial (degree 5) kernel



Σχήμα 18: Κατηγοριοποίηση SVM με χρήση πολυωνυμικού πυρήνα βαθμού 5.

SVM with polynomial (degree 10) kernel



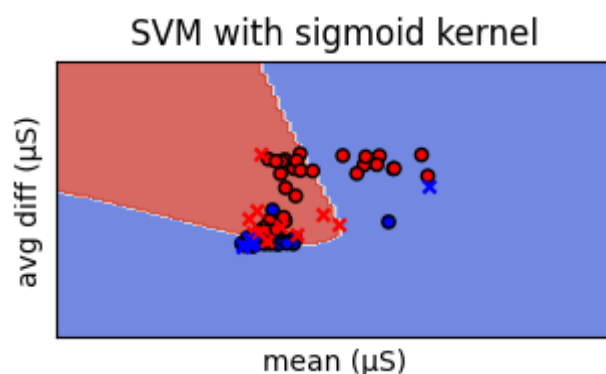
Σχήμα 19: Κατηγοριοποίηση SVM με χρήση πολυωνυμικού πυρήνα βαθμού 10.

Μια υψηλότερη τιμή βαθμού (degree) παράγει πιο περίπλοκο όριο το οποίο είναι πιθανό να αυξήσει την ακρίβεια του μοντέλου ωστόσο αυξάνεται ταυτόχρονα και ο κίνδυνος υπερπροσαρμογής (overfitting) το οποίο θα έχει ως επακόλουθο μείωση της ικανότητας του μοντέλου να κατατάσσει στην σωστή κλάση νέα δείγματα.

6.2.4. Κατηγοριοποίηση με χρήση πυρήνα σιγμοειδούς (sigmoid kernel)

Η επόμενη περίπτωση αφορά την χρήση πυρήνα σιγμοειδούς (sigmoid kernel) για τον διαχωρισμό των δεδομένων σε κλάσεις. Λειτουργεί με την εφαρμογή μιας σιγμοειδούς συνάρτησης στο βαθμωτό γινόμενο των ζευγών σημείων δεδομένων στον αρχικό χώρο χαρακτηριστικών. Η σιγμοειδής συνάρτηση μετατρέπει την είσοδο σε μια περιοχή μεταξύ 0 και 1, δημιουργώντας ένα όριο απόφασης που μπορεί να συλλάβει μη γραμμικές σχέσεις στα δεδομένα.

Τα αποτελέσματα της χρήσης αυτού του πυρήνα φαίνονται παρακάτω στο [Σχήμα 20](#).



Σχήμα 20: Κατηγοριοποίηση SVM με χρήση σιγμοειδούς πυρήνα (Sigmoid Kernel)

Μπορούμε να δούμε ότι το όριο απόφασης που λαμβάνεται με τον σιγμοειδή πυρήνα εμφανίζονται ακανόνιστα. Το όριο απόφασης προσπαθεί να διαχωρίσει τις κλάσεις τοποθετώντας μια καμπύλη σχήματος σιγμοειδούς, με αποτέλεσμα ένα σύνθετο όριο που μπορεί να μην γενικεύεται καλά σε αόρατα δεδομένα. Από αυτό το παράδειγμα γίνεται προφανές ότι ο σιγμοειδής πυρήνας έχει πολύ συγκεκριμένες περιπτώσεις χρήσης, όταν

ασχολείται με δεδομένα που η κατανομή τους στον χώρο δεδομένων εμφανίζει σιγμοειδές σχήμα.

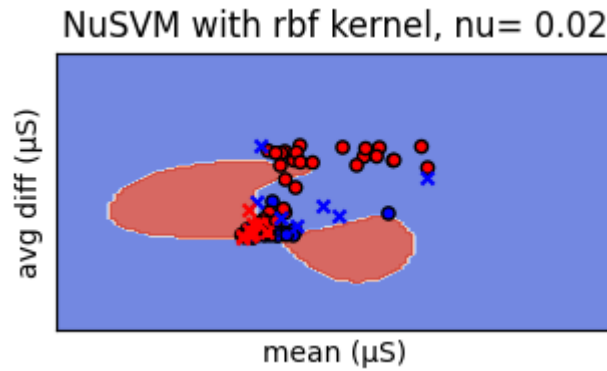
Για τις επόμενες τρεις περιπτώσεις που μελετήθηκαν χρησιμοποιήθηκε το μοντέλο NuSVC της Python. Το NuSVC εισάγει μια νέα παράμετρο (παράμετρος ν , ν parameter) η οποία λαμβάνει τιμές από 0 έως 1. Υψηλότερες τιμές της παραμέτρου οδηγούν στην κατασκευή ορίου απόφασης όπου αφενός επιτρέπεται να υπάρχουν όλο και λιγότερα δεδομένα στον χώρο του περιθωρίου (margin) και αφετέρου μειώνεται ο αριθμός των διανυσμάτων υποστήριξης.

Υπενθυμίζεται ότι διανύσματα υποστήριξης αποτελούν τα δείγματα που είναι πλησιέστερα στο υπερεπίπεδο (hyperplane) που κατασκευάζει ως όριο απόφασης το μοντέλο.

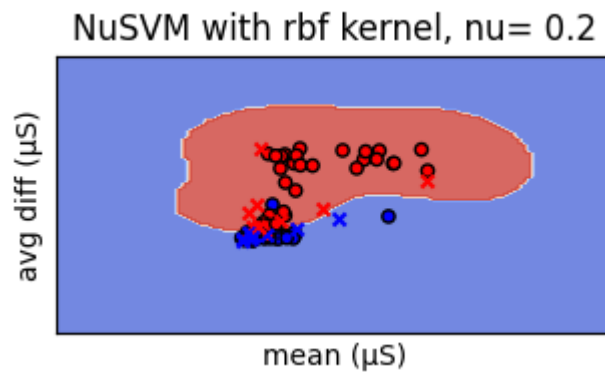
Όταν η τιμή της παραμέτρου μειώνεται τότε το μοντέλο ανέχεται μεγαλύτερο αριθμό λανθασμένων κατηγοριοποιήσεων κάτι που οδηγεί σε αυξημένο αριθμό διανυσμάτων υποστήριξης ταυτόχρονα επιτρέπει την ύπαρξη περισσότερων δεδομένων στον χώρο του περιθωρίου. Αυτό έχει ως αποτέλεσμα ότι το όριο απόφασης μπορεί να γίνει λιγότερο ακριβές και αυξάνεται ο κίνδυνος υπερπροσαρμογής (overfitting). Πρέπει να τονιστεί ότι υπάρχει περίπτωση το μοντέλο να μην μπορεί να κατατάξει τα δεδομένα με βάση την τιμή της παραμέτρου που ζητήθηκε και ως εκ τούτου η διαδικασία της κατάταξης οδηγείται σε αποτυχία.

6.2.5. NuSVM με πυρήνα συνάρτησης ακτινικής βάσης (RBF Kernel)

Ως πρώτη περίπτωση στο NuSVM μελετάται η χρήση πυρήνα συνάρτησης ακτινικής βάσης (RBF kernel) με τιμές της παραμέτρου ν 0.02 και 0.2. Τα αποτελέσματα εμφανίζονται παρακάτω στο [Σχήμα 21](#) και στο [Σχήμα 22](#).



Σχήμα 21: Κατηγοριοποίηση NuSVM με χρήση πυρήνα συνάρτησης ακτινικής βάσης και τιμή παραμέτρου nu ίση με 0.02.



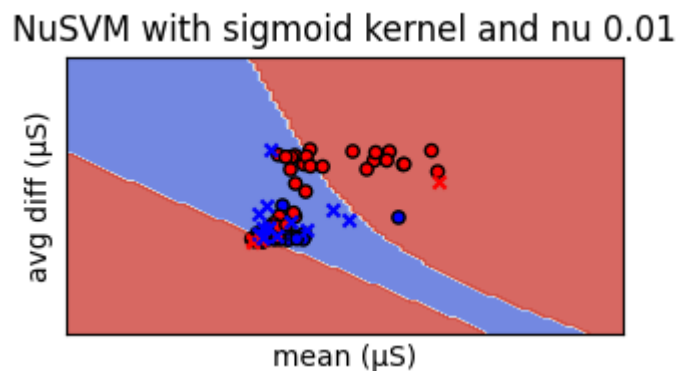
Σχήμα 22: Κατηγοριοποίηση NuSVM με χρήση πυρήνα συνάρτησης ακτινικής βάσης και τιμή παραμέτρου nu ίση με 0.2.

Από τα παραπάνω αποτελέσματα φαίνεται η επίδραση της τιμής της παραμέτρου nu στην κατηγοριοποίηση των αποτελεσμάτων. Στην μικρότερη τιμή nu (0.02) παρουσιάζονται περισσότερα δείγματα εκπαίδευσης που αφορούν τους αγχωμένους (κόκκινοι κύκλοι) σε περιοχή όπου το NuSVM έχει ορίσει ότι θα ανήκουν τα δείγματα των μη αγχωμένων (μπλε περιοχή). Αντίθετα, στην δεύτερη περίπτωση όπου η τιμή του nu είναι μεγαλύτερη παρατηρείται ότι τα περισσότερα δείγματα έχουν ανατεθεί στις περιοχές όπου είναι αναμενόμενο. Συμπερασματικά, επιβεβαιώνεται η θεωρητική άποψη ότι μείωση της τιμής της παραμέτρου nu οδηγεί σε λιγότερο ακριβές όριο απόφασης και ότι επιτρέπει την ανάθεση σε λάθος κλάση περισσότερων σημείων δοκιμής.

6.2.6. NuSVM με πυρήνα σιγμοειδούς (Sigmoid Kernel)

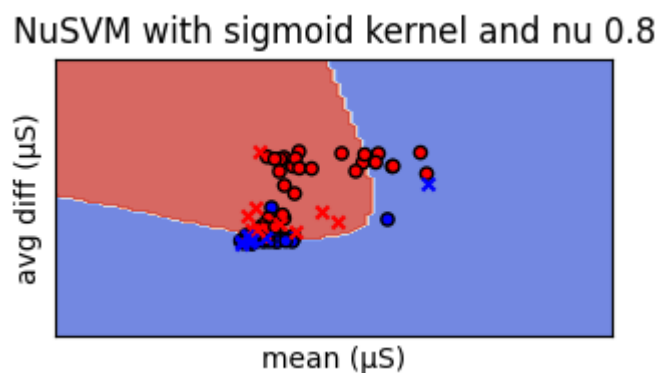
Στα επόμενα πειράματα μελετάται η χρήση NuSVM σε συνδυασμό με τον πυρήνα σιγμοειδούς και δύο διαφορετικές τιμές της παραμέτρου ν .

Στο [Σχήμα 23](#) φαίνεται η κατηγοριοποίηση με χρήση πυρήνα σιγμοειδούς και τιμή ν ίση με 0.01.



Σχήμα 23: Κατηγοριοποίηση NuSVM με χρήση πυρήνα σιγμοειδούς και τιμή ν ίση με 0.01

Στο [Σχήμα 24](#) φαίνεται η κατηγοριοποίηση με χρήση πυρήνα σιγμοειδούς και τιμή ν ίση με 0.8.



Σχήμα 24: Κατηγοριοποίηση NuSVM με χρήση πυρήνα σιγμοειδούς και τιμή ν ίση με 0.8

Γίνεται και εδώ αντιληπτή η συμπεριφορά που αναμένεται από την θεωρητική μελέτη για την επίδραση της τιμής της παραμέτρου ν στην κατηγοριοποίηση των δειγμάτων.

Για την τιμή της παραμέτρου 0.01 παρατηρείται ότι το όριο απόφασης είναι λιγότερο ακριβές και ότι υπάρχει μεγαλύτερη ανεκτικότητα στην παρουσία δειγμάτων εκπαίδευσης στην λάθος περιοχή του ορίου απόφασης.

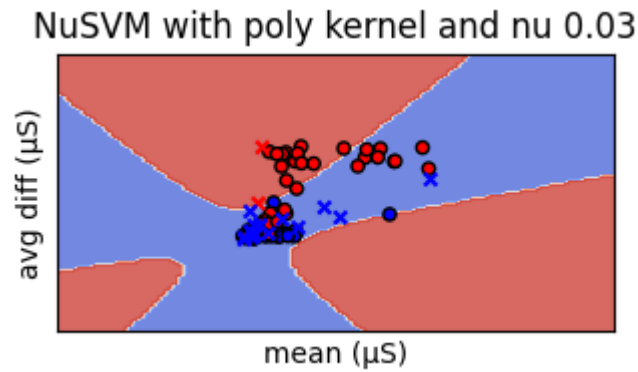
Επιπλέον, παρατηρείται ομοιότητα στο όριο απόφασης που προέκυψε σε προηγούμενη μέτρηση από την χρήση απλού SVM με χρήση σιγμοειδούς πυρήνα ([Σχήμα 20](#)). Ωστόσο, εντοπίζονται μικρές διαφορές οι οποίες έχουν ως αίτιο την εισαγωγή επίσης παραμέτρου ν η οποία θέτει περιορισμούς στον αριθμό των δειγμάτων που εντάσσονται σε λάθος κατηγορία και στον αριθμό των διανυσμάτων υποστήριξης.

Τέλος, στην παρούσα περίπτωση παρατηρήθηκε ότι για τιμή ν ίση με 1.0 το μοντέλο δεν μπορεί να κατηγοριοποιήσει τα δείγματα και αποτυγχάνει. Το αποτέλεσμα αυτό είναι επίσης αναμενόμενο καθώς η τιμή αυτή αποτελεί την μέγιστη τιμή που μπορεί να λάβει η τιμή ν κάτι το οποίο οδηγεί σε μηδενική ανοχή στην ύπαρξη δειγμάτων συνόλου εκπαίδευσης που έχουν ανατεθεί σε λανθασμένη κλάση. Ως εκ τούτου, η απόπειρα κατηγοριοποίησης έχει μεγάλο βαθμό δυσκολίας και αποτυγχάνει.

6.2.7. NuSVM με χρήση πολυωνυμικού πυρήνα (Polynomial Kernel)

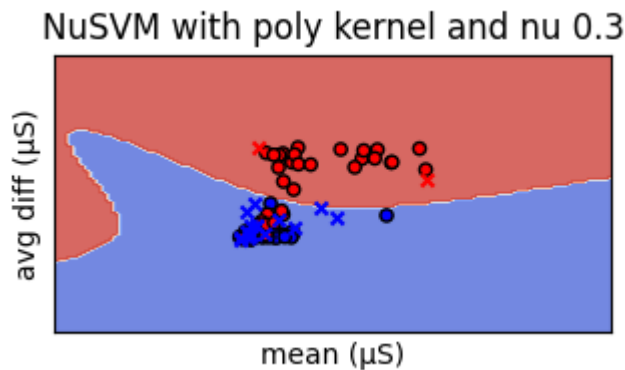
Τέλος, έγινε κατηγοριοποίηση με χρήση NuSVM σε συνδυασμό με πολυωνυμικό πυρήνα (Polynomial Kernel) και δύο διαφορετικές τιμές της παραμέτρου ν .

Στο [Σχήμα 25](#) παρουσιάζεται η κατηγοριοποίηση που προέκυψε από την χρήση του παραπάνω μοντέλου και με τιμή της παραμέτρου ν ίση με 0.03.



Σχήμα 25: Κατηγοριοποίηση με χρήση NuSVM σε συνδυασμό με πολυωνυμικό πυρήνα και τιμή παραμέτρου ν ίση με 0.03.

Στο [Σχήμα 26](#) παρουσιάζεται η κατηγοριοποίηση που προέκυψε από την χρήση του παραπάνω μοντέλου και με τιμή της παραμέτρου ν ίση με 0.3.



Σχήμα 26: Κατηγοριοποίηση με χρήση NuSVM σε συνδυασμό με πολυωνυμικό πυρήνα και τιμή παραμέτρου ν ίση με 0.3.

Όπως και στα προηγούμενα παραδείγματα με το μοντέλο NuSVM γίνεται αντιληπτή η επίδραση της παραμέτρου ν στην κατηγοριοποίηση των δειγμάτων. Για την μικρότερη τιμή της παραμέτρου ν παρατηρείται εκ νέου την μεγαλύτερη ανοχή του μοντέλου στην ύπαρξη δειγμάτων εκπαίδευσης σε διαφορετική κλάση από αυτή που θα έπρεπε να ανήκουν. Επίσης είναι ορατή η ομοιότητα με την κατηγοριοποίηση από το SVM σε συνδυασμό με πολυωνυμικό πυρήνα χωρίς την χρήση της παραμέτρου ν ([Σχήμα 18](#), [Σχήμα 19](#)).

Στη συνέχεια, στο [Κεφάλαιο 7](#) θα παρουσιαστούν οι επιδόσεις στην κατάταξη των αποτελεσμάτων των διάφορων μοντέλων SVM που παρουσιάστηκαν στο παρόν κεφάλαιο.

7. Αξιολόγηση Μοντέλων

Στο παρόν κεφάλαιο θα παρουσιαστούν οι επιδόσεις ακρίβειας των μοντέλων. Ως δείγμα δοκιμών επιλέχθηκαν τα δεδομένα από τις μετρήσεις GSR ατόμων που υποβλήθηκαν σε διαγνωστικό έλεγχο για Covid-19.

Για την μέτρηση της ακρίβειας χρησιμοποιήθηκαν η βαθμολογία ακριβείας και η βαθμολογία F1.

- Η **βαθμολογία ακριβείας** συγκρίνει τους αναμενόμενους χαρακτηρισμούς των δειγμάτων με την κατηγοριοποίηση που προέκυψε από το SVM και το αποτέλεσμα εκφράζεται ως το ποσοστό ομοιότητας αναμενόμενων τιμών και τιμών που προέκυψαν από την κατηγοριοποίηση.
- Η **βαθμολογία F1** χρησιμοποιεί τα μεγέθη ακρίβεια (Precision) και ανάκλησης (Recall). Κατόπιν υπολογίζει τον αρμονικό μέσο των δύο αυτών μεγεθών. Μπορεί να χρησιμοποιηθεί μόνο σε περίπτωση κατηγοριοποίησης σε δύο κλάσεις. Σε αυτήν την περίπτωση οι κλάσεις χωρίζονται στην θετική (positive) και την αρνητική (negative). Στην παρούσα διπλωματική ως θετική κλάση ορίστηκαν οι αγχωμένοι εξεταζόμενοι και ως αρνητική οι μη αγχωμένοι. Τα μεγέθη ακρίβεια (Precision) και ανάκληση (Recall) καθώς και η βαθμολογία F1 υπολογίζονται από τους παρακάτω τύπους.

$$Precision (P) = \frac{TP}{TP + FP} \quad (1)$$

$$Recall (R) = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = \frac{2 * P * R}{P + R} \quad (3)$$

όπου:

TP ο αριθμός των δειγμάτων που ορθώς ανατέθηκαν στην θετική κλάση (True Positive),

FP ο αριθμός των δειγμάτων που λανθασμένα ανατέθηκαν στην θετική κλάση (False Positive),

FN ο αριθμός των δειγμάτων που λανθασμένα ανατέθηκαν στην αρνητική κλάση (False Negative)

Η ακρίβεια (Precision) υποδεικνύει την ικανότητα του μοντέλου να κάνει σωστά θετικές προβλέψεις, χωρίς να ταξινομεί ψευδώς τις αρνητικές περιπτώσεις ως θετικές. Με άλλα λόγια, μετρά την ακρίβεια των θετικών προβλέψεων.

Η ανάκληση (Recall) υποδεικνύει την ικανότητα του μοντέλου να καταγράφει όλες τις θετικές περιπτώσεις που υπάρχουν στο σύνολο δεδομένων, χωρίς να λείπει κάποια. Με άλλα λόγια, μετρά την πληρότητα των θετικών προβλέψεων.

Η βαθμολογία F1 είναι προτιμότερο να χρησιμοποιείται όταν δεν έχουμε παρόμοιο αριθμό δειγμάτων ανά κλάσεις (μη ισορροπία κλάσεων) και όταν έχει μεγάλη σημασία το να ελαχιστοποιηθεί ο αριθμός των λανθασμένων κατατάξεων στην αρνητική κλάση (false negatives, ψευδώς αρνητικά δείγματα) σε αντίθετη περίπτωση συνίσταται η χρήση της βαθμολογίας ακρίβειας.

7.1. Κατηγοριοποιήσεις με τα μοντέλα SVM

7.1.1. Απλό γραμμικό SVM

Στην πρώτη δοκιμή κατάταξης με το γραμμικό SVM το ποσοστό των δειγμάτων του συνόλου δοκιμής που κατηγοριοποιήθηκαν ως αγχωμένοι είναι 11,11 % ενώ το ποσοστό των μη αγχωμένων είναι 88,89%.

Η βαθμολογία ακρίβειας είναι 40% ενώ η βαθμολογία F1 είναι 25%. Η σημαντική αυτή διαφορά μεταξύ των δύο βαθμολογιών είναι ένδειξη ότι το μοντέλο με την χρήση αυτού του πυρήνα είναι σε κατάσταση υπερπροσαρμογής (overfitting). Ακόμα μια πιθανή αιτία για την διαφορά των δύο τιμών είναι το ότι το μοντέλο ευνοεί την μια κλάση (μη αγχωμένοι) έναντι της άλλης.

7.1.2. Εξειδικευμένο Γραμμικό SVM (LinearSVM)

Με το LinearSVM τα δείγματα που κατηγοριοποιήθηκαν ως αγχωμένοι είναι 27,78% και οι μη αγχωμένοι 72,22%.

Η βαθμολογία ακριβείας είναι 50% ενώ η βαθμολογία F1 είναι 44,44%. Οι τιμές των δυο βαθμολογιών είναι περίπου ίσες οπότε το συμπέρασμα είναι ότι δεν υπάρχουν πολλά ψευδώς αρνητικά δείγματα και ότι το μοντέλο έχει καλή γενική επίδοση.

7.1.3. Κατηγοριοποίηση με χρήση Πυρήνα Συνάρτησης Ακτινικής Βάσης (RBF kernel)

Στην περίπτωση συνάρτησης ακτινικής βάσης οι μετρήσεις που κατηγοριοποιήθηκαν ως αγχωμένοι είναι 27,78% και οι μη αγχωμένοι είναι 72,22%.

Η βαθμολογία ακρίβειας είναι 50% και η βαθμολογία F1 είναι 44,44%. Όπως και στο προηγούμενο παράδειγμα οι τιμές των δύο βαθμολογιών είναι

παραπλήσιες οπότε το συμπέρασμα είναι ότι το μοντέλο πετυχαίνει καλή κατηγοριοποίηση και δεν υπάρχει ζήτημα υπερπροσαρμογής (overfitting).

7.1.4. Κατηγοριοποίηση με χρήση πολυωνυμικού πυρήνα βαθμού 5 και βαθμού 10

Στην περίπτωση χρήσης πολυωνυμικού πυρήνα τόσο με βαθμό 5 όσο και με βαθμό 10, 16,67% των δειγμάτων κατηγοριοποιήθηκαν ως αγχωμένοι ενώ 83,33% ως μη αγχωμένοι.

Η βαθμολογία ακρίβειας είναι 40% ενώ η βαθμολογία F1 είναι 25% και για τις δύο περιπτώσεις.

Παρά την αλλαγή της τιμής του βαθμού, δεν παρατηρείται αλλαγή στα ποσοστά των δύο βαθμολογιών. Από το παραπάνω φαινόμενο γίνεται αντιληπτό ότι η αναπροσαρμογή του ορίου λόγω αλλαγής του βαθμού δεν συνεπάγεται αυτόματα και αύξηση της ακρίβειας. Αυτό οφείλεται στην κατανομή των δειγμάτων δοκιμών στον χώρο χαρακτηριστικών. Τα παραπάνω επιβεβαιώνονται και από το [Σχήμα 18](#) και το [Σχήμα 19](#) παρατηρείται ότι δεν άλλαξε η κατηγοριοποίηση των δειγμάτων αλλά το όριο απόφασης μεταβλήθηκε. Κατά συνέπεια, αν υπήρχαν δείγματα του συνόλου δοκιμής σε περιοχές όπου στην πρώτη περίπτωση (βαθμός 5) ανήκαν στην μη αγχωμένη πλευρά του ορίου απόφασης (μπλε περιοχή) και στην δεύτερη περίπτωση (βαθμός 10) ανήκαν στην πλευρά του ορίου απόφασης όπου κατατάσσονται οι αγχωμένοι (κόκκινη περιοχή), γίνεται αντιληπτό ότι η ύπαρξη δειγμάτων σε τέτοιες περιοχές θα μετέβαλλε και τις δύο βαθμολογίες (επίδοσης και F1).

7.1.5. Κατηγοριοποίηση με χρήση πυρήνα σιγμοειδούς

Στην περίπτωση της κατηγοριοποίησης με χρήση πυρήνα σιγμοειδούς το ποσοστό των αγχωμένων δειγμάτων του συνόλου δοκιμών ήταν 61,11% και οι μη αγχωμένοι 38,89%.

Η βαθμολογία ακρίβειας είναι 70% και η βαθμολογία F1 είναι 72,73%. Οι συγκεκριμένες τιμές είναι πολύ ικανοποιητικές. Δεν παρουσιάζουν μεγάλη διαφορά μεταξύ τους κάτι που σημαίνει ότι το μοντέλο έχει πετύχει μια καλή ισορροπία μεταξύ του να αναγνωρίζει αποτελεσματικά τις σωστές κλάσεις για τα δεδομένα ενώ ελαχιστοποιεί τα ψευδώς θετικά και τα ψευδώς αρνητικά, οδηγώντας σε αξιόπιστες και ακριβείς προβλέψεις.

7.1.6. Κατηγοριοποίηση με χρήση NuSVM, πυρήνα συνάρτησης ακτινικής βάσης (RBF Kernel) και $\nu = 0.02$

Με χρήση NuSVM, πυρήνα RBF και τιμή $\nu = 0.02$ το 61,11% των δειγμάτων κατηγοριοποιήθηκαν ως αγχωμένοι ενώ το 38,89% ως μη αγχωμένοι.

Επετεύχθη βαθμολογία ακρίβειας 60% και βαθμολογία F1 75%. Οι παραπάνω τιμές κρίνονται ως πολύ ικανοποιητικές. Η βαθμολογία F1 είναι υψηλότερη από την βαθμολογία ακρίβειας κάτι που σημαίνει ότι το μοντέλο SVM επιτυγχάνει υψηλή ακρίβεια (λίγα ψευδώς θετικά) ενώ θυσιάζει κάποια ανάκληση (επιτρέποντας περισσότερα ψευδώς αρνητικά). Αυτό θα μπορούσε να συμβεί όταν το μοντέλο είναι συντηρητικό στην πρόβλεψη θετικών περιπτώσεων, οδηγώντας σε μικρότερο αριθμό ψευδώς θετικών, αλλά ενδεχομένως να λείπουν κάποια πραγματικά θετικά.

7.1.7. Κατηγοριοποίηση με χρήση NuSVM, πυρήνα συνάρτησης ακτινικής βάσης (RBF Kernel) και $\nu = 0.2$

Με χρήση NuSVM πυρήνα RBF και τιμή $\nu=0.2$ το 50% των δειγμάτων κατηγοριοποιείται στην κατηγορία των αγχωμένων ενώ το υπόλοιπο 50% ως μη αγχωμένοι.

Η βαθμολογία ακρίβειας είναι 70% και βαθμολογία F1 72,73%. Όπως και προηγουμένως τιμές των βαθμολογιών κρίνονται ως πολύ ικανοποιητικές. Επίσης, η διαφορά στις δυο τιμές είναι πολύ μικρή. Συμπερασματικά το

μοντέλο λειτουργεί με πολύ ικανοποιητικό τρόπο και υπάρχει ισορροπία μεταξύ ακρίβειας και ανάκλησης.

7.1.8. Κατηγοριοποίηση με χρήση NuSVM, πυρήνα σιγμοειδούς (Sigmoid Kernel) και $\nu = 0.01$

Με τον συνδυασμό NuSVM, πυρήνα σιγμοειδούς και τιμή παραμέτρου ν 0.01 το 22,22% των δειγμάτων κατηγοριοποιείται ως αγχωμένοι ενώ το 77,78% ως μη αγχωμένοι.

Οι βαθμολογίες ακρίβειας και F1 είναι 30% και 36,36% αντίστοιχα. Οι διαφορές στις δύο τιμές είναι μικρή κάτι που σημαίνει ότι το μοντέλο εξισορροπεί την ακρίβεια και την ανάκληση οπότε και εντοπίζει τα θετικά δείγματα χωρίς να χάνει αρκετά από αυτά. Εντούτοις οι τιμές είναι χαμηλές οπότε συνολικά ο συγκεκριμένος συνδυασμός δεν πετυχαίνει καλή κατηγοριοποίηση.

7.1.9. Κατηγοριοποίηση με χρήση NuSVM, πυρήνα σιγμοειδούς (Sigmoid Kernel) και $\nu = 0.8$

Με τον συνδυασμό NuSVM, πυρήνα σιγμοειδούς και τιμή παραμέτρου ν 0.8 το 55,56% των δειγμάτων κατηγοριοποιείται ως αγχωμένοι ενώ το 44,44% ως μη αγχωμένοι.

Οι βαθμολογίες ακρίβειας και F1 έχουν αμφότερες την ίδια τιμή η οποία είναι 60%. Η τιμή κρίνεται ως σχετικά ικανοποιητική. Επιπλέον, το ότι οι δύο τιμές είναι ίσες σημαίνει ότι υπάρχει απόλυτη ισορροπία μεταξύ ακρίβειας και ανάκλησης οπότε το μοντέλο μπορεί να ανιχνεύει ικανοποιητικά της περιπτώσεις που ανήκουν στην θετική κλάση (κλάση των αγχωμένων).

7.1.10. Κατηγοριοποίηση με χρήση NuSVM, πολυωνυμικού πυρήνα (Polynomial Kernel) και $\nu = 0.03$

Στην περίπτωση χρήσης NuSVM, πολυωνυμικού πυρήνα και τιμής $\nu = 0,03$ οι αγχωμένοι αποτελούν το 11,11% του συνόλου δοκιμών και οι μη αγχωμένοι το 88,89%.

Η βαθμολογία ακρίβειας είναι 30% και η βαθμολογία F1 είναι 0%. Μια βαθμολογία ακρίβειας 30% υποδηλώνει κακή συνολική απόδοση, καθώς το μοντέλο ταξινόμησε σωστά μόνο το 30% των περιπτώσεων. Μια βαθμολογία F1 0% υποδεικνύει ότι το μοντέλο απέτυχε εντελώς να προσδιορίσει σωστά τις θετικές περιπτώσεις, με αποτέλεσμα είτε έλλειψη πραγματικών θετικών (ακρίβεια = 0) είτε έλλειψη ανάκλησης (ανάκληση = 0) ή και τα δύο. Η έλλειψη ανάκλησης, στο πλαίσιο των μοντέλων ταξινόμησης, σημαίνει ότι το μοντέλο αποτυγχάνει να συλλάβει όλες τις περιπτώσεις της θετικής κλάσης (αληθινά θετικά) που υπάρχουν στο σύνολο δεδομένων.

Συμπερασματικά, ο παραπάνω συνδυασμός δεν πετυχαίνει καλά αποτελέσματα κατηγοριοποίησης.

7.1.11. Κατηγοριοποίηση με χρήση NuSVM, πολυωνυμικού πυρήνα (Polynomial Kernel) και $\nu = 0.3$

Στην περίπτωση χρήσης NuSVM, πολυωνυμικού πυρήνα και τιμής $\nu = 0,3$ οι αγχωμένοι αποτελούν το 11,11% του συνόλου δοκιμών και οι μη αγχωμένοι το 88,89% όπως ακριβώς και στην προηγούμενη περίπτωση.

Η βαθμολογία ακρίβειας είναι 40% και η βαθμολογία F1 είναι 25%. Παρατηρείται ότι η βαθμολογία ακρίβειας είναι περίπου ίδια με πριν ενώ η βαθμολογία F1 αυξήθηκε σε έναν βαθμό. Αυτό σημαίνει ότι το μοντέλο για ακόμα μια φορά δεν έχει καλή συνολική απόδοση. Εντούτοις, στην παρούσα περίπτωση το μοντέλο καταφέρνει να κατηγοριοποιήσει σωστά κάποια δεδομένα που ανήκουν στην θετική κλάση (αγχωμένοι) οπότε και παρατηρείται αύξηση της βαθμολογίας F1 παρόλο που η τιμή της συνεχίζει να παραμένει χαμηλή.

Συμπερασματικά, και αυτή η περίπτωση χρήσης δεν πετυχαίνει ικανοποιητικά αποτελέσματα κατηγοριοποίησης.

Στο επόμενο κεφάλαιο ([Κεφάλαιο 8](#)) θα παρουσιαστούν τα γενικότερα συμπεράσματα της αξιολόγησης και θα αναφερθούν προτάσεις για μελλοντική έρευνα και επέκταση της παρούσας διπλωματικής εργασίας.

8. Συμπεράσματα και Προτάσεις για Μελλοντική Έρευνα

8.1. Γενικά Συμπεράσματα Αξιολόγησης

Η μεγαλύτερη τιμή βαθμολογίας ακρίβειας ήταν 70% και ανιχνεύθηκε στην κατηγοριοποίηση SVM με πυρήνα σιγμοειδούς (sigmoid kernel) και στην περίπτωση χρήσης NuSVM με χρήση πυρήνα συνάρτησης ακτινικής βάσης (RBF kernel) και τιμή παραμέτρου ν ίση με 0,2.

Η μεγαλύτερη τιμή βαθμολογίας F1 ήταν 75% και παρατηρήθηκε στην περίπτωση χρήσης NuSVM με χρήση πυρήνα συνάρτησης ακτινικής βάσης (RBF kernel) και τιμή παραμέτρου ν ίση με 0.02.

Αντίστοιχα, η χαμηλότερη τιμή βαθμολογίας ακρίβειας ήταν 30% και παρατηρήθηκε στην περίπτωση χρήσης NuSVM σε συνδυασμό με πυρήνα σιγμοειδούς (sigmoid kernel) και πολυωνυμικό πυρήνα (polynomial kernel) με τιμές παραμέτρου ν ίσες με 0,01 και 0,03 αντίστοιχα.

Η χαμηλότερη τιμή βαθμολογίας F1 ήταν 0% και παρατηρήθηκε κατά την χρήση NuSVM με χρήση πολυωνυμικού πυρήνα και τιμή παραμέτρου ν ίση με 0,03. Σε αυτήν την περίπτωση, είτε το μοντέλο ταξινόμησε σχεδόν όλες τις αρνητικές περιπτώσεις ως θετικές (ακρίβεια – precision = 0) ή το μοντέλο δεν κατάφερε να ανιχνεύσει τις περισσότερες θετικές περιπτώσεις (ύπαρξη άγχους) οπότε και η ανάκληση (recall) έχει τιμή 0.

Τέλος, στις περιπτώσεις κατηγοριοποίησης, μια υψηλή τιμή βαθμολογίας ακρίβειας συνοδευόταν από υψηλή τιμή βαθμολογίας F1 και αντίστοιχα στα μοντέλα όπου παρατηρούνταν χαμηλές τιμές βαθμολογίας ακρίβειας παρατηρούνταν και χαμηλές τιμές βαθμολογίας F1.

8.2. Προτάσεις για μελλοντική έρευνα

Ο συνδυασμός μετρήσεων GSR και αλγορίθμων μηχανικής μάθησης αποτελεί ένα πολλά υποσχόμενο πεδίο, ιδιαίτερα για την προώθηση της εξατομικευμένης υγειονομικής περίθαλψης. Με την πρόοδο της τεχνολογίας, οι καινοτόμες προσεγγίσεις που συνδυάζουν την μηχανική μάθηση με τα βιομετρικά δεδομένα όπως το GSR προσφέρουν την δυνατότητα ανίχνευσης στρες με υψηλή ακρίβεια και σε πραγματικό χρόνο. Για την υλοποίηση αυτών των εφαρμογών θα μπορούσε να χρησιμοποιηθεί μια ευρεία γκάμα τεχνικών μηχανικής μάθησης όπως η βαθιά μάθηση (deep learning), η εκμάθηση συνόλου (ensemble learning), και της ενισχυμένης μάθησης (reinforcement learning).

Οι παραπάνω τεχνικές έχουν την δυνατότητα να ανιχνεύουν και να αναλύουν πολύπλοκα μοτίβα στα δεδομένα μετρήσεων GSR και να συντελέσουν σε σημαντικό βαθμό στην εξαγωγή σημαντικών γνώσεων σχετικά με τα επίπεδα άγχους ενός ατόμου.

Επιπλέον μελλοντική έρευνα μπορεί να συνδυάσει και άλλα βιομετρικά σήματα όπως η μεταβλητότητα καρδιακού ρυθμού (Heart Rate Variability - HRV) και δεδομένα ηλεκτροεγκεφαλογράφηματος (EEG) για να ενισχύσει την αποτελεσματικότητα και την ακρίβεια των μοντέλων ανίχνευσης άγχους.

Τα παραπάνω σε συνδυασμό με την πρόοδο της τεχνολογίας σχεδιασμού και κατασκευής φορητών αισθητήρων και στους αλγορίθμους ανάλυσης δεδομένων από βιομετρικά σήματα μπορούν να διαδραματίσουν καθοριστικό ρόλο στην ανάπτυξη μη επεμβατικών, εύχρηστων και χαμηλού κόστους λύσεων παρακολούθησης των επιπέδων άγχους οι οποίες μπορούν να ενσωματωθούν με ομαλό τρόπο στην καθημερινή ζωή.

Συνεχίζοντας την τροχιά της έρευνας για το συνδυασμό αλγορίθμων Μηχανικής Μάθησης με μετρήσεις GSR για ανίχνευση στρες, οι μελλοντικές προσπάθειες θα μπορούσαν να επικεντρωθούν στη βελτίωση της ερμηνευσιμότητας και της επεξήγησης αυτών των μοντέλων.

Διασαφηνίζοντας τα υποκείμενα χαρακτηριστικά και τις διαδικασίες λήψης αποφάσεων των αλγορίθμων Μηχανικής Μάθησης, οι ερευνητές μπορούν να ενισχύσουν την εμπιστοσύνη και την αποδοχή αυτών των συστημάτων σε κλινικά και πρακτικά περιβάλλοντα. Επιπλέον, διαχρονικές μελέτες που ενσωματώνουν σύνολα δεδομένων μεγάλης κλίμακας και διαφορετικούς πληθυσμούς μπορούν να παρέχουν πολύτιμες γνώσεις σχετικά με τη γενίκευση και την αξιοπιστία των μοντέλων ανίχνευσης στρες σε διαφορετικά δημογραφικά στοιχεία και περιβάλλοντα.

Δεοντολογικά ζητήματα, όπως το απόρρητο δεδομένων και η αλγοριθμική μεροληψία, απαιτούν επίσης ιδιαίτερη προσοχή για να διασφαλιστεί η υπεύθυνη ανάπτυξη και εφαρμογή των τεχνολογιών ανίχνευσης στρες. Από τα παραπάνω γίνεται εύκολα κατανοητό ότι η μελλοντική έρευνα σε αυτόν τον τομέα έχει τη δυνατότητα να ενδυναμώσει τα άτομα με αξιόπιστες γνώσεις για τα επίπεδα άγχους τους, προωθώντας την προληπτική φροντίδα και καλλιεργώντας μια κουλτούρα ευημερίας στην κοινωνία.

Συμπερασματικά η συνέχιση της έρευνας σε αυτόν τον τομέα έχει την προοπτική να αλλάξει ριζικά τον τρόπο με τον οποίο διαχειρίζεται και κατανοείται το άγχος και έτσι να οδηγήσει σε πιο αποτελεσματικές παρεμβάσεις οι οποίες θα βελτιώσουν την ψυχική και σωματική υγεία οδηγώντας έτσι στην βελτίωση της ζωής των ανθρώπων κάτι το οποίο αποτελεί βασικό στόχο της ιατρικής επιστήμης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] EODY - “Ελονοσία και ταξίδι”. <https://eody.gov.gr> [online] Available at: https://eody.gov.gr/wp-content/uploads/2018/12/elonosia_taksidi.pdf [Accessed December 2018]
- [2] Shaw, R., Kim, Y. kyun, & Hua, J. (2020). Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia. *Progress in Disaster Science*, 6. <https://doi.org/10.1016/j.pdisas.2020.100090>
- [3] Our World In Data - <https://ourworldindata.org/> [online]. Available at: <https://ourworldindata.org/grapher/covid-19-testing-policy>
- [4] Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2), 156–166. <https://ieeexplore.ieee.org/document/1438384>
- [5] de Santos Sierra, A., Sánchez Ávila, C., Guerra Casanova, J., & Bailador Del Pozo, G. (2011). A stress-detection system based on physiological signals and fuzzy logic. *IEEE Transactions on Industrial Electronics*, 58(10), 4857–4865. <https://ieeexplore.ieee.org/document/5680639>
- [6] J. Andren and P. Funk, “A case-based approach using behavioral biometrics to determine a user’s stress level,” in *Proc. ICCBR Workshops*, 2005, pp. 9-17
- [7] Dinges, D. F., Venkataraman, S., McGlinchey, E. L., & Metaxas, D. N. (2007). Monitoring of facial stress during space flight: Optical computer recognition combining discriminative and generative methods. *Acta Astronautica*, 60(4–7), 341– 350. <https://doi.org/10.1016/j.actaastro.2006.09.003>

- [8] Jiang, M., & Wang, Z. (2009). A Method for Stress Detection Based on FCM Algorithm. 2009 2nd International Congress on Image and Signal Processing, 1–5. <https://doi.org/10.1109/CISP.2009.5304150>
- [9] “Where to place GSR electrodes”. Tobii Connect. June 26, 2023. Available at: https://connect.tobii.com/s/article/where-to-place-gsr-electrodes?language=en_US
- [10] Roth, W. T., Dawson, M. E., & Fillion, D. L. (2012). Publication recommendations for electrodermal measurements: Publication standards for EDA. *Psychophysiology*, 49(8), 1017–1034. <https://doi.org/10.1111/j.1469-8986.2012.01384.x>
- [11] Fowles, D. C., Christie, M. J., Edelberg, R., GRINGS, W. W., Lykken, D. T., & Venables, P. H. (1981). Publication Recommendations for Electrodermal Measurements. *Psychophysiology*, 18(3), 232–239. Available at: <https://doi.org/10.1111/j.1469-8986.1981.tb03024.x>
- [12] Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1), 80–91. Available at: <https://doi.org/10.1016/j.jneumeth.2010.04.028>
- [13] Patil, V.K., Pawar, V.R., Randive, S. *et al.* From face detection to emotion recognition on the framework of Raspberry pi and galvanic skin response sensor for visual and physiological biosignals. *Journal of Electrical Systems and Inf Technol* 10, 24 (2023). <https://doi.org/10.1186/s43067-023-00085-2>
- [14] Tucci, L., & Burns, E. (2023, September 15). *What is machine learning and how does it work? In-depth guide*. Enterprise AI. <https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML>
- [15] Zhou, Victor (2019-12-20). "[Machine Learning for Beginners: An Introduction to Neural Networks](#)". *Medium*. [Archived](#) from the original on 2022-03-09. Retrieved 2021-08-15.
- [16] *Machine learning, explained | MIT Sloan*. (2021, April 21). MIT Sloan. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>

- [17] Zhong, S., Zhang, K., Bagheri, M., Burken, J. G., Gu, A., Li, B., Ma, X., Marrone, B. L., Ren, Z. J., Schrier, J., Shi, W., Tan, H., Wang, T., Wang, X., Wong, B. M., Xiao, X., Yu, X., Zhu, J. J., & Zhang, H. (2021). Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environmental Science and Technology*, 55(19), 12741-12754. <https://doi.org/10.1021/acs.est.1c01339>
- [18] Hsieh, W. W. (2022). Evolution of machine learning in environmental science— A prespective. *Environmental Data Science*, 1. <https://doi.org/10.1017/eds.2022.2>
- [19] Fahd, K., Venkatraman, S., Miah, S. J., & Ahmed, K. (2021). Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature. *Education and Information Technologies*, 27(3), 3743–3775. <https://doi.org/10.1007/s10639-021-10741-7>
- [20] Omar, M. (2022). Application of Machine Learning (ML) to Address Cybersecurity Threats. In: Machine Learning for [Cybersecurity](#). *SpringerBriefs in Computer Science*. Springer, Cham. https://doi.org/10.1007/978-3-031-15893-3_1
- [21] Georgas, A., Georgas, K., & Hristoforou, E. (2023b). Advancements in SARS-CoV-2 Testing: Enhancing Accessibility through Machine Learning-Enhanced Biosensors. *Micromachines*, 14(8), 1518. <https://doi.org/10.3390/mi14081518>
- [22] Rahmani AM, Yousefpoor E, Yousefpoor MS, Mehmood Z, Haider A, Hosseinzadeh M, Ali Naqvi R. Machine Learning (ML) in Medicine: Review, Applications, and Challenges. *Mathematics*. 2021; 9(22):2970. <https://doi.org/10.3390/math9222970>
- [23] Reddy, B. M. (2023). Machine learning for drug discovery and manufacturing. In *Advanced technologies and societal change* (pp. 3–30). https://doi.org/10.1007/978-981-99-0377-1_1
- [24] Tripathi, N., Goshisht, M. K., Sahu, S. K., & Arora, C. (2021). Applications of artificial intelligence to drug design and discovery in the big data era: a

comprehensive review. *Molecular Diversity*, 25(3), 1643–1664.
<https://doi.org/10.1007/s11030-021-10237-z>

[25] Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferrán, E. A., Lee, G., Li, B., Madabhushi, A., Shah, P. K., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery*, 18(6), 463–477. <https://doi.org/10.1038/s41573-019-0024-5>

[26] Kolluri, S., Lin, J., Liu, R. *et al.* Machine Learning and Artificial Intelligence in Pharmaceutical Research and Development: a Review. *AAPS J* 24, 19 (2022). <https://doi.org/10.1208/s12248-021-00644-3>

[27] Miller, M.I., Shih, L.C. & Kolachalama, V.B. Machine Learning in Clinical Trials: A Primer with Applications to Neurology. *Neurotherapeutics* 20, 1066–1080 (2023). <https://doi.org/10.1007/s13311-023-01384-2>

[28] Ardabili SF, Mosavi A, Ghamisi P, Ferdinand F, Varkonyi-Koczy AR, Reuter U, Rabczuk T, Atkinson PM. COVID-19 Outbreak Prediction with Machine Learning. *Algorithms*. 2020; 13(10):249. <https://doi.org/10.3390/a13100249>

[29] Heidari, A., Jafari Navimipour, N., Unal, M. *et al.* Machine learning applications for COVID-19 outbreak management. *Neural Comput & Applic* 34, 15313–15348 (2022). <https://doi.org/10.1007/s00521-022-07424-w>

[30] Liu, Z., Zhang, J., Hou, Y., Zhang, X., Li, G., & Xiang, Y. (2021a). Machine learning for multimodal Electronic Health Records-based Research: Challenges and Perspectives. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2111.04898>

[31] Tufail, S., Riggs, H., Tariq, M., & Sarwat, A. I. (2023). Advancements and Challenges in Machine Learning: A comprehensive review of models, libraries, applications, and algorithms. *Electronics (Basel)*, 12(8), 1789. <https://doi.org/10.3390/electronics12081789>

[32] Basudhar, A., & Missoum, S. (2010). An improved adaptive sampling scheme for the construction of explicit boundaries. *Structural and Multidisciplinary Optimization (Print)*, 42(4), 517–529. <https://doi.org/10.1007/s00158-010-0511-0>

- [33] Sahu, S. K., M. B. N. V. Prasad, and B. K. Tripathy. "A support vector machine binary classification and image segmentation of remote sensing data of Chilika Lagloon." *International Journal of Research in Information Technology* 3.5 (2015): 191-204.
- [34] Kim, Eric. "Everything you wanted to know about the kernel trick." January 9, 2013. https://www.eric-kim.net/eric-kim-net/posts/1/kernel_trick.html
- [35] Ayata, D., Yaslan, Y., & Kamaşak, M. (2017). Emotion Recognition via Galvanic Skin Response: Comparison of Machine Learning Algorithms and Feature Extraction Methods. *IU-Journal of Electrical & Electronics Engineering*, 17(1), 3147-3156.
- [36] Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.
- [37] Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145–172. <https://doi.org/10.1037/0033-295x.110.1.145>
- [38] Lang, P., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biological Psychiatry*, 44(12), 1248–1263. [https://doi.org/10.1016/s0006-3223\(98\)00275-3](https://doi.org/10.1016/s0006-3223(98)00275-3)
- [39] Mehrabian, Albert (2007). *Nonverbal communication*. pp. 10–15. [ISBN 978-0-202-30966-8](https://doi.org/10.1016/0006-3223(98)00275-3)
- [40] Ratneshwar, S.; Glen, David (2003). *The why of consumption: contemporary perspectives on consumer motives*. p. 39. [ISBN 978-0-415-31617-0](https://doi.org/10.1016/0006-3223(98)00275-3).
- [41] Becker, Christian; Kopp, Stefan; et al. (2008). "Why Emotions should be Integrated into Conversational Agents". In Nishida, Toyoaki (ed.). *Conversational informatics: an engineering approach*. pp. 49–59, 61. [ISBN 978-0-4700-2699-1](https://doi.org/10.1016/0006-3223(98)00275-3).
- [42] Graziotin, Daniel & Wang, Xiaofeng & Abrahamsson, Pekka. (2015). Understanding the Affect of Developers: Theoretical Background and Guidelines for Psychoempirical Software Engineering. <https://doi.org/10.1145/2804381.2804386>.

- [43] Feidakis, M., Rangoussi, M., Kasnesis, P., Patrikakis, C. Z., Kogias, D. G., & Charitopoulos, A. (2019). Affective assessment in distance learning: a semi-explicit approach. *The International Journal of Technologies in Learning*, 26(1), 19–34. <https://doi.org/10.18848/2327-0144/cgp/v26i01/19-34>
- [44] S. Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," in *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, Jan.-March 2012, doi: 10.1109/T-AFFC.2011.15
- [45] Maverick, A. (2023, April 3). What is Empirical Mode Decomposition - AI Maverick - Medium. *Medium*. <https://samanemami.medium.com/what-is-empirical-mode-decomposition-3ec89115db6b>
- [46] Akram, Muhammad Usman; DAR, MUHAMMAD NAJAM; Rahim, Amna (2022), "Young Adult's Affective Data (YAAD) Using ECG and GSR Signals", *Mendeley Data*, V4, doi: 10.17632/g2p7vwxyn2.4
- [47] Smerdov, A., Zhou, B., Lukowicz, P., & Somov, A. (2020). Collection and validation of psychophysiological data from professional and amateur players: a multimodal esports dataset. *arXiv preprint arXiv:2011.00958*.