



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

Καταγραφή και Ανάλυση Βιομετρικών  
Δεικτών για Υποστήριξη Ασθενών σε  
Θέματα Ψυχικής Υγείας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

Μαρία Παναγιώτου

Επιβλέπων: Πέτρος Μαραγκός  
Καθηγητής Ε.Μ.Π.

ΕΡΓΑΣΤΗΡΙΟ ΟΡΑΣΗΣ ΥΠΟΛΟΓΙΣΤΩΝ, ΕΠΙΚΟΙΝΩΝΙΑΣ ΛΟΓΟΥ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑΣ  
ΣΗΜΑΤΩΝ  
Αθήνα, Φεβρουάριος 2022





Εθνικό Μετσόβιο Πολυτεχνείο  
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  
Τομέας Σημάτων, Ελέγχου και Ρομποτικής  
Εργαστήριο Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων

# Καταγραφή και Ανάλυση Βιομετρικών Δεικτών για Υποστήριξη Ασθενών σε Θέματα Ψυχικής Υγείας

## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

Μαρία Παναγιώτου

Επιβλέπων: Πέτρος Μαραγκός  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 28<sup>η</sup> Φεβρουαρίου, 2022.

.....  
Πέτρος Μαραγκός  
Καθηγητής Ε.Μ.Π.

.....  
Αθανασιος Ροντογιάννης  
Αναπληρωτής Καθηγητής Ε.Μ.Π.

.....  
Γεράσιμος Ποταμιανός  
Αναπληρωτής Καθηγητής  
Παν/μιο Θεσσαλίας

Αθήνα, Φεβρουάριος 2022

.....  
**ΜΑΡΙΑ ΠΑΝΑΓΙΩΤΟΥ**  
Διπλωματούχος Ηλεκτρολόγος Μηχανικός  
και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © – All rights reserved Μαρία Παναγιώτου, 2022.  
Με επιφύλαξη παντός δικαιώματος.

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

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



# Περίληψη

Οι διαταραχές του φάσματος της σχιζοφρένειας είναι χρόνιες παθήσεις που πάρα τις έρευνες που γίνονται τα τελευταία χρόνια οι αίτιες τους παραμένουν ασαφείς. Η χρήση φορητών συσκευών (wearables) και οι συλλογή των δεδομένων μέσω αυτών μας επιτρέπει την αξιόπιστη καταγραφή πληθώρας βιομετρικών δεικτών, οι οποίοι μας δίνουν τη δυνατότητα να μαζέψουμε μεγάλο όγκο δεδομένων και να τον αξιοποιήσουμε κατάλληλα για τη έγκαιρη πρόβλεψη ψυχωτικών διαταραχών.

Σε αυτή τη μελέτη εξετάζεται η εφαρμογή μοντέλων βαθιάς μάθησης για την ανίχνευση υποτροπών σε ασθενείς με ψυχωτικές διαταραχές. Συγκεκριμένα, χρησιμοποιούνται φυσιολογικά σήματα όπως η κίνηση των ασθενών και βιοσήματα όπως η παρακολούθηση του καρδιακού ρυθμού, που προέρχονται και τα δύο από έξυπνα ρολόγια (smartwatches). Τα δεδομένα μας προέρχονται από δέκα ασθενείς που πάσχουν με διαταραχές του φάσματος της σχιζοφρένειας και τα επεξεργαζόμαστε με σκοπό να εξάγουμε τις πιο κατάλληλες αναπαραστάσεις και χαρακτηριστικά. Κάθε επεισόδιο υποτροπής χαρακτηρίζεται και επισημαίνεται κατάλληλα από τους κλινικούς ιατρούς ως χαμηλό, μέτριο ή υψηλό.

Για την αντιμετώπιση του προβλήματος ανίχνευσης των υποτροπών, υλοποιούνται τέσσερις διαφορετικές αρχιτεκτονικές κωδικοποιητή - αποκωδικοποιητή, βασισμένες σε convolutional, fully connected, recurrent neural networks autoencoder και transformer, και εκπαιδεύονται είτε εξατομικευμένα για κάθε ασθενή είτε γενικά με βάση τα συνολικά δεδομένα των ασθενών. Τα μοντέλα μας έχουν ως απώτερο σκοπό να μάθουν τις συμπεριφορές των ασθενών και να μπορούν να ξεχωρίζουν μία φυσιολογική συμπεριφορά από τη συμπεριφορά της περιόδου υποτροπής.

Τα αποτελέσματα των διάφορων μοντέλων που υλοποιήθηκαν συγκρίνονται μεταξύ τους, με τα μοντέλα convolutional και fully connected neural network να έχουν τις καλύτερες αποδόσεις. Τέλος, για τους ασθενείς με περισσότερα από ένα επεισόδιο υποτροπής με διαφορετική σοβαρότητα υποτροπής, πειραματιστήκαμε για να δούμε πόσο σημαντική είναι οι κατηγορία της υποτροπής που έχει ο ασθενείς, για την ανίχνευση της.

**Λέξεις Κλειδιά** — Ψυχωτικές Διαταραχές, Ανίχνευση Ανωμαλιών, Autoencoders, Transformers, Convolutional Neural Networks, Fully Connected Neural Network, Recurrent Neural Network, Χρονοσειρές





# Abstract

Schizophrenia spectrum disorders are chronic diseases and despite extensive research in recent years, their causes remain unclear. The use of wearables and the collection of data through them, allows us to reliably record a variety of biometric indicators, which enable us to collect large volumes of data and use them appropriately for the timely prediction of psychotic disorders.

This study examines the application of deep learning models to detect relapses in patients with psychotic disorders. Specifically, physiological signals such as patients' movement and biosignals such as heart rate monitoring, both originated by smartwatches, are used. Our data comes from ten patients with schizophrenia spectrum disorders from which we extract features by processing them accordingly. Each relapse' episode is appropriately characterized by clinicians as low, moderate or high.

To face up with anomalies' detection problem, four different encoder-decoder architectures are implemented, based on convolutional, fully connected, recurrent neural networks autoencoder and transformer, and are trained either individually for each patient or generally based on the overall patients data. Our ultimate goal is to learn patients' behaviors and to be able to distinguish a normal behavior from the behavior of the relapse period.

The results of the various models implemented are compared each other, with the convolutional and fully connected neural network autoencoder models having the highest efficiencies. Finally, patients that have more than one relapse' episode with different severity level, an experiment was performed with the goal of detecting the importance of the severity level for the patient relapse.

**Keywords** — Psychotic Disorders, Anomaly Detection, Autoencoders, Transformers, Convolutional Neural Networks, Fully Connected Neural Network, Recurrent Neural Network, Time-series



# Ευχαριστίες

Καταρχάς, θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ. Πέτρο Μαραγκό, για την ευκαιρία που μου έδωσε και την εμπιστοσύνη που μου έδειξε με την ανάθεση αυτής της διπλωματικής εργασίας. Επίσης, θα ήθελα να ευχαριστήσω τη μεταδιδακτορική ερευνήτρια Δρ. Νάνσυ Ζλατίντση και τους διδακτορικούς φοιτητές Αργύρη Ρουμελιώτη και Παναγιώτη Φιλντίση για την συνεχή καθοδήγηση και άμεση στήριξη τους. Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου, όλους του συμφοιτητές και φίλους μου που με στήριξαν και με βοήθησαν καθ' όλη την διάρκεια των φοιτητικών μου χρόνων.

Παναγιώτου Μαρία  
Φεβρουάριος 2022



# Περιεχόμενα

Περιεχόμενα	xiii
Λίστα Σχημάτων	xv
Κατάλογος Πινάκων	xviii
<b>1 Εισαγωγή</b>	<b>1</b>
1.1 Η σημασία του προβλήματος	2
1.2 Αντιμετώπιση του προβλήματος	2
1.3 Διαφορετικές Προσεγγίσεις και Εφαρμογές	3
1.3.1 Επεξεργασία σήματος στη Ψυχική Υγεία	5
1.3.2 Σχετικές έρευνες με νευρωνικά δίκτυα	8
1.4 Κίνητρα και Συνεισφορές	13
1.5 Διάρθρωση της Διπλωματικής εργασίας	14
<b>2 Θεωρητικό Υπόβαθρο</b>	<b>15</b>
2.1 Επεξεργασίας Σήματος	15
2.1.1 Θεωρία Πιθανοτήτων και Στατιστικής	15
2.1.2 Γραμμική Ανάλυση (Linear Analysis)	16
2.1.3 Μη Γραμμική Ανάλυση (Non-Linear Analysis)	18
2.2 Μηχανική Μάθηση	22
2.3 Νευρωνικά Δίκτυα	23
2.4 Βαθιά Νευρωνικά Δίκτυα	25
2.4.1 Πλήρως Συνδεδεμένα Νευρωνικά Δίκτυα (FNN)	25
2.4.2 Συνελικτικά Νευρωνικά Δίκτυα (CNNs)	25
2.4.3 Sequence Modeling: Αναδρομικά Νευρωνικά Δίκτυα (RNN)	28
2.4.4 Transformer Network	34
2.4.5 Autoencoders	36
<b>3 Δεδομένα και Προεπεξεργασία</b>	<b>39</b>
3.1 Συλλογή Δεδομένων και Πειραματικό Πρωτόκολλο	39
3.2 Ανάλυση Δεδομένων	41
3.3 Προεπεξεργασία Δεδομένων	52
<b>4 Περιγραφή Μεθόδου</b>	<b>55</b>
4.1 Περιγραφή Μεθόδου	55

4.2	Αρχιτεκτονικές Μοντέλων . . . . .	56
4.2.1	Fully Connected Neural Network Autoencoder (FNN AE) . . . . .	56
4.2.2	Convolutional Neural Network Autoencoder (CNN AE) . . . . .	57
4.2.3	Recurrent Neural Network (RNN) . . . . .	58
4.2.4	Transformer . . . . .	58
4.3	Εκπαίδευση . . . . .	60
4.4	Υπολογισμός Anomaly score . . . . .	60
4.5	Ανίχνευση Ανωμαλιών . . . . .	62
<b>5</b>	<b>Πειράματα και Αξιολόγηση</b>	<b>63</b>
5.1	Μετρικές Αξιολόγησης . . . . .	63
5.2	Πειράματα και Αξιολόγηση . . . . .	66
5.2.1	Εξατομικευμένα Μοντέλα . . . . .	66
5.2.2	Global Μοντέλα . . . . .	76
5.2.3	Σοβαρότητα Υποτροπής . . . . .	80
5.2.4	Πειράματα με thresholding . . . . .	81
<b>6</b>	<b>Συμπεράσματα και Μελλοντικές Επεκτάσεις</b>	<b>83</b>
6.1	Σύνοψη και Συμπεράσματα . . . . .	83
6.2	Μελλοντικές Επεκτάσεις . . . . .	84
6.3	Υποβολή Συνεδριακού Άρθρου . . . . .	85
<b>A</b>	<b>Βιβλιογραφία</b>	<b>87</b>

# Λίστα Σχημάτων

1.3.1 Αρχιτεκτονική FNN AD [Adl+20]	11
1.3.2 Αρχιτεκτονική GRU sequence-to-sequence [Adl+20]	11
2.1.1 Παράδειγμα Poincaré plot που περιγράφει λεπτομερώς τη διαδικασία προσαρμογής της έλλειψης. Το σύστημα συντεταγμένων $x_1$ και $x_2$ είναι προσαρμοσμένος στις $45^\circ$ ως προς τον κανονικό άξονα. Η τυπική απόκλιση της απόστασης των σημείων από κάθε άξονα καθορίζει το πλάτος ( $SD1$ ) και το μήκος ( $SD2$ ) της έλλειψης [BPK01].	22
2.3.1 Απλό νευρωνικό δίκτυο [Ola15]	24
2.4.1 Πλήρως συνδεδεμένο νευρωνικό δίκτυο (FNN) [Ola15]	25
2.4.2 Αρχιτεκτονική Convolutional Neural Network [AM17]	26
2.4.3 Convolutional Layer [Yak+18]	27
2.4.4 Pooling Layer [RW17]	28
2.4.5 Απλό RNN [Ola15]	29
2.4.6 Αμφίδρομο αναδρομικό νευρωνικό δίκτυο (Bi-RNN) [Ola15]	30
2.4.7 Εσωτερική Δομή ενός LSTM layer. [Ola15]	30
2.4.8 Forget Gate [Ola15]	31
2.4.9 Input Gate [Ola15]	32
2.4.10 Output Gate [Ola15]	32
2.4.11 Αρχιτεκτονική GRU [Ola15]	34
2.4.12 Αρχιτεκτονική Transformer Encoder-Decoder [Vas+17]	35
2.4.13 Multi-head Self Attention [Vas+17]	36
2.4.14 Αρχιτεκτονική Autoencoder [Alo+19]	36
2.4.15 Αρχιτεκτονική Sequence-to-sequence Autoencoder [Kos19]	38
3.2.1 Παράδειγμα χαρακτηριστικών, που χρησιμοποιήθηκαν σε αυτή τη μελέτη, κατά τη διάρκεια της ημέρας ενός ασθενή.	43
3.2.2 Χαρακτηριστικό μέσης ενέργειας γραμμικής επιτάχυνσης του ασθενή #10. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γραμμικής επιτάχυνσης για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.	44
3.2.3 Χαρακτηριστικό νόρμα γωνιακής επιτάχυνσης του ασθενή #10. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γωνιακής επιτάχυνσης για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.	45
3.2.4 Χαρακτηριστικό μέσης τιμής καρδιακού ρυθμού του ασθενή #10. Στην πάνω γραφική παρουσιάζεται ο καρδιακός ρυθμός για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.	45

3.2.5 Χαρακτηριστικό του διαστήματος RR του καρδιακού ρυθμού του ασθενή #10. Στην πάνω γραφική παρουσιάζεται το διάστημα RR του καρδιακού ρυθμού για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής. . . . .	46
3.2.6 Χαρακτηριστικό SD1 της μεταβλητότητας του καρδιακού ρυθμού του ασθενή #10. Στην πάνω γραφική παρουσιάζεται το SD1 για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής. . . . .	46
3.2.7 Χαρακτηριστικό των κανονικοποιημένων ζώνων LF και HF της μεταβλητότητας του καρδιακού ρυθμού. Στην πάνω γραφική παρουσιάζονται οι κανονικοποιημένοι LF και HF του ασθενή #10 για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής. . . . .	47
3.2.8 Χαρακτηριστικό μέσης ενέργειας γραμμικής επιτάχυνσης του ασθενή #7. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γραμμικής επιτάχυνσης για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής. . . . .	48
3.2.9 Χαρακτηριστικό μέσης ενέργειας γωνιακής επιτάχυνσης του ασθενή #7. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γωνιακής επιτάχυνσης για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής. . . . .	48
3.2.10 Χαρακτηριστικό μέσης τιμής καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζεται ο καρδιακός ρυθμός για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής. . . . .	49
3.2.11 Χαρακτηριστικό του διαστήματος RR του καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζεται το διάστημα RR του καρδιακού ρυθμού για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής. . . . .	50
3.2.12 Χαρακτηριστικό SD1 της μεταβλητότητας του καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζεται το SD1 για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής. . . . .	50
3.2.13 Χαρακτηριστικό των κανονικοποιημένων ζωνών LF και HF της μεταβλητότητας του καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζονται οι κανονικοποιημένες ζώνες LF και HF για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής. . . . .	51
3.3.1 Συνολικές μέρες στο σύνολο δεδομένων για κάθε ασθενή πριν και μετά την προεπεξεργασία των δεδομένων. . . . .	53
4.1.1 Περιγραφή Μεθόδου . . . . .	55
4.2.1 Αρχιτεκτονική Fully Connected Neural Network Autoencoder . . . . .	57
4.2.2 Αρχιτεκτονική Convolutional Neural Network Autoencoder . . . . .	58
4.2.3 Αρχιτεκτονική GRU Seq2Seq Autoencoder with Attention . . . . .	59
4.2.4 Αρχιτεκτονική Transformer . . . . .	59



5.1.1 Παράδειγμα ROC curve από τον ασθενή #6 του FNN AE μοντέλου, με ROC AUC Score 0.75. Παρατηρούμε ότι όσο μεταβάλλονται προοδευτικά οι πιθανές οριακές τιμές διαχωρισμού, τόσο μεγαλύτερος είναι ο λόγος του TPR προς το FPR, κάτι το οποίο είναι αναμενόμενο, αφού όσο μεγαλύτερο είναι το threshold τόσο μεγαλύτερο είναι και το ποσοστό των σωστά ταξινομημένων σημείων. . . . .	64
5.1.2 Παράδειγμα Precision-Recall curve από το ασθενή #6 του FNN AE μοντέλου, με PR AUC Score 0.91. Παρατηρούμε ότι όσο αυξάνεται το recall, τόσο μειώνεται το precision και αντίστροφα. Με την καμπύλη αυτή, μπορούμε να οπτικοποιήσουμε τον τρόπο με τον οποίο η επιλογή του threshold επηρεάζει την απόδοση των μοντέλων μας και επίσης στο να επιλέξουμε το καλύτερο threshold. Στο συγκεκριμένο σχήμα παρατηρούμε ότι το FNN AE μοντέλο φτάνει σε Recall 0.8 χωρίς να προβλέψει ψευδής υποτροπές. . . . .	65
5.2.1 Γραφική παράσταση του Anomaly Score του ασθενή #1 στο μοντέλο CNN στο διάστημα 5 λεπτών. Το anomaly score δεξιά της διαχωριστικής γραμμής, είναι οι μέρες στις οποίες ο ασθενής είχε διαγνωστεί με ψυχωτική υποτροπή, ενώ το anomaly score αριστερά της διαχωριστικής γραμμής είναι οι μέρες όπου ο ασθενής δεν είχε κάποιο σύμπτωμα υποτροπής ('κανονικές' μέρες). Παρατηρούμε ότι τα anomaly scores στις μέρες υποτροπής είναι αρκετά υψηλότερα από τις κανονικές μέρες. Να σημειώσουμε ότι τα Points στον άξονα- $x$ είναι οι μέρες, οι οποίες δεν είναι συνεχόμενες, δηλαδή τα απεικονιζόμενα anomaly scores δεν είναι συνεχή. . . . .	68
5.2.2 Median PR AUC για τα μοντέλα FNN AE, CNN AE, Transformer και GRU AE στα διαφορετικά granularities (5 λεπτών, 10 λεπτών και 1 ώρας). . . . .	71
5.2.3 Median ROC AUC για τα μοντέλα FNN AE, CNN AE, Transformer και GRU AE στα διαφορετικά granularities (5 λεπτών, 10 λεπτών και 1 ώρας). . . . .	71
5.2.4 Σφάλματα Ανακατασκευής (median) για τα εξατομικευμένα μοντέλα με δεδομένα διαστήματος διάρκειας 5 λεπτών. . . . .	73
5.2.5 Median PR AUC score για τα εξατομικευμένα μοντέλα μας, χρησιμοποιώντας τα Raw δεδομένα (στήλη με σκούρο μπλε χρώμα) και τα δεδομένα με τα 10 χαρακτηριστικά (γαλάζια στήλη). . . . .	75
5.2.6 Median ROC AUC score για τα εξατομικευμένα μοντέλα μας, χρησιμοποιώντας τα Raw δεδομένα (στήλη με σκούρο μπλε χρώμα) και τα δεδομένα με τα 10 χαρακτηριστικά (γαλάζια στήλη). . . . .	75
5.2.7 PR AUC για το global μοντέλο FNN το οποίο όμως αξιολογείται μεμονωμένα ο κάθε ασθενή και για το εξολοκλήρου εξατομικευμένο μοντέλο CNN. . . . .	78
5.2.8 ROC AUC για το global μοντέλο FNN το οποίο όμως αξιολογείται μεμονωμένα ο κάθε ασθενή και για το εξολοκλήρου εξατομικευμένο μοντέλο CNN. . . . .	79
5.2.9 Γραφική παράσταση του Anomaly Score του ασθενή #2 στο μοντέλο CNN στο διάστημα 5 λεπτών. Το anomaly score δεξιά της διαχωριστικής γραμμής, είναι οι μέρες στις οποίες ο ασθενής είχε διαγνωστεί με ψυχωτική υποτροπή, ενώ το anomaly score αριστερά της διαχωριστικής γραμμής είναι οι μέρες όπου ο ασθενής δεν είχε κάποιο σύμπτωμα υποτροπής ('κανονικές' μέρες). Παρατηρούμε ότι οι μέρες ψυχωτικής υποτροπής για τις οποίες έχουμε δεδομένα είναι μόνο 2, στις οποίες δεν βλέπουμε το αναμενόμενο, δηλαδή υψηλότερα anomaly scores στις μέρες υποτροπής από τις κανονικές μέρες. Να σημειώσουμε ότι τα Points στον άξονα- $x$ είναι οι μέρες, οι οποίες δεν είναι συνεχόμενες, δηλαδή τα απεικονιζόμενα anomaly scores δεν είναι συνεχή. . . . .	79



# Κατάλογος Πινάκων

3.1	Δημογραφικά στοιχεία για τους ασθενείς με Σχιζοφρένεια (SZ), διπολική διαταραχή I (BP), Schizoaffective disorder, Brief Psychotic Episode και Schizophreniform disorder. Ο πίνακας περιλαμβάνει στοιχεία μόνο για τους 10 από τους 24 ασθενείς, αφού αυτά είναι και τα δεδομένα που χρησιμοποιήθηκαν για την συγκεκριμένη εργασία. . . . .	40
3.2	Υποτροπές ασθενών ανά κατηγορία υποτροπής σε μέρες σύμφωνα με τις επισημειώσεις των ιατρών. . . . .	41
3.3	Αρχικό Dataset. Πίνακας με τις μέρες στο Dataset ανά κατηγορία πριν τη προεπεξεργασία δεδομένων. . . . .	41
3.4	Πίνακας με τις μέρες στο Dataset ανά κατηγορία υποτροπής μετά τη προεπεξεργασία δεδομένων. Με έντονη γραφή επισημειώνονται οι ασθενείς οι οποίοι έχουν υποστεί περισσότερες από μία ψυχωτικές υποτροπές, με διαφορετική σοβαρότητα υποτροπής. Τα δεδομένα των ασθενών #5 και #6 από την περίοδο που είχαν υποστεί relapse χαμηλής σοβαρότητας, αποτελούνται μόνο από 2 και 12 ώρες, αντίστοιχα . . . . .	53
4.1	Οι παράμετροι για τα μοντέλα GRU Seq2Seq Autoencoder με Attention μηχανισμό και Transformer περιγράφονται από τους υποπίνακες a και b. . . . .	61
4.2	Οι παράμετροι για τα μοντέλα Fully Connected Autoencoder και Convolutional Autoencoder περιγράφονται από τους υποπίνακες a και b. . . . .	61
5.1	Αποτελέσματα για <b>PR-AUC</b> σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	67
5.2	Αποτελέσματα για <b>ROC-AUC</b> σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	67
5.3	Αποτελέσματα για <b>PR-AUC</b> σε δεδομένα διαστήματος 10 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	69

5.4	Αποτελέσματα για <b>ROC-AUC</b> σε δεδομένα διαστήματος 10 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	69
5.5	Αποτελέσματα για <b>PR-AUC</b> σε δεδομένα διαστήματος μιας ώρας για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	70
5.6	Αποτελέσματα για <b>ROC-AUC</b> σε δεδομένα διαστήματος μιας ώρας για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	70
5.7	Τα Σφάλμα Ανακατασκευής των εξατομικευμένων μοντέλων μας περιγράφονται από τους Υποπίνακες a, b, c και d. . . . .	72
5.8	Αποτελέσματα για <b>PR-AUC</b> σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα, τα οποία εκπαιδεύτηκαν και αξιολογήθηκαν στα raw δεδομένα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	74
5.9	Αποτελέσματα για <b>ROC-AUC</b> σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα, τα οποία εκπαιδεύτηκαν και αξιολογήθηκαν στα raw δεδομένα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	74
5.10	Αποτελέσματα για <b>PR-AUC</b> σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα global μοντέλα που αξιολογήθηκαν ξεχωριστά για κάθε ασθενή αλλά και για όλους τους ασθενείς μαζί (επισημαίνεται ως Global). Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	76
5.11	Αποτελέσματα για <b>ROC-AUC</b> σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα global μοντέλα που αξιολογήθηκαν ξεχωριστά για κάθε ασθενή αλλά και για όλους τους ασθενείς μαζί (επισημαίνεται ως Global). Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή. . . . .	77
5.12	Σφάλμα Ανακατασκευής - Mean Square Error στα εξατομικευμένα μοντέλα, για Low και Moderate υποτροπές ξεχωριστά. . . . .	80
5.13	Σφάλμα Ανακατασκευής - Mean Square Error στα global μοντέλα με αξιολόγηση στα εξατομικευμένα δεδομένα, για Low και Moderate υποτροπές ξεχωριστά. . . . .	80
5.14	Σφάλμα Ανακατασκευής - Mean Square Error στα global μοντέλα, για Low, Moderate και Severe υποτροπές ξεχωριστά. . . . .	81
5.15	Αποτελέσματα για τα εξατομικευμένα μοντέλα που περιγράφτηκαν στην Ενότητα 4.2, σε δεδομένα διαστήματος 5 λεπτών. Παρουσιάζεται συνολική η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. . . . .	81
5.16	Σύγκριση απόδοσης με το άρθρο [Adl+20]. . . . .	82

# Κεφάλαιο 1

## Εισαγωγή

Η τεχνολογική πρόοδος που έχει σημειωθεί τα τελευταία χρόνια στις φορητές συσκευές έχει πλέον επιτρέψει την αξιόπιστη καταγραφή πληθώρας βιομετρικών δεικτών για λόγους υγείας και μας δίνει τη δυνατότητα μη παρεμβατικής απόκτησης φυσιολογικών δεδομένων, που θα μπορούσαν στην πραγματικότητα να μετατρέψουν την νοσοκομειοκεντρική πρακτική υγειονομικής περίθαλψης σε προληπτική, εξατομικευμένη φροντίδα. Με το σχεδιασμό νέων ευφυών ηλεκτρονικών υπηρεσιών μπορούν να αντιμετωπιστούν διάφορα ζητήματα υγείας και ευεξίας σε ασθενείς με ψυχικές διαταραχές, όπως η πρόληψη της υποτροπής μέσω του εντοπισμού προειδοποιητικών συμπτωμάτων.

Ο μεγάλος όγκος των δεδομένων που συλλέγονται, καθιστούν την τεχνητή νοημοσύνη (Artificial Intelligence - AI) σημαντικό τεχνολογικό σύμμαχο για την έρευνα σε ψυχικές διαταραχές, συμπεριλαμβανομένων εφαρμογών όπως μαγνητική τομογραφία, ηλεκτροεγκεφαλογραφία και κινησικές διαγνώσεις [Liu+20]. Τα ακατέργαστα δεδομένα όμως που συλλέγονται, δεν θα ήταν χρήσιμα χωρίς την επεξεργασία σήματος. Η υπολογιστική επεξεργασία σήματος είναι στην πραγματικότητα ο θεμέλιος λίθος, ο οποίος υποστηρίζει τη χαρτογράφηση από ακατέργαστα δεδομένα αισθητήρων έως αναπαραστάσεις συμπεριφορών και ψυχικών καταστάσεων. Η διεργασία ξεκινά με πρωτογενή σήματα από φορητούς αισθητήρες. Στη συνέχεια, τα σήματα πρέπει να μοντελοποιηθούν κατάλληλα για την εξαγωγή σημαντικών πληροφοριών, ενώ ως τελικό στάδιο, χρησιμοποιείται η μηχανική ή η βαθιά μάθηση για την εξαγωγή συμπερασμάτων σχετικά με τις ψυχικές καταστάσεις για την υποστήριξη της ανθρώπινης ή αυτόνομης λήψης αποφάσεων. Η εφαρμογή της μηχανικής μάθησης στον τομέα της ψυχικής υγείας είναι ακόμη υπό διερεύνηση και υψίστης σημασίας στην ψυχιατρική έρευνα [Adl+20; Elu+21; Gar+21; Wan+20].

Με τη διπλωματική αυτή, στοχεύουμε στην ανάπτυξη ενός ευφυούς συστήματος επεξεργασίας δεδομένων και μοντέλων βαθιάς μάθησης για την ανίχνευση υποτροπών σε ασθενείς με ψυχικές διαταραχές (Σχιζοφρένεια και Διπολική Διαταραχή Ι). Η εργασία αυτή πραγματοποιείται στο πλαίσιο του τρέχοντος έργου έρευνας και ανάπτυξης e-Prevention που στοχεύει σε καινοτόμες υπηρεσίες ηλεκτρονικής υγείας για την αποτελεσματική παρακολούθηση των ασθενών [Mag+20]. Η καταγραφή των βιομετρικών δεικτών γίνεται από πραγματικούς χρήστες, οι οποίοι πάσχουν από σχιζοφρένεια ή διπολική διαταραχή, με τη χρήση έξυπνων ρολογιών (smart-watches). Συγκεκριμένα, χρησιμοποιούνται φυσιολογικά σήματα όπως η κίνηση των ασθενών και βιοσήματα προερχόμενα από την παρακολούθηση καρδιακού ρυθμού. Στη συνέχεια, εφαρ-

μόζονται τεχνικές επεξεργασίας σήματος για την εξαγωγή χαρακτηριστικών, τα οποία θα μας βοηθήσουν στην ανίχνευση των υποτροπών. Το τελικό στάδιο του συστήματος είναι η ανάπτυξη βαθιών νευρωνικών δικτύων κωδικοποιητή-αποκωδικοποιητή (encoder-decoder) για τη ανίχνευση υποτροπών.

## 1.1 Η σημασία του προβλήματος

Η λέξη σχιζοφρενία αποτελείται από τα συνθετικά ‘σχίζο’ και ‘φρένια’ τα οποία σημαίνουν διάσπαση και μυαλό αντίστοιχα, δηλαδή είναι μία διαταραχή που προκαλεί σοβαρό κατακερματισμό της σκέψης και της προσωπικότητας [And95]. Η διπολική διαταραχή είναι μια σοβαρή διαταραχή της διάθεσης. Χαρακτηρίζεται από επαναλαμβανόμενα επεισόδια ευερέθιστης διάθεσης και κατάθλιψης, τα οποία συνοδεύονται από αλλαγές στη δραστηριότητα ή την ενέργεια, δηλαδή συνήθως βιώνουν μια αύξηση στην ενέργεια και μειωμένη ανάγκη για ύπνο, και σχετίζονται με χαρακτηριστικά γνωσιακά, σωματικά και συμπεριφορικά [AHS12]. Η ψύχωση είναι ένα φάσμα διαταραχών που αποτελούν τη βάση διαφορετικών αιτιολογικών μηχανισμών που δρουν στο Κεντρικό Νευρικό Σύστημα (ΚΝΣ). Τα συμπτώματα των ψυχωτικών διαταραχών περιλαμβάνουν τις παραισθήσεις, τις ψευδαισθήσεις, αλλαγές στη βούληση και τη νευρογνωσία, και τη συναισθηματική δυσρύθμιση [KP09]. Οι διαταραχές του φάσματος της σχιζοφρένειας (Schizophrenia Spectrum Disorder - SSD) είναι χρόνιες καταστάσεις και η σοβαρότητα των συμπτωμάτων και των λειτουργικών βλαβών εκδηλώνονται κατά τη διάρκεια της ασθένειας. Παρά την εκτεταμένη έρευνα τα τελευταία 60 χρόνια στη νευροβιολογία και τη νευροφυσιολογία των ψυχικών διαταραχών, οι αιτίες τους παραμένουν ασαφείς, ενώ αξιόπιστοι βιοδείκτες, είτε για τη διάγνωση είτε για την πρόβλεψη της πορείας της ψυχωτικής συμπτωματολογίας δεν έχουν βρεθεί ακόμη. Η χρήση τέτοιων δεικτών για την έγκαιρη διάγνωση και την πρόληψη ψυχωτικών υποτροπών είναι πλέον ένας από τους σημαντικότερους ερευνητικούς τομείς στη ψυχιατρική [EBM95; Kou+12; McG+14]. Με βάση το γεγονός ότι η διαδικασία της ψύχωσης είναι συνεχής και η υποτροπή είναι μια βιολογική διαδικασία που εξελίσσεται με την πάροδο του χρόνου [McC+86; Gae+93; Wie+98], θα ήταν αναμενόμενο να παρατηρηθούν αλλαγές σε τέτοιους δείκτες που σχετίζονται και πιθανόν να προηγούνται της έναρξης ή/και επιδείνωσης ψυχωτικών καταστάσεων. Μερικά από τα τυπικά πρώιμα προειδοποιητικά σημάδια ψυχικών ασθενειών περιλαμβάνουν ακαμψία, τρόμο, σπασμωδικές κινήσεις των χεριών, εγκατάλειψη των υπαίθριων δραστηριοτήτων κ.λ.π. [Max+16].

## 1.2 Αντιμετώπιση του προβλήματος

Η ανάπτυξη ολοκληρωμένων συστημάτων απομακρυσμένης παρακολούθησης για την ανίχνευση προειδοποιητικών ενδείξεων επικείμενων συμπτωματικών υποτροπών θα επέτρεπε στους κλινικούς ιατρούς να παρέμβουν πριν επιδεινωθεί η κατάσταση του ασθενούς [Adl+20] και θα μπορούσε να αποτρέψει τις ψυχωτικές υποτροπές πριν εκφραστούν πλήρως τα συμπτώματα του ασθενή. Επομένως, η τεχνολογία των αισθητήρων μπορεί να χρησιμοποιηθεί για να κατανοήσει τα πρότυπα της καθημερινής συμπεριφοράς που σχετίζονται με τη ψυχική υγεία, συμπεριλαμβανομένου του ύπνου, της διάθεσης και του στρες με απώτερο σκοπό την εξαγωγή βασικών δεικτών που να μπορούν ενδεχομένως να βοηθήσουν το ιατρικό προσωπικό στην παρακολούθηση ασθενών με ψυχικές διαταραχές [Mag+20].



Στη δική μας έρευνα, θα επικεντρωθούμε σε παθητικά δεδομένα που συλλέγονται μέσω ενός έξυπνου ρολογιού που παρέχουμε σε όλους τους ασθενείς, τα οποία είναι διακριτικό, μικρού βάρους, μπορεί να φορεθεί και να χρησιμοποιηθεί για παρακολούθηση ενώ οι ασθενείς εκτελούν όλες τις καθημερινές τους δραστηριότητες. Ειδικότερα, χρησιμοποιούμε φυσιολογικά σήματα κίνησης και παρακολούθησης της καρδιάς από τα έξυπνα ρολόγια, για να εντοπίσουμε κατάλληλους βιοδείκτες που μπορούν να βοηθήσουν στην πρόβλεψη ψυχωτικών υποτροπών. Επιπλέον, κάθε ασθενής παρακολουθείται από κλινικούς ιατρούς, των οποίων οι επισημειώσεις αφορούν την κατάσταση του ασθενούς η οποία χαρακτηρίζεται ως σταθερή ή υποτροπιάζουσα (όπου και δηλώνεται η συγκεκριμένη περίοδος της υποτροπής καθώς και η σοβαρότητά της ως χαμηλή (low), μέση (moderate) ή υψηλή (severe)). Ο προσδιορισμός των υποτροπών (δηλ. η εμφάνιση ψευδαισθήσεων σε ασθενείς με ψυχωτική διαταραχή, ή κατάθλιψη ή μανία) είναι πολυπαραγοντικός. Συγκεκριμένα, για τον προσδιορισμό της υποτροπής, οι ιατροί αξιολογούν και λαμβάνουν υπόψη τα εξής: 1) τις μηνιαίες συναντήσεις τους με τους ασθενείς που βοηθούν στον προσδιορισμό της διάρκειας και της σοβαρότητας της υποτροπής, καθώς και της αιτίας εμφάνισής της, 2) την αξιολόγηση με την κλίμακα PANSS (Positive and Negative Syndrome Scale), η οποία δίνει πολύτιμες πληροφορίες για την ίδια την υποτροπή, 3) τις πληροφορίες που προκύπτουν από την επικοινωνία με τον θεράποντα ιατρό, την οικογένεια ή τον φροντιστή του ασθενούς, και με το νοσοκομείο, σε περιπτώσεις που ο ασθενής χρειάστηκε νοσηλεία.

Για την αντιμετώπιση του προβλήματος ανίχνευσης ανωμαλιών, εφαρμόζονται διάφορα μοντέλα βαθιάς μάθησης κωδικοποιητή-αποκωδικοποιητή όπως οι transformer, τα πλήρως συνδεδεμένων νευρωνικών δικτύων, τα συνελικτικά νευρωνικά δίκτυα και τα αναδρομικά νευρωνικά δίκτυα (Recurrent Neural Network - RNN). Υπάρχουν πολλά πλεονεκτήματα στην εφαρμογή αυτών των μοντέλων. Από τη μια πλευρά, έχουμε τα sequence-to-sequence δίκτυα, όπως τον transformer και το RNN, στα οποία χρησιμοποιούνται οι self-attention μηχανισμοί για να μάθουν πολύπλοκα μοτίβα και δυναμικές [Wu+20; Li+19]. Παρ' όλα αυτά, τα RNN έχουν αποδειχθεί δύσκολα στην εκπαίδευση (προβλήματα που σχετίζονται με τη σύγκλιση), ενώ δεν μπορούν να παραλληλιστούν, οδηγώντας σε υψηλούς χρόνους απόκτησης αποτελεσμάτων. Γι' αυτό, έχουμε τα πλήρως συνδεδεμένα και συνελικτικά νευρωνικά δίκτυα, τα οποία είναι πλήρως παραλληλισμένα και μπορούν να καταγράψουν αποτελεσματικά μεγάλο περιεχόμενο, με αποτέλεσμα τη σταθερή αναπαράσταση ολόκληρης της ακολουθίας. Στην περίπτωση μας, εξετάσαμε εξατομικευμένα και global μοντέλα για δέκα ασθενείς, με τα οποία έχουμε λάβει υποσχόμενα αποτελέσματα για την ανίχνευση ψυχωτικών επεισοδίων.

### 1.3 Διαφορετικές Προσεγγίσεις και Εφαρμογές

Πρόσφατα, ο όρος 'ψηφιακός φαινοτυπισμός' [Tor+16] διευθετήθηκε για να περιγράψει την τεράστια εξατομικευμένη συλλογή πολυάριθμων συμπεριφορικών και βιομετρικών σημάτων μέσω των αισθητήρων φορητών συσκευών, όπως τα έξυπνα ρολόγια (smartwatches) και τα fitness trackers. Οι δείκτες που αποκτώνται από αυτά τα σήματα, αξιοποιούνται ήδη στη γενική ιατρική και έχει γίνει και η εισαγωγή τους στην κλινική ψυχιατρική [AMC17]. Ιδιαίτερα, αποτελούν αντικείμενο εκμετάλλευσης σε ψυχικές ασθένειες, όπως η σχιζοφρένεια [Wan+16], η κατάθλιψη [Sae+15] και η διπολική διαταραχή [Gra+15]. Η πρόβλεψη υποτροπών στη σχιζοφρένεια μελετάται ευρέως από διάφορους ερευνητές. Οι μελέτες κυμαίνονται από την κατανόηση της συμπεριφοράς αυτών των ατόμων μέσω στατιστικών μοντέλων και νευρωνικών δικτύων μέχρι και την έγκαιρη πρόβλεψη υποτροπών. Για την έγκαιρη πρόβλεψη υποτροπών πρέπει πρώτα να γίνει η

ανίχνευση της υποτροπής. Η υποτροπή θεωρείται ως ανωμαλία σε μια ακολουθία από δεδομένα. Γενικά, αρκετές μελέτες έχουν διεξαχθεί για την κατανόηση των θεωρητικών και πρακτικών εφαρμογών βαθιάς μάθησης στη ιατρική, με στόχο την ανίχνευση ανωμαλιών (anomaly detection) [Wan+16; Cow+18; Sat+18]. Η εύρεση σπάνιων συμβάντων/ανωμαλιών σε τομείς όπως η ανάλυση ιατρικής εικόνας, τα αρχεία κλινικής ηλεκτροεγκεφαλογραφίας (EEG), επιτρέπουν τη διάγνωση και την παροχή προληπτικών θεραπειών για μια ποικιλία ιατρικών καταστάσεων [CC19]. Η μέθοδος του anomaly detection είναι ουσιαστικά η ανίχνευση κάποιας ανωμαλίας. Ο Hawkins [Haw80] ορίζει μια ανωμαλία ως μια παρατήρηση που αποκλίνει τόσο σημαντικά από άλλες παρατηρήσεις ώστε να εγείρει υποψίες ότι δημιουργήθηκε από διαφορετικό ‘μηχανισμό’. Στη δική μας μελέτη, θα εφαρμόσουμε τη μέθοδο του anomaly detection στη ψυχιατρική, με σκοπό να εντοπίσουμε συμπεριφορές στα δεδομένα, οι οποίες θα παρεκκλίνουν από τις φυσιολογικές.

Στην πρόσφατη βιβλιογραφία, η μελέτη της σχιζοφρένειας μέσω του ‘ψηφιακού φαινοτύπου’ βασίστηκε κυρίως σε δεδομένα που συλλέχθηκαν από smartphones [Adl+20; Hen+20; Hen+21]. Στη μελέτη [Hen+21], χρησιμοποιήθηκαν δεδομένα από smartphones για τη πρόβλεψη κλινικών υποτροπών σε άτομα με σχιζοφρένεια. Τα δεδομένα συλλέχθηκαν από τους αισθητήρες του smartphone και από εφαρμογές που εγκαταστάθηκαν σε αυτό για τη συλλογή δεδομένων EMA (Ecological Momentary Assessment). Το EMA είναι ουσιαστικά κάποια ερωτηματολόγια που πρέπει να απαντήσουν οι ασθενείς, τα οποία σχετίζονται με τη κατάθλιψη, το άγχος, τον ύπνο, την κοινωνικότητα και τη ψύχωση τους [Wan+16; Bar+18]. Τα δεδομένα αποτελούνται από 83 συμμετέχοντες που είχαν διαγνωστεί με σχιζοφρένεια και 43 που ήταν υγιείς. Επιλέχθηκαν να χρησιμοποιηθούν χαρακτηριστικά κινητικότητας (από το GPS), κοινωνικότητας (από τα αρχεία κλήσεων/μηνυμάτων), το EMA (Ecological Momentary Assessment), του χρόνου που η οθόνη είναι ενεργή, της διάρκειας του ύπνου και των γνωσιακών διεργασιών (όπως η οπτική προσοχή και η εναλλαγή εργασιών, από Jewels Trails A και B [Liu+19]), κάποια από τα οποία αποδείχθηκε ότι είναι σημαντικά για τη μελέτη μοντέλων πρόβλεψης υποτροπών. Τα χαρακτηριστικά αυτά κανονικοποιήθηκαν για κάθε συμμετέχοντα, καθορίζοντας μια βασική ακολουθία αυτών, έτσι ώστε όταν τα χαρακτηριστικά αποκλίνουν σημαντικά από αυτή τη ακολουθία, να ανιχνευθούν ως ημέρες υποτροπής. Τα αποτελέσματα στη πρόβλεψη υποτροπής στη συγκεκριμένη έρευνα είναι 89% ευαισθησία (sensitivity) και 75% ειδικότητα (specificity). Στη προαναφερόμενη έρευνα, εκτός από τα παθητικά δεδομένα χρησιμοποιήθηκαν και δεδομένα από το EMA, γεγονός που περιπλέκει πολλές φορές το αποτέλεσμα, αφού καθώς οι ασθενείς πλησιάζουν στη υποτροπή και βιώνουν αλλαγές στις γνωσιακές διεργασίες χάνουν τη εικόνα των συμπτωμάτων τους και αναφέρουν διαφορετικά συμπτώματα από τα πραγματικά [Hen+21].

Υπάρχουν εργασίες που προσπάθησαν να αντιμετωπίσουν τα διάφορα προβλήματα σχετικά με τις ψυχικές διαταραχές, προσφέροντας πολλές υποσχόμενες πληροφορίες χρησιμοποιώντας δεδομένα από αισθητήρες. Στο [Max+16] χρησιμοποιήθηκαν δεδομένα επιταχυνσιόμετρου και ήχου από smartphones, για την ταξινόμηση της κατάστασης 5 ασθενών με διπολική διαταραχή, που συλλέχθηκαν σε χρονικό διάστημα 12 εβδομάδων, οι πραγματικές τους δραστηριότητες, δείχνοντας τη δυνατότητα ταξινόμησης με μεγάλη σιγουριά της πορείας των επεισοδίων της διάθεσης ή των υποτροπών. Στο [Cha+17] χρησιμοποιήθηκαν στατιστικές μέθοδοι για τον χαρακτηρισμό προτύπων δραστηριότητας που προέρχονται από επιταχυνσιόμετρο (που συλλέγονται σε διάστημα μιας εβδομάδας) από 99 ενήλικες με ετερογενείς ψυχικές ασθένειες, υποστηρίζοντας ότι τα πρότυπα δραστηριότητας ποικίλλουν μεταξύ διαφορετικών διαταραχών, ενώ



η πλούσια φύση της ανθρώπινης κίνησης που καταγράφεται κατά τη διάρκεια της εγρήγορσης και του ύπνου, θα μπορούσε να ποσοτικοποιηθεί. Στη πιλοτική μελέτη [Bar+18] διάρκειας 3 μηνών, εφαρμόζεται ανίχνευση ανωμαλιών σε ενεργά και παθητικά δεδομένα από smartphones σε 15 άτομα με σχιζοφρένεια. Από τα 15 άτομα, πέντε παρουσίασαν κλινική υποτροπή και μόνο τα τρία από αυτά είχαν επαρκή δεδομένα για ανίχνευση ανωμαλιών. Για την ανίχνευση ανωμαλιών εφαρμόστηκε ένας στατιστικός έλεγχος εμπνευσμένος από την προσέγγιση του Filzmoser [FRG04]. Παρατηρώντας τις αλλαγές στα πρότυπα κινητικότητας και στην κοινωνική συμπεριφορά, αυτών των ασθενών, ήταν σε θέση να εντοπίσουν στατιστικά σημαντικές ανωμαλίες στη συμπεριφορά των ασθενών τις ημέρες πριν από την υποτροπή. Στο άρθρο [Fil+20], γίνεται μία στατιστική ανάλυση σε δεδομένα που συλλέγονται από smartwatch μέσω μιας πλατφόρμας [Mag+20], για να δείξουν πώς τα παθητικά δεδομένα μπορούν να χρησιμοποιηθούν για να συμπεράνουν διαφορές μεταξύ ατόμων με ψυχικές διαταραχές και υγιών ατόμων.

Γενικά, η ανίχνευση ανωμαλιών είναι το πεδίο όπου ανήκει η πρόβλεψη της υποτροπής. Σε αυτό το πεδίο η χρήση της τεχνητής νοημοσύνης είναι ευρεία [Pan+21] και πολλές εφαρμογές της εντοπίστηκαν για την ανίχνευση ιατρικών ανωμαλιών, όπως ανίχνευση επιληπτικών επεισοδίων, ανωμαλίες σε εικόνες ενδοσκόπησης και καρδιακού ρυθμού [Fer+20]. Σε μία άλλη σχετική έρευνα [Adl+20] που αναλύεται εκτενέστερα πιο κάτω, χρησιμοποιώντας και αυτή παθητικά δεδομένα από το σύστημα CrossCheck [Wan+16] υλοποιήθηκε ποικιλία μοντέλων ανίχνευσης ανωμαλιών κωδικοποιητή-αποκωδικοποιητή, όπως πλήρως συνδεδεμένα νευρωνικά δίκτυα (Fully Connected Neural Network - FNN) και Gated Recurrent Unit (GRU) sequence-to-sequence νευρωνικά δίκτυα. Αυτές οι αρχιτεκτονικές όπως τα πλήρως συνδεδεμένα νευρωνικά δίκτυα και sequence to sequence μοντέλα, εφαρμόζονται σε μεγάλο βαθμό σε διαδοχικά δεδομένα όπως σε αλληλουχίες DNA, στην επεξεργασία φυσικής γλώσσας, σε δεδομένα ομιλίας [Gar+21], σε βίντεο καθώς και σε δεδομένα χρονοσειρών [Ret+].

Στο κεφάλαιο αυτό θα αναλύσουμε εκτενέστερα κάποιες από τις έρευνες που ήδη αναφέρθηκαν και σχετίζονται με διαταραχές του φάσματος της σχιζοφρένειας. Αναλύουμε στατιστικά μοντέλα και νευρωνικά δίκτυα που θα μας βοηθήσουν στη πρόβλεψη υποτροπών σε άτομα με ψυχικές διαταραχές με σκοπό τη δημιουργία της δικής μας προσέγγισης για την έγκαιρη πρόβλεψη υποτροπών.

### 1.3.1 Επεξεργασία σήματος στη Ψυχική Υγεία

Η επεξεργασία σήματος είναι ο ακρογωνιαίος λίθος που υποστηρίζει τη συλλογή δεδομένων με αποτέλεσμα να μπορεί να εξάγει τις αναπαραστάσεις συμπεριφορών και νοητικών καταστάσεων. Δηλαδή, η μέθοδος επεξεργασίας σήματος συμπεριφοράς (Behavioral Signal Processing (BSP)) επιδιώκει να ποσοτικοποιήσει ποιοτικά χαρακτηριστικές συμπεριφορές εξάγοντας πληροφορίες από βιοσήματα.

Το πρώτο στάδιο είναι η συλλογή δεδομένων, για την οποία επιθυμούμε δεδομένα υψηλής ποιότητας και υψηλής συνοχής. Δεδομένου ότι η συμπεριφορά είναι εγγενώς πολυτροπική, πρέπει να συλλέγουμε πολυτροπικά δεδομένα. Αυτή η συλλογή πρέπει να είναι φυσική και να μην επηρεάζει την ίδια τη συμπεριφορά. Δηλαδή συλλέγουμε πρώτα ακατέργαστα (raw) σήματα, π.χ., από φυσιολογικούς αισθητήρες, αφού είναι σημαντικό να υπάρχει διαρκής παρακολούθηση της συμπεριφοράς. Επιπλέον, πρέπει να ληφθεί υπόψη για τη μοντελοποίηση, ότι η συμπεριφορά ενός ατόμου δεν είναι απομονωμένη, αλλά επηρεάζεται από τις δράσεις των επικοινωνιακών αλληλεπιδράσεων

και το περιβάλλον που βρίσκεται [Bon+17].

Σε έρευνες που σχετίζονται με την ιατρική όπως η ανίχνευση του άγχους [DM21], το παραλήρημα [Oh+18], η κολπική μαρμαρυγή [Ebr+18], η άπνοια [Uca+18], κ.α, έχουν προταθεί διάφορες αναπαραστάσεις χαρακτηριστικών χρησιμοποιώντας δεδομένα από wearables. Μια από τις πιο συχνά χρησιμοποιούμενες αναπαραστάσεις χαρακτηριστικών που έχει χρησιμοποιηθεί σε έρευνες που σχετίζονται με τις διάφορες διαταραχές του φάσματος της σχιζοφρένειας [Hen+10; Vos+06], είναι η μεταβλητότητα του καρδιακού ρυθμού (Heart Rate Variability - HRV). Το HRV είναι το φυσιολογικό φαινόμενο της διακύμανσης του χρονικού διαστήματος μεταξύ των διαδοχικών καρδιακών παλμών σε χιλιοστά του δευτερολέπτου. Αρκετές μελέτες έχουν δείξει ότι τα υγιή άτομα εμφανίζουν υψηλό βαθμό HRV, κάτι που βοηθά στην γρήγορη προσαρμογή στις φυσικές ή ψυχολογικές απαιτήσεις του περιβάλλοντος, ενώ τα χαμηλότερα επίπεδα και οι μειώσεις του HRV σχετίζονται με καρδιακές βλάβες [Sin95], οι οποίες μπορεί να υποδηλώνουν δυσπροσαρμοστικές αποκρίσεις σε στρεσογόνους παράγοντες [CFV05], ενώ συνδέονται επίσης έντονα με διαταραχές της διάθεσης και έχουν αναφερθεί σε άτομα που πάσχουν από κατάθλιψη [Age+02]. Επίσης, έχει αποδειχθεί ότι οι καρδιακοί ασθενείς με πιο σοβαρή κατάθλιψη εμφανίζουν χαμηλότερο HRV σε σύγκριση με εκείνους με λιγότερο σοβαρή κατάθλιψη [Kri+97], ενώ η επιτυχημένη θεραπεία της κατάθλιψης τείνει να συνοδεύεται από αυξημένο HRV [CA02]. Επιπλέον, οι αγχώδεις διαταραχές έχουν συσχετιστεί με χαμηλό HRV [FT98], ενώ δυσλειτουργία του HRV έχει επίσης παρατηρηθεί στη σχιζοφρένεια, [Bär+05; Boe+06]. Οι συχνότερες γραμμικές μέθοδοι για την ανάλυση του HRV είναι ο μέσος καρδιακός ρυθμός (Heart Rate - HR) [Hen+10] και ο μέσος HRV [KSB05], η τυπική απόκλιση (Standard Deviation - SD) και ο μέσος όρος των διαδοχικών διαφορών (Root Mean Squared Successive Differences - RMSSD) [Vos+06], το pNNx [Vos+06], καθώς και διάφορα φασματικά χαρακτηριστικά, συγκεκριμένα, το LF (Low frequency) και το HF (High frequency), δηλαδή και η ισχύς στη ζώνη χαμηλής ή υψηλής συχνότητας, μαζί με την αναλογία LF-προς-HF [Vos+06; KSB05]. Για μια ανασκόπηση των γραμμικών χαρακτηριστικών βλ. [Can+20].

Την τελευταία δεκαετία, έχει δοθεί αυξανόμενη έμφαση στην εφαρμογή μη γραμμικών μεθόδων για τον χαρακτηρισμό της καρδιακής λειτουργίας. Οι διακυμάνσεις στον καρδιακό ρυθμό έχουν αναφερθεί ότι παρουσιάζουν ακανόνιστη μεταβλητότητα που υποδηλώνουν μη γραμμική συμπεριφορά [Mäk+02]. Τέτοιες μέθοδοι περιλαμβάνουν ανάλυση αποτεινόμενης διακύμανσης (detrended fluctuation analysis) [Pen+94], κατά προσέγγιση εντροπία (Approximate Entropy) [Pin91], εντροπία δείγματος (Sample Entropy) [RLM04], διαγράμματα Poincare (Poincare Plots) [BPK01], συμβολική δυναμική [Vos+06; TBC05; Hui+09], φράκταλ διάσταση (Fractal Dimension) [MM82; Fal04; LB09] και εκθέτες Lyapunov [KY79], μεταξύ άλλων. Υπάρχουν ενδείξεις ότι οι μη γραμμικές μετρήσεις είναι ανώτεροι προγνωστικοί παράγοντες της καρδιακής και αυτόνομης δυσλειτουργίας [Bau+04; Bau+05; Hoy+06], σε σύγκριση με πιο παραδοσιακές αναλύσεις στο πεδίο του χρόνου και της συχνότητας, ωστόσο δεν έχουν ακόμη εξεταστεί εκτενώς σε ερευνητικούς τομείς που περιλαμβάνουν άτομα με ψυχικές διαταραχές.

Στο [Bär+07], εκτός από τις γραμμικές παραμέτρους HRV, χρησιμοποιήθηκαν μη γραμμικά μέτρα πολυπλοκότητας (δηλαδή, φράκταλ διάσταση, συμπίεση και κατά προσέγγιση εντροπία) για να περιγράψουν την κανονικότητα του καρδιακού ρυθμού σε ασθενείς με οξεία σχιζοφρένεια. Είκοσι ασθενείς με παρανοϊκή σχιζοφρένεια και είκοσι υγιείς, συμμετείχαν για τη συλλογή δεδομένων ηλεκτροκαρδιογραφήματος για 30 συνεχόμενα λεπτά, με σκοπό να συγκριθούν τα δεδομένα τους με αυτά των ασθενών. Η στατιστική ανάλυση έδειξε ότι μόνο η φράκταλ διάσ-

ταση παρέμεινε σημαντικά αλλαγμένη και φαίνεται να είναι μια σχετικά πιο σημαντική ανεξάρτητη παράμετρος του HR. Στο [Hen+10], 47 ασθενείς που εισήχθησαν για νοσηλεία είτε με διπολική διαταραχή είτε με σχιζοφρένεια επιλέχθηκαν για να εξετάσουν τις αλλοιώσεις του HRV τους σε σύγκριση με 23 υγιείς συμμετέχοντες. Ο HRV ποσοτικοποιήθηκε χρησιμοποιώντας γραμμικά και μη γραμμικά χαρακτηριστικά, τα οποία έδειξαν ότι τα άτομα που έπασχαν ιδιαίτερα από διπολική διαταραχή εμφάνισαν σημαντική μείωση του HRV σε σύγκριση με τους υγιείς συμμετέχοντες, ενώ η μειωμένη εντροπία συσχετίστηκε με αυξημένη επιθετικότητα και μειωμένη προσωπική υγιεινή μεταξύ άλλων. Οι εκθέτες Entropy και Lyapunov επίσης χρησιμοποιήθηκαν στα άρθρα [Bär+07; Mag+20], όπου παρουσιάστηκε μειωμένο μοτίβο πολυπλοκότητας ή μικρότερη μεταβλητότητα σε καταθλιπτικούς ή σχιζοφρενικούς ασθενείς σε σύγκριση με τα υγιείς άτομα. Στο [Vos+06] παρουσιάζεται μια ανασκόπηση γραμμικών και μη γραμμικών μεθόδων για HRV, υποδεικνύοντας τα πιο σημαντικά χαρακτηριστικά, ενώ τα κοινά χρησιμοποιούμενα χαρακτηριστικά HRV και οι διαγνωστικές τους χρήσεις βρίσκονται στο [KSB05].

Σε πρόσφατη έρευνα [Fil+20], υλοποιήθηκε μια αυστηρή στατιστική ανάλυση των βραχυχρόνιων χαρακτηριστικών που εξάγονται από φορητές συσκευές, κατά τη διάρκεια μακροχρόνιας συνεχούς παρακολούθησης ασθενών με ψυχικές διαταραχές και υγιείς συμμετέχοντες, αντίστοιχα. Τα δεδομένα που χρησιμοποιήθηκαν εξάγονται από ένα έξυπνο ρολόι που φοριέται όλο το 24ωρο εκτός από τη φόρτιση (περίπου 2 ώρες). Με χρήση τεχνικών επεξεργασίας σήματος εξάγονται χαρακτηριστικά προκειμένου να ληφθεί μια γενική εκτίμηση για το πως διαφοροποιούνται υγιείς άτομα και ασθενείς με ψυχωτικές διαταραχές. Τα χαρακτηριστικά που εξήχθησαν παρουσιάζονται λεπτομερώς στη επόμενη ενότητα. Τα δεδομένα αποτελούνται από μετρήσεις της επιτάχυνσης, της γωνιακής ταχύτητας και τον καρδιακό ρυθμό που λήφθηκαν από 23 υγιείς εθελοντές και 22 ασθενείς με ψυχική διαταραχή (9 με σχιζοφρένεια, 8 με διπολική διαταραχή I, 3 με σύντομο ψυχωτικό επεισόδιο και 2 με σχιζοσυναισθηματική διαταραχή). Με την ανάλυση των εξαχθέντων χαρακτηριστικών παρατηρήθηκαν σημαντικές διαφορές ανάμεσα στους εθελοντές και στους ασθενείς, σχετικά με τη συμπεριφορά της κίνησης όπως επίσης σε ορισμένους δείκτες της καρδιακής λειτουργίας τόσο κατά τη διάρκεια της εγρήγορσης όσο και του ύπνου. Για να εξαχθούν χαρακτηριστικά από τα δεδομένα που συλλέγονται, εφαρμόζονται ανάλυσης βραχέος χρόνου σε 'μεγαλύτερη' κλίμακα από ό,τι στα παραδοσιακά σήματα (όπως ο ήχος), δεδομένου της μεγάλης χρονικής τους διάρκειας. Για δεδομένα κίνησης (accelerometer - acc και gyroscope - gyr), το παράθυρο ανάλυσης είναι 10 λεπτά χωρίς επικάλυψη, αφού σε προηγούμενη έρευνα είχε βρεθεί ότι ήταν βέλτιστο για τη διάκριση βραχυπρόθεσμων προτύπων [Ret+], ενώ για δεδομένα HRV χρησιμοποιείται ένα παράθυρο μιας ώρας, αφού το HRV από το σήμα 5Hz που καταγράφεται από το ρολόι συνίστανται στην απόρριψη των ίδιων διαδοχικών τιμών, διασφαλίζοντας ότι για κάθε διάστημα μιας ώρας, η λαμβανόμενη ακολουθία διαστημάτων RR αθροίζονται σε τουλάχιστον 54 λεπτά (ένα εμπειρικό όριο που αντιστοιχεί στο 90% έγκυρων καρδιακών δεδομένων). Στα αποτελέσματα αυτής της μελέτης, υπάρχουν μεγάλες διαφορές στη τυπική απόκλιση της εντροπίας του δείγματος (Sample Entropy) [RM00], στο μέσο όρο του Higuchi fractal dimension [Hig88], καθώς και διάφορα στατιστικά στοιχεία που προέρχονται από την ανάλυση Multi-scale Fractal Dimension (MFD) [Mar94], κατά τη διάρκεια του ύπνου. Αντίστοιχα, κατά τη διάρκεια της εγρήγορσης, μεγάλες διαφορές παρατηρήθηκαν στο μέσο όρο της εντροπίας του δείγματος, στο standard deviation (*std*) των poincare χαρακτηριστικών [BPK01] και πάλι σε διάφορα στατιστικά στοιχεία του MFD [Mar94; Fil+20]. Τα πιο πάνω χαρακτηριστικά περιγράφονται εκτενώς στη Ενότητα 2.1, αφού κάποια από αυτά χρησιμοποιήθηκαν στη δική μας μελέτη για την ανίχνευση των υποτροπών.

Τέλος, καθώς έχουμε εντοπίσει τις κατάλληλες αναπαραστάσεις της συμπεριφοράς που προέρχονται από τα σήματα, μπορεί να γίνει χρήση μηχανικής μάθησης για την εξαγωγή συμπερασμάτων σχετικά με τη ψυχική κατάσταση του ατόμου. Παρ' όλα αυτά, οι αποφάσεις για την υγεία ενός ατόμου είναι τόσο ζωτικής σημασίας που ακόμη και αν ένα σύστημα είναι σε θέση να διαγνώσει μια διαταραχή με 100% ακρίβεια, εάν αυτό το σύστημα είναι ένα μαύρο κουτί, δηλαδή η διαδικασία λήψης αποφάσεων είναι ασαφής, πολλές φορές οι κλινικοί ιατροί είναι δύσκολο να το εμπιστευτούν. Και δικαίως, δεδομένου ότι υπάρχουν σίγουρα παγίδες στις οποίες μπορούν να πέσουν οι προσεγγίσεις μηχανικής μάθησης. Ένας τρόπος αντιμετώπισης αυτού, είναι μέσω μιας προσέγγισης, η οποία να ενσωματώνει την ανθρώπινη γνώση στη δημιουργία των χαρακτηριστικών και στη μοντελοποίηση, δηλαδή να ληφθούν υπόψη τόσο τα δεδομένα όσο και το ιατρικό φαινόμενο το οποίο μοντελοποιείται (π.χ., στη δική μας περίπτωση η ψυχική υποτροπή). Η ημιεπιβλεπόμενη μάθηση μπορεί να είναι μια ενδιάμεση λύση, αφού συνδυάζει δεδομένα τα οποία έχουν κάποια ετικέτα (label) και δεδομένα τα οποία δεν έχουν ετικέτα, με σκοπό την καλύτερη εκπαίδευση μοντέλων μηχανική μάθησης και, ταυτόχρονα, να εμπεριέχει τη γνώση από ανθρώπινους ειδικούς [Bon+17].

### 1.3.2 Σχετικές έρευνες με νευρωνικά δίκτυα

Σε αντίθεση με τις στατιστικές προσεγγίσεις, τα νευρωνικά δίκτυα δεν απαιτούν υποθέσεις σχετικά με την υποκείμενη κατανομή των δεδομένων και είναι συχνά ιδανικά σε σύγκριση με τις κλασικές τεχνικές μηχανικής μάθησης, επειδή μπορούν να παρέχουν ακριβείς προβλέψεις χωρίς την ανάγκη σύνθετων χαρακτηριστικών. Οι ερευνητές έχουν αρχίσει να χρησιμοποιούν ανίχνευση ανωμαλιών για να προβλέψουν προειδοποιητικά σημάδια που σχετίζονται με τις υποτροπές.

Οι πιο καινοτόμες προσεγγίσεις για την ανίχνευση ανωμαλιών από δεδομένα χρονοσειρών χρησιμοποιούν μοντέλα νευρωνικών δικτύων κωδικοποιητή-αποκωδικοποιητή (Encoder-Decoder). Στη έρευνα [Mal+16], στην οποία και βασίστηκε ένα κομμάτι της συγκεκριμένης εργασίας, προτείνεται ένα μοντέλο Long Short-term memory (LSTM) Encoder-Decoder το οποίο μαθαίνει να ανακατασκευάζει την 'κανονική' συμπεριφορά των χρονοσειρών και στη συνέχεια χρησιμοποιεί το σφάλμα ανακατασκευής για τον εντοπισμό των ανωμαλιών. Το καινοτόμο στοιχείο της συγκεκριμένης μελέτης, είναι ότι τα δεδομένα αποτελούνται από χρονοσειρές πολλαπλών αισθητήρων (Multi-sensor) και έτσι η ανίχνευση ανωμαλιών γίνεται πάνω σε αυτά τα δεδομένα. Δηλαδή, ο κωδικοποιητής μαθαίνει μια διανυσματική αναπαράσταση των χρονοσειρών εισόδου και ο αποκωδικοποιητής χρησιμοποιεί αυτήν την αναπαράσταση για να την ανακατασκευάσει. Το συγκεκριμένο μοντέλο αξιολογήθηκε σε τέσσερα διαφορετικά σύνολο δεδομένων από τον πραγματικό κόσμο, με το καλύτερο αποτέλεσμα να έχει Precision, Recall,  $F_0.1$  και  $TPR/FPR$  0.96, 0.18, 0.93 και 7.6, αντίστοιχα. Επιπλέον, στη συγκεκριμένη εργασία παρατηρήθηκε ότι αυτό το μοντέλο είναι σε θέση να ανιχνεύει ανωμαλίες ακόμη και σε πιο απρόβλεπτες χρονοσειρές όπως επίσης και να ανιχνεύει ανωμαλίες σε χρονοσειρές με πολύ μεγάλο μήκος. Στο [Cow+18] δημιουργήθηκε ένας Gated Recurrent Unit (GRU) Encoder-Decoder, ο οποίος εκπαιδεύεται για να ανακατασκευάζει τα δεδομένα εισόδου του, με σκοπό την έγκαιρη ανίχνευση αναπνευστικών ασθενειών που αναπτύσσονται σε γουρούνια μέσω δεδομένων από περιβαλλοντικούς αισθητήρες. Τα δεδομένα που χρησιμοποιήθηκαν για αυτή την ανίχνευση, συλλέγονται από τρεις αισθητήρες, οι οποίοι μετράνε τη θερμοκρασία, την υγρασία και το διοξείδιο του άνθρακα ( $CO_2$ ). Το μοντέλο εκπαιδεύεται χρησιμοποιώντας τα περιβαλλοντικά δεδομένα που δεν οδήγησαν σε εμφάνιση



αναπνευστικής νόσου στα γουρούνια, τα οποία επισημαίνονται ως ‘κανονικά’ και αξιολογείται σε δεδομένα τα οποία περιέχουν ‘κανονικά’ και ‘ανώμαλα’ δεδομένα (δηλ. δεδομένα που οδήγησαν σε εμφάνιση αναπνευστικής νόσου στα γουρούνια). Τα ‘ανώμαλα’ δεδομένα είχαν μεγαλύτερο σφάλμα ανακατασκευής από τα κανονικά. Η ανίχνευση της ανωμαλίας γίνεται με βάση ενός κατωφλιού, το οποίο εντοπίστηκε χρησιμοποιώντας τον αλγόριθμο βελτιστοποίησης Particle Swarm Optimisation (PSO) [SE98]. Τα αποτελέσματα από τη μέθοδο αυτής της μελέτης, έδειξαν ότι το μοντέλο GRU AE σε συνδυασμό με τον ανιχνευτή ανωμαλιών βελτιστοποιημένο με PSO, δείχνει ισχυρές δυνατότητες στην ικανότητά του να ανιχνεύει ανωμαλίες που θα οδηγήσουν σε εμφανίσεις αναπνευστικών ασθενειών στα γουρούνια. Αυτό το μοντέλο μπόρεσε να ξεπεράσει άλλες συγκρίσιμες μεθοδολογίες στην ακρίβεια, την ανάκληση και τη συνολική απόδοση. Το άρθρο [Elu+21] διερευνά τη χρήση τεχνικών βαθιών νευρωνικών δικτύων και μηχανικής μάθησης όπως το πολυστρωματικό perceptron (Multi Layer Perceptron - MLP), τη μηχανή υποστήριξης διανυσμάτων (Support Vector Machine - SVM), το random forest (RF) και το δέντρο αποφάσεων (Decision Tree - DT), για τον εντοπισμό μοτίβων στα δεδομένα ασθενών, για την κατάλληλη ταξινόμηση των ασθενειών τους. Υλοποιήθηκε ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο βαθιάς μάθησης, το οποίο αποτελείται από 3 επίπεδα, με αριθμών νευρώνων 20–20–40, αντίστοιχα. Στη συγκεκριμένη μελέτη το σύνολο δεδομένων αποτελείται από 500 ασθενείς που πάσχουν από ασθένειες σχετικές με ψυχωτικές διαταραχές όπως, διπολική διαταραχή, αγγειακή άνοια, διαταραχή ελλειμματικής προσοχής/υπερκινητικότητας, αϋπνία και σχιζοφρένεια. Για την εκπαίδευση του πλήρους συνδεδεμένου νευρωνικού δικτύου, χρησιμοποιήθηκαν τα ιατρικά αρχεία των ασθενών τα οποία περιείχαν 15 χαρακτηριστικά που σχετίζονταν με τα δημογραφικά στοιχεία (φύλο, ηλικία, θρησκεία, επάγγελμα και οικογενειακή κατάσταση), με οικογενειακά ζητήματα (απώλεια γονέα, ιστορικό τέτοιας πάθησης στη οικογένεια, οικογενειακή κατάσταση, διαζύγιο), με κληρονομικότητα (γενετική πάθηση) και αν είχες κάποιο τραυματισμό στο κεφάλι. Επίσης, με τα ίδια δεδομένα εκπαιδεύτηκαν τα μοντέλα MLP, SVM, RF και DT. Τα αποτελέσματα έδειξαν ότι το προτεινόμενο βαθύ νευρωνικό δίκτυο είχε τη καλύτερη απόδοση με ακρίβεια 75.17% σε δεδομένα όταν υπήρχε ανισορροπία κλάσης, ενώ η ακρίβεια του μοντέλου MLP είναι 58.44%, στα ίδια δεδομένα. Αντίθετα, την καλύτερη απόδοση στις τεχνικές μηχανικής μάθησης παρουσίασε το μοντέλο random forest, χρησιμοποιώντας το σύνολο δεδομένων χωρίς ανισορροπία κλάσης και το αποτέλεσμά του, σε σύγκριση με τις τεχνικές βαθιάς μάθησης, είναι 64.1% και 55.87%, αντίστοιχα. Στη συγκεκριμένη έρευνα, παρατηρήθηκε επίσης ότι η ηλικία του ασθενούς είναι το χαρακτηριστικό που συμβάλλει περισσότερο στην απόδοση του μοντέλου ενώ το διαζύγιο είναι το λιγότερο. Τέλος, η μελέτη αυτή αποκαλύπτει ότι υπάρχει υψηλή τάση για έναν ασθενή με διπολική διαταραχή να έχει αϋπνία.

Η αναγνώριση ατόμου (Person ID) από φορητά σήματα μπορεί να παρέχει χρήσιμες πληροφορίες για ασθενείς με ψυχικές διαταραχές, αναγνωρίζοντας τη διακριτική συμπεριφορά αυτών των ατόμων. Δηλαδή, δημιουργώντας ένα αξιόπιστο προφίλ συμπεριφοράς και κατά συνέπεια ανιχνεύοντας σημαντικές αλλαγές στη συμπεριφορά λόγω της κατάστασής τους. Στο άρθρο [Ret+] διερευνάται η ταυτοποίηση προσώπων από δεδομένα που συλλέγονται από ένα έξυπνο ρολόι Samsung Gear S3 Frontier, από 20 εθελοντές. Τα δεδομένα που εξάγονται από το ρολόι είναι η γωνιακή ταχύτητα η γραμμική επιτάχυνση και μετρήσεις καρδιακού ρυθμού, τα οποία ταξινομήθηκαν σε τρεις κλάσης (ύπνος, περπάτημα, το άτομο δεν κοιμάται ούτε περπατά), όπου αυτά με την σειρά τους μας βοηθούν στο να αντιστοιχηθεί ο αισθητήρας με το αντίστοιχο άτομο. Τα δεδομένα που χρησιμοποιήθηκαν στη συγκεκριμένη έρευνα είναι τα ίδια που στη συνέχεια θα χρησιμοποιήσουμε στη παρούσα μελέτη. Οι ερευνητές ανέπτυξαν ένα βαθύ νευρωνικό δίκ-

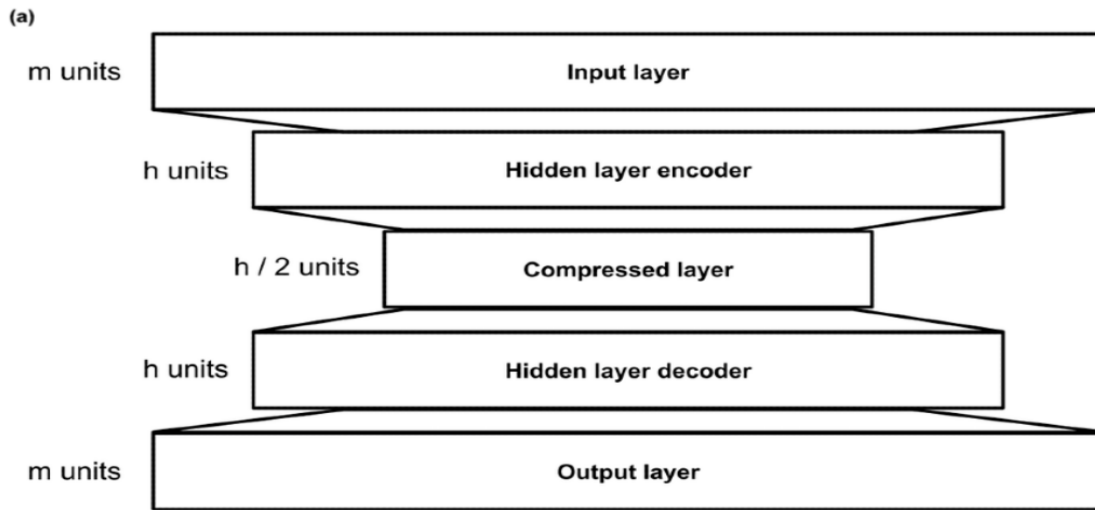
τυο που λαμβάνει ως είσοδο τα τρία σήματα από τους αισθητήρες και επιστρέφει το Person ID. Η αρχιτεκτονική του νευρωνικού δίκτυα περιλαμβάνει τρία επίπεδα. Στο πρώτο επίπεδο τα δεδομένα υποβάλλονται σε επεξεργασία από ένα μπλοκ συνελικτικών επιπέδων μιας διάστασης (1D convolution layer) αφού τα δεδομένα που εξήχθησαν από το smartwatch δεν είναι συγχρονισμένα και μπορεί να εμφανίζουν μικρές χρονικές μεταβολές μεταξύ τους, επομένως και διαφορετικό μήκος σήματος. Στο 2ο επίπεδο πραγματοποιείται συγχώνευση των δεδομένων. Μετά τη συγχώνευση εφαρμόζεται το μπλοκ τελικής συνέλιξης, που αποτελείται από πολλαπλά συνελικτικά επίπεδα μιας διάστασης μαζί με batch normalization και μια μη γραμμικότητα τη ReLU. Όλα τα συνελικτικά δίκτυα έχουν πυρήνα μήκους 5. Τα blocks περιέχουν dropout layers μετά από κάθε ReLU για την αποφυγή του overfitting. Τέλος, εφαρμόζεται average pooling ακολουθούμενο από ένα πλήρως συνδεδεμένο επίπεδο, για τη μείωση του μήκους της ακολουθίας. Παρατηρήθηκε ότι η κατάσταση του περπατήματος έχει τα καλύτερα αποτελέσματα ενώ τα χειρότερα είναι του ύπνου. Όταν οι ερευνητές δοκίμασαν να ανταλλάξουν ρολόγια σε δύο εθελοντές για δύο επιπλέον μέρες είχαν ως αποτέλεσμα τα διαστήματα με εντατική κίνηση να ταξινομηθούν σωστά, ενώ τα διαστήματα χωρίς κίνηση (π.χ., ύπνος) να ταξινομούνται πολύ συχνά στον προηγούμενο ιδιοκτήτη του ρολογιού. Αυτό οφείλεται στο θόρυβο των αισθητήρων. Για τη λύση αυτού του προβλήματος οι συγγραφείς αποφάσισαν να προσθέσουν στα δεδομένα επιπλέον τεχνητό θόρυβο ικανό να καλύψει τον θόρυβο των αρχικών αισθητήρων, με αποτέλεσμα να δημιουργηθούν δίκτυα ανεξάρτητα από τους αισθητήρες.

Μια πρόσφατη μελέτη [Bir+19] χρησιμοποίησε δεδομένα από δημοσιεύσεις στο Facebook για την ανίχνευση γλωσσικών και συμπεριφορικών αλλαγών που οδηγούν στην ψυχωτική υποτροπή, σε άτομα με ψυχωτικές διαταραχές. Τα δεδομένα προέρχονται από 51 συμμετέχοντες που είχαν υποστεί πρόσφατα ψυχωτική υποτροπή. Για την ανίχνευση της υποτροπής δημιουργήθηκε ένας SVM ταξινομητής μίας κατηγορίας (1-class SVM Classifier), ο οποίος ταξινομεί τις περιόδους υγείας και σχετικής υγείας (δηλαδή 1 μήνα πριν την υποτροπή) εξατομικευμένα για κάθε ασθενή. Ο ταξινομητής αυτός πέτυχε ειδικότητα (specificity) 0,71 στην πρόβλεψη υποτροπής. Παρατηρήθηκε ότι ένα μήνα πριν την υποτροπή οι ασθενείς στις αναρτήσεις τους στο Facebook, χρησιμοποιούν αυξημένα λέξεις που ανήκουν στις βρισιές, στο θυμό και κατηγορίες θανάτου και χρησιμοποιούν μειωμένα λέξεις που ανήκουν στις κατηγορίες εργασία, φίλος και υγεία.

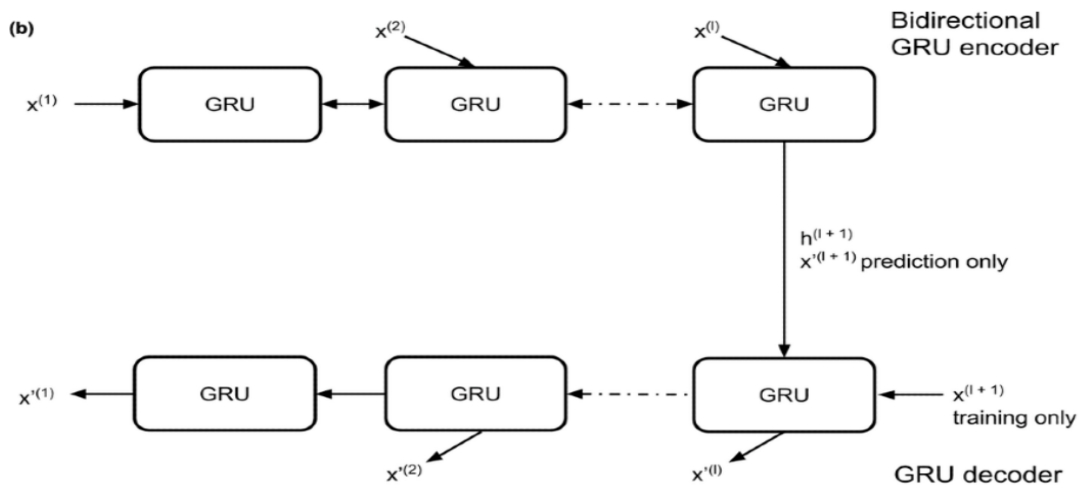
Στη έρευνα [Adl+20], οι ερευνητές δημιούργησαν μια ποικιλία μοντέλων ανίχνευσης ανωμαλιών βασισμένη στο νευρωνικό δίκτυο κωδικοποιητή-αποκωδικοποιητή (Encoder-Decoder) με σκοπό να προβλέψουν έγκαιρα προειδοποιητικά σημάδια ψυχωτικής υποτροπής χρησιμοποιώντας δεδομένα παθητικής ανίχνευσης που συλλέγονται από ένα smartphone από 60 άτομα με ψυχική διαταραχή για διάρκεια 12 μηνών. Τα παθητικά δεδομένα που χρησιμοποιήθηκαν από το smartphone ήταν μετρήσεις από τη επιτάχυνση, τις τοποθεσίες, τη διάρκεια του ύπνου, πόσες εφαρμογές χρησιμοποιήθηκαν, τη διάρκεια και τον αριθμό των συνομιλιών, διάρκεια χρήσης του κινητού και τη κοινωνική επίδραση με μηνύματα του ασθενή. Επιπλέον χρησιμοποιήθηκε ως χαρακτηριστικό η ημέρα της εβδομάδας και η ώρα της ημέρας. Τα παραπάνω δεδομένα συλλέχθηκαν μέσω του συστήματος CrossCheck [Wan+16], ενός συστήματος ανίχνευσης παθητικών δεδομένων από smartphones.

Στη συγκεκριμένη εργασία, αναπτύχθηκαν δύο νευρωνικά δίκτυα, ένα πλήρως συνδεδεμένο μοντέλο αυτόματου κωδικοποιητή νευρωνικών δικτύων (FNN AD) και ένα μοντέλο GRU Seq2Seq [Bal12], τα οποία εκπαιδεύτηκαν για να μάθουν όρια ανωμαλίας για συγκεκριμένους συμμετέχοντες με βάση το σφάλμα ανακατασκευής του μοντέλου. Συγκεκριμένα, στα δύο μοντέλα ως εί-

σοδος εισάγεται μια ακολουθία χαρακτηριστικών ανά ώρα για κάθε συμμετέχοντα. Το δίκτυο FNN AD στην έξοδο του προβλέπει τη ανακατασκευασμένη ακολουθία χαρακτηριστικών που πρέπει να είναι ίδια με τη αρχική ενώ το GRU Seq2Seq προβλέπει τη αντίστροφη ανακατασκευασμένη ακολουθία. Για την εκπαίδευση του μοντέλου τα δεδομένα κάθε συμμετέχοντα χωρίστηκαν



Σχήμα 1.3.1: Αρχιτεκτονική FNN AD [Adl+20]



Σχήμα 1.3.2: Αρχιτεκτονική GRU sequence-to-sequence [Adl+20]

σε 4 σύνολα. Το πρώτο σύνολο δεδομένων είναι το training set και αποτελείται από δεδομένα που ανήκουν σε ημέρες που δεν είναι κοντά στη περίοδο υποτροπής. Το δεύτερο σύνολο δεδομένων είναι το cross-validation set που περιλαμβάνει δεδομένα που ανήκουν σε ημέρες που δεν είναι κοντά στη περίοδο υποτροπής, όπως στο προηγούμενο σύνολο. Το τρίτο σύνολο δεδομένων είναι και αυτό ένα cross-validation set αλλά αποτελείται από κάποια δεδομένα που ανήκουν σε ημέρες που δεν είναι κοντά στη περίοδο υποτροπής και κάποια άλλα που ανήκουν σε ημέρες κοντά στη περίοδο υποτροπής. Τέλος, το τέταρτο σύνολο δεδομένων είναι το test

set που περιλαμβάνει δεδομένα που ανήκουν τόσο στις ημέρες που δεν είναι κοντά στη περίοδο υποτροπής όσο και στις ημέρες κοντά στη υποτροπή.

Στη συνέχεια, γίνεται η εκπαίδευση του μοντέλου χρησιμοποιώντας το training set, δηλαδή τα δεδομένων που ανήκουν σε ημέρες φυσιολογικές, μακριά από τη περίοδο υποτροπής. Το εκπαιδευμένο μοντέλο εφαρμόζεται στα υπόλοιπα σύνολα δεδομένων για να τα ανακατασκευάσει. Για κάθε σύνολο δεδομένων υπολογίζεται το απόλυτο σφάλμα μεταξύ των στοιχείων  $x(i)$  της ακολουθίας εισόδου και των αντίστοιχων στοιχείων  $x'(i)$  του ανακατασκευασμένου μέρους ( $e(i) = |x(i) - x'(i)|$ ). Ακολουθώντας, υπολογίζεται ανά ώρα η βαθμολογία ανωμαλίας  $s^i$  για τα διανύσματα σφάλματος με χρήση της απόστασης Mahalanobis. Επίσης, ανά ημέρα υπολογίζεται η μέση βαθμολογία ανωμαλίας  $s^d$  από το μέσο όρο των βαθμολογιών ανωμαλίας ανά ώρα. Μία μέρα θεωρείται ότι είναι ανώμαλη όταν  $s^d > \tau$  ενώ όταν το  $s^d \leq \tau$  θεωρείται φυσιολογική. Η επιλογή του κατωφλιού  $\tau$  έγινε εφαρμόζοντας τα δεδομένα του cross-validation set έτσι ώστε να βρεθεί το βέλτιστο κατώφλι  $\tau$  για κάθε συμμετέχοντα. Ορίστηκε ώστε να μεγιστοποιήσει τη ποσότητα  $TPR/FPR$  όπου  $TPR$  το True Positive Rate και  $FPR$  το False Positive Rate. Το  $TPR$  ορίζεται ως η ποσότητα του πόσες μέρες ανιχνεύτηκαν σωστά ως ανώμαλες και το  $FPR$  ορίζεται ως η ποσότητα του πόσες μέρες ανιχνεύτηκαν λάθος ως ανώμαλες. Για να θεωρείται μια μέρα ότι ανιχνεύθηκε σωστά ως ανώμαλη πρέπει τα δεδομένα να ανήκουν στο σύνολο των ημερών κοντά σε υποτροπή ενώ για να θεωρείται μια μέρα ότι ανιχνεύτηκε λάθος ως ανώμαλη πρέπει τα δεδομένα να ανήκουν στο σύνολο των ημερών που δεν είναι κοντά στη υποτροπή.

Το μοντέλο FNN AE είχε την καλύτερη απόδοση και πέτυχε μια μέση ευαισθησία 0.25 (IQR 0.15 – 1.00) και ειδικότητα 0.88 (IQR 0.14 – 0.96). Για την επιτυχή όμως υλοποίηση ενός νευρωνικού δικτύου που βασίζεται σε ένα σύστημα ανίχνευσης υποτροπών εντός της συμπεριφορικής υγείας, πρέπει κανείς να δείξει όχι μόνο καλά αποτελέσματα στην ανίχνευση, αλλά και να παρέχει μια διαδικασία για την ερμηνεία των συμπεριφορών που οδηγούν σε ανωμαλία/υποτροπή και την παροχή κλινικής ερμηνείας για αυτές τις συμπεριφορές. Επομένως, έγινε μία εκ των υστέρων ανάλυση για τη κλινική ερμηνεία των ανιχνευμένων ανωμαλιών στο πλαίσιο των διαταραχών του φάσματος της σχιζοφρένειας. Παρατηρήθηκε ότι μόλις ο αλγόριθμος εντοπίσει τις ανώμαλες ημέρες, μπορεί επίσης να εντοπίσει μία συγκεκριμένη συμπεριφορά για κάθε ασθενή. Αυτή η προσέγγιση θα μπορούσε να χρησιμοποιηθεί έτσι ώστε στο μέλλον να εκπαιδευτούν εξατομικευμένα μοντέλα για τη ανίχνευση της συγκεκριμένης συμπεριφοράς για κάθε ασθενή που έχει ως επακόλουθο τη ανίχνευση της υποτροπής [Adl+20].

Στην εργασία μας, θα χρησιμοποιηθούν δεδομένη χρονοσειρών από πολλαπλούς αισθητήρες. Τα δεδομένα χρονοσειρών είναι διαδομένα σε πολλούς επιστημονικούς και μηχανικούς κλάδους. Η πρόβλεψη χρονοσειρών είναι ένα κρίσιμο έργο στη μοντελοποίηση δεδομένων χρονοσειρών και είναι ένας σημαντικός τομέας της μηχανικής μάθησης. Στο άρθρο [Wu+20] παρουσιάζεται μια νέα προσέγγιση στη πρόβλεψη χρονοσειρών, που μας έχει εμπνεύσει και για την δική μας εργασία, η οποία βασίζεται στον transformer. Η αρχιτεκτονική αυτή χρησιμοποιεί μηχανισμούς αυτοπροσοχής (self-attention mechanisms) για την εκμάθηση πολύπλοκων μοτίβων και δυναμικών από τα δεδομένα χρονοσειρών. Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση του μοντέλου είναι ο λόγος ILI (Influenza Like Illness). Ο λόγος ILI είναι ουσιαστικά ο αριθμός των ασθενών που έχουν συμπτώματα γρίπης ως προς τον συνολικό αριθμό επισκέψεων των ασθενών στους ιατρούς για εκείνη την εβδομάδα, τα οποία υπάρχουν για όλες τις πολιτείες των Ηνωμένων Πολιτειών Αμερικής (ΗΠΑ). Το μοντέλο εκπαιδεύεται για να προβλέπει 4 μελλοντικές εβδομαδιαίες αναλογίες ILI από 10 εβδομαδιαία σημεία δεδομένων. Η αρχιτεκτονική transformer (η οποία θα



περιγραφή στην Ενότητα 2.4.4) είχε την καλύτερη επίδοση, σε σύγκριση με άλλα μοντέλα όπως ARIMA, LSTM, Seq2Seq with attention, με 0.928 και 0.588 Pearson Correlation και RMSE, αντίστοιχα.

## 1.4 Κίνητρα και Συνεισφορές

Το κύριο κίνητρο αυτής της διπλωματικής εργασίας, είναι η βοήθεια που μπορούμε να προσφέρουμε σε ασθενείς που πάσχουν από διαταραχές του φάσματος της σχιζοφρένειας, ανιχνεύοντας πότε έχουν υποτροπή με μελλοντική επέκταση στην πρόβλεψη της με σκοπό την αποφυγή της. Το παρόν πρόβλημα θεωρούμε ότι είναι μεγάλης καινοτομίας και σημαντικότητας, μιας και οι παραπλήσιες έρευνες είναι ελάχιστες και αφενός χρησιμοποιούν διαφορετικού τύπου δεδομένα, όπως δεδομένα από κινητά τηλέφωνα και αφετέρου τα δεδομένα αυτά στις περισσότερες έρευνες είναι βραχυπρόθεσμα, δηλαδή διαρκούν για ένα μικρό χρονικό διάστημα μιας εβδομάδας ή κάποιον ημερών/ωρών. Σε αντίθεση, στη δική μας μελέτη χρησιμοποιούμε παθητικά δεδομένα μακροχρόνιας συνεχούς παρακολούθησης περίπου 2 χρόνων, που συλλέχθηκαν με τη χρήση smartwatch μέσω του έργου e-Prevention ήταν το έναυσμα για την έναρξη αυτής της διπλωματικής. Στηριζόμενοι στις ενδείξεις από έρευνες όπως [Fil+20; Mag+20], στις οποίες έχουν εντοπιστεί διάφορα χαρακτηριστικά που επηρεάζουν τους ασθενείς με ψυχωτικές διαταραχές, χρησιμοποιούμε και επεκτείνουμε κάποια από αυτά με σκοπό να ανιχνεύσουμε υποτροπές. Επιπλέον, λαμβάνοντας υπόψη ότι οι ιατροί χαρακτηρίζουν τη σοβαρότητα της υποτροπής κάθε ασθενή (ως low, moderate ή severe), μας βοηθάει να εκμεταλλευτούμε αυτή την πληροφορία για να ελέγξουμε τα μοντέλα μας αν επηρεάζονται στην ανίχνευση της υποτροπής ανάλογα της σοβαρότητας της.

Οι συνεισφορές της διπλωματικής εργασίας, μπορούν να χωριστούν σε τρεις κατηγορίες, στα δεδομένα, την προεπεξεργασία δεδομένων και τα νευρωνικά δίκτυα οι οποίες αναλύονται ως εξής:

- Το καινοτόμο σύνολο δεδομένων της μελέτης μας, αποτελείται από μεγάλο όγκο δεδομένων που συλλέγονται καθημερινά εδώ και δύο χρόνια συνεχούς παρακολούθησης μέσω αισθητήρων από έξυπνα ρολόγια.
- Εξάγουμε τα ακατέργαστα δεδομένα μας από το smartwatch και εφαρμόζοντας μεθόδους επεξεργασίας σήματος εξάγουμε χρήσιμες αναπαραστάσεις και χαρακτηριστικά όπως η μεταβλητότητα του καρδιακού ρυθμού κ.α. Επιπλέον, δημιουργήσαμε ένα καινούργιο σύνολο δεδομένων με βάση τα εξαγόμενα χαρακτηριστικά, το οποίο αποτελείται από δεδομένα 10 ασθενών, και το επεξεργαστήκαμε με σκοπό να δημιουργήσουμε συνεχόμενες ακολουθίες χρονοσειρών οι οποίες δεν έχουν χρονικές απώλειες.
- Προσεγγίσαμε την επίλυση του προβλήματος μας ως anomaly detection πρόβλημα, με σκοπό να εντοπίσουμε τις υποτροπές σαν ανωμαλίες. Έτσι, αναπτύξαμε τέσσερις autoencoder αρχιτεκτονικές, με σκοπό να μάθουν τις φυσιολογικές συμπεριφορές των ασθενών έτσι ώστε να μπορούν να ανιχνεύουν τις μη φυσιολογικές (υποτροπές). Εφαρμόσαμε πειραματική αξιολόγηση των προτεινόμενων μεθόδων για την ανίχνευση υποτροπών με τη χρήση των autoencoder αρχιτεκτονικών και εξάγαμε ενδιαφέροντα αποτελέσματα μέσα από εκτεταμένα πειράματα για την ανίχνευση των υποτροπών. Συγκεκριμένα, παρατηρήσαμε ότι οι αρχιτεκτονικές Fully Connected και Convolutional Autoencoder

έχουν γενικότερα καλύτερες αποδόσεις. Επιπρόσθετα, οι αναπαραστάσεις των χαρακτηριστικών που εξάγαμε μας δίνουν καλύτερα αποτελέσματα από τα ακατέργαστα δεδομένα. Αξίζει να σημειωθεί, ότι οι σοβαρότητα της υποτροπής μας μας δίνει σημαντική πληροφορία σχετικά με την ανίχνευση της υποτροπής. Τέλος, καταλήξαμε στο συμπέρασμα ότι ο όγκος και το πλήθος των ασθενών είναι ο σημαντικότερος παράγοντας για να έχουμε αξιόπιστα αποτελέσματα.

## 1.5 Διάρθρωση της Διπλωματικής εργασίας

Η δομή της διπλωματικής εργασίας οργανώνεται ως εξής:

- **Κεφάλαιο 2:** Στο κεφάλαιο 2 παρουσιάζονται έννοιες σχετικές με τα θεωρητικά εργαλεία που θα χρησιμοποιηθούν για τη επίλυση του προβλήματος. Συγκεκριμένα, θα αναλυθούν έννοιες και μεθοδολογίες επεξεργασίας σήματος, όπως γραμμική και μη γραμμική ανάλυση σήματος. Επίσης, θα αναλυθούν έννοιες νευρωνικών δικτύων και αρχιτεκτονικών βαθιάς μάθησης, όπως πλήρως συνδεδεμένα νευρωνικά δίκτυα, συνελικτικά νευρωνικά δίκτυα, αναδρομικά νευρωνικά δίκτυα και transformer, τα οποία χρησιμοποιήθηκαν για την υλοποίηση αυτής της εργασίας.
- **Κεφάλαιο 3:** Στο κεφάλαιο 3 παρουσιάζεται η συλλογή παθητικών δεδομένων από έξυπνα ρολόγια, για ασθενείς που πάσχουν από ψυχικές διαταραχές. Ακολούθως, γίνεται μία ανάλυση των δεδομένων και εξαγωγή χαρακτηριστικών, τα οποία θα χρησιμοποιήσουμε για την εκπαίδευση των μοντέλων βαθιάς μάθησης που έχουμε υλοποίηση. Στη συνέχεια, περιγράφεται η προεπεξεργασία των δεδομένων, δηλαδή οι τεχνικές που χρησιμοποιήθηκαν για να συμπληρώσουμε τα κενά που είχαμε στα δεδομένα, όπως επίσης και η δημιουργία των πολυμεταβλητών χρονοσειρών μας για το τελικό σύνολο δεδομένων. Το τελικό σύνολο δεδομένων θα χρησιμοποιηθεί με απώτερο σκοπό την εκπαίδευση και αξιολόγηση των μοντέλων μας για την ανίχνευση υποτροπών σε ασθενείς με ψυχικές διαταραχές.
- **Κεφάλαιο 4:** Στο κεφάλαιο 4 παρουσιάζεται η μέθοδος που σχεδιάστηκε και υλοποιήθηκε για την ανίχνευση ανωμαλιών. Δηλαδή, περιγράφονται οι τέσσερις αρχιτεκτονικές βαθιάς μάθησης που υλοποιήθηκαν, τα πλήρως συνδεδεμένα νευρωνικά δίκτυα αυτόματου κωδικοποιητή (Autoencoder - AE), τα συνελικτικά νευρωνικά δίκτυα AE, την GRU AE και τον transformer. Επίσης, παρουσιάζεται η διαδικασία εκπαίδευσης των μοντέλων μας, οι υπερπαραμέτροι που επιλέχθηκαν και η μέθοδος ανίχνευσης ανωμαλιών.
- **Κεφάλαιο 5:** Στο κεφάλαιο 5 αρχικά παρουσιάζονται οι μετρικές αξιολόγησης που χρησιμοποιήθηκαν για την αξιολόγηση ολόκληρης της μεθόδου που παρουσιάστηκε. Στη συνέχεια, πειραματιστήκαμε με εξατομικευμένα και γενικά μοντέλα πάνω στα raw δεδομένα αλλά και πάνω στα επεξεργασμένα χαρακτηριστικά. Δοκιμάσαμε διαφορετικά granularities στα δεδομένα και διαφορετικούς τρόπους ανίχνευσης ανωμαλιών. Τα αποτελέσματά μας είναι υποσχόμενα, με τα δύο από τα τέσσερα μοντέλα να ξεχωρίζουν. Τέλος, γίνεται μια διεξοδική ανάλυση για να ελέγξουμε πόσο επηρεάζει η σοβαρότητα της υποτροπής, τον εντοπισμό της.
- **Κεφάλαιο 6:** Στο κεφάλαιο 6 παρουσιάζονται τα συμπεράσματα και πιθανές μελλοντικές επεκτάσεις αυτής της διπλωματικής εργασίας.

# Κεφάλαιο 2

## Θεωρητικό Υπόβαθρο

### 2.1 Επεξεργασίας Σήματος

Η Επεξεργασία Σήματος ασχολείται με την αναπαράσταση, τη μεταφορά και το χειρισμό των σημάτων και της πληροφορίας που περιέχουν. Συγκεκριμένα, στη Ψηφιακή Επεξεργασία Σήματος (ΨΕΣ ή Digital Signal Processing - DSP) τα σήματα αναπαρίστανται από ακολουθίες αριθμών πεπερασμένης ακρίβειας και η επεξεργασία υλοποιείται με ψηφιακούς υπολογισμούς. Η χρήση της DSP οδηγεί στην πρόοδο σε πολλούς τομείς της ιατρικής. Στη δική μας περίπτωση, με τη βοήθεια της επεξεργασίας σήματος, από τα ακατέργαστα δεδομένα αισθητήρων οδηγούμαστε στις αναπαραστάσεις συμπεριφορών και ψυχικών καταστάσεων.

Ως σήμα (signal) ορίζεται κάθε μέγεθος το οποίο περιέχει πληροφορία. Μαθηματικά, τα σήματα παριστάνονται ως συναρτήσεις μιας ή περισσότερων ανεξάρτητων μεταβλητών. Τα χρονοσυνεχή σήματα, ορίζονται στο συνεχές σύνολο του χρόνου και παριστάνονται από συνεχή ανεξάρτητη μεταβλητή. Στο πραγματικό κόσμο τα περισσότερα σήματα είναι χρονοσυνεχή, παρόλο που στις περισσότερες περιπτώσεις είναι προτιμότερο να εργαζόμαστε με ψηφιακά σήματα που μπορούν να επεξεργαστούν από τους υπολογιστές [Opp99].

#### 2.1.1 Θεωρία Πιθανοτήτων και Στατιστικής

Τα σήματα που παράγονται από ένα αισθητήρα (όπως ο καρδιακός ρυθμός) είναι τυχαία σήματα (random signal). Δεδομένου ότι μια τυχαία διαδικασία ή ένα τυχαίο σήμα μπορεί να περιγραφεί μόνο πιθανολογικά, θα αναλύσουμε κάποια βασική θεωρία πιθανοτήτων.

#### Τυχαίες Μεταβλητές

Η Συνάρτηση πυκνότητας πιθανότητας (Probability density function - PDF) μιας τυχαίας μεταβλητής  $X$  συμβολίζεται με  $f_X(x)$  και χρησιμοποιείται για να καθορίσει την πιθανότητα η τυχαία μεταβλητή να εμπίπτει σε ένα συγκεκριμένο εύρος τιμών, σε αντίθεση με τη λήψη οποιασδήποτε τιμής. Αυτή η πιθανότητα δίνεται από το ολοκλήρωμα της PDF αυτής της μεταβλητής σε αυτό το εύρος, Εξίσωση 2.1.1. Η συνάρτηση πυκνότητας πιθανότητας είναι παντού μη αρνητική

και το ολοκλήρωμα της σε ολόκληρο τον χώρο είναι ίσο με 1, Εξίσωση 2.1.2.

$$F_X(x) = \int_{-\infty}^{\infty} f_X(x) dx \quad (2.1.1)$$

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (2.1.2)$$

Για την αξιολόγηση της συμπεριφοράς τυχαίων πειραμάτων χρησιμοποιούνται οι μετρικές της αναμενόμενης τιμής (mean) και της διακύμανσης (variance). Η αναμενόμενη τιμή  $\mu(X)$  (expected value, mean) μιας τυχαίας μεταβλητής  $X$  ορίζεται ως:

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx \quad (2.1.3)$$

Η διακύμανση (variance) μιας τυχαίας μεταβλητής  $X$  ορίζεται ως:

$$\text{var}[X] = \sigma_X^2 = E[(x - \mu_X)^2] = \int_{-\infty}^{\infty} (X - \mu_x)^2 f_X(x) dx \quad (2.1.4)$$

Το  $\sigma_X$  ονομάζεται τυπική απόκλιση (standard deviation) της τυχαίας μεταβλητής  $X$ . Διαισθητικά, η διακύμανση και η τυπική απόκλιση μετρούν τον βαθμό της τυχειότητας της μεταβλητής  $X$ .

### Gaussian Τυχαίες μεταβλητές

Η Gaussian τυχαίες μεταβλητές έχουν κύριο ρόλο στη θεωρία πιθανοτήτων. Η PDF μιας gaussian τυχαίας μεταβλητής  $x$  ορίζεται ως:

$$f_X(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right\} \quad (2.1.5)$$

όπου  $\mu_x$  και  $\sigma_x^2$  είναι το mean και το variance της μεταβλητής  $x$  αντίστοιχα [Hay09].

## 2.1.2 Γραμμική Ανάλυση (Linear Analysis)

### Ανάλυση Βραχέος Χρόνου (Short-time analysis)

Οποιαδήποτε βραχέος χρόνου ανάλυση (short-time analysis)  $T\{\}$  του σήματος  $x[m]$  εφαρμόζεται ένα παράθυρο  $w[n]$  στο σήμα:

$$x_{\hat{n}} = \sum_{m=-\infty}^{\infty} T\{(x[m]w[\hat{n} - m])\} \quad (2.1.6)$$

όπου:  $x_{\hat{n}}$ : short-time signal,  $w[\hat{n} - m]$ : παράθυρο,  $n$ : το δείγμα στο οποίο επικεντρώνεται το παράθυρο ανάλυσης.

### Short-time Energy

Η ενέργεια βραχέος χρόνου (Short-time Energy - STE) είναι η ενέργεια του σήματος, δηλαδή είναι μια απλή και αποτελεσματική παράμετρος ταξινόμησης αν έχω κάποια 'δράση' στο σήμα.

Η STE ενός σήματος  $x$  εκφράζεται με τη βοήθεια ενός κινούμενου παραθύρου  $w$  σύμφωνα με την εξίσωση:

$$E_{\hat{n}} = \sum_{m=-\infty}^{\infty} (x[m]w[\hat{n} - m])^2 \quad (2.1.7)$$

### Short-time autocorrelation function

Η αυτοσυσχέτιση (autocorrelation) χρησιμοποιείται για την εύρεση περιοδικοτήτων σε συναρτήσεις και είναι η βάση πολλών μεθόδων ανάλυσης φάσματος (spectrum analysis). Μαθηματικά εκφράζεται ως:

$$\phi_{\hat{n}}[l] = \sum_{m=-\infty}^{\infty} x[m]\hat{w}_l[\hat{n} - m] \quad (2.1.8)$$

όπου,  $\hat{w}_l[m] = w[m]w[m + l]$ .

Να σημειώσουμε ότι η αυτοσυσχέτιση στο lag 0 είναι ίση με το short-time energy, άρα:

$$E_{\hat{n}} = \sum_{m=-\infty}^{\infty} (x[m]w[\hat{n} - m])^2 = \phi_{\hat{n}}[0] \quad (2.1.9)$$

### Short-time Fourier Transform

Ο μετασχηματισμός Fourier μας δίνει το φάσμα (spectrum) των συχνοτήτων ενός σήματος. Ο βραχυχρόνιος μετασχηματισμός Fourier (STFT) [Rab78] είναι μια ακολουθία μετασχηματισμών Fourier ενός παραθυρόμενου σήματος. Ο STFT παρέχει τις χρονικά εντοπισμένες πληροφορίες συχνότητας (time-localized frequency) για καταστάσεις στις οποίες τα στοιχεία συχνότητας ενός σήματος ποικίλλουν με την πάροδο του χρόνου, ενώ ο τυπικός μετασχηματισμός Fourier παρέχει τις πληροφορίες συχνότητας που υπολογίζονται κατά μέσο όρο σε ολόκληρο το χρονικό διάστημα του σήματος. Ο Short-Time Fourier Transform (STFT) ορίζεται ως:

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m)w(n - m)e^{-j\omega m} \quad (2.1.10)$$

### Spectral Analysis

Πολλές χρονοσειρές δείχνουν περιοδική συμπεριφορά. Αυτή η περιοδική συμπεριφορά μπορεί να είναι πολύ περίπλοκη. Η φασματική ανάλυση (spectral analysis) είναι μια τεχνική που μας επιτρέπει να ανακαλύψουμε υποκείμενες περιοδικότητες. Για να πραγματοποιήσουμε φασματική ανάλυση, πρέπει πρώτα να μετατρέψουμε δεδομένα από το πεδίο του χρόνου στο πεδίο της συχνότητας. Συγκεκριμένα, η φασματική ανάλυση διακρίνει τρία στοιχεία μιας χρονοσειράς: την τάση (trend) (μακροπρόθεσμη αλλαγή), τις κυκλικές ή ταλαντευόμενες συναρτήσεις και τα τυχαία ή ακανόνιστα στοιχεία. Επίσης, διασπά τη σειρά σε αθροιστικές ή απλές κυματομορφές που περιγράφονται από μια σειρά ημιτονοειδών και συνημιτονικών κυμάτων [CC08].

Στη δική μας μελέτη χρησιμοποιήθηκε το spectral χαρακτηριστικό της μεταβλητότητας του καρδιακού ρυθμού (HRV). Το HRV είναι το φυσιολογικό φαινόμενο της διακύμανσης των καρδιακών παλμών. Δεδομένου ότι το HRV είναι ένα εξ ορισμού μη ομοιόμορφα δειγματοληπτικό

σήμα, εκτελέστηκε φασματική ανάλυση χρησιμοποιώντας το περιοδόγραμμα του Lomb-Scargle [Fil+20].

Το περιοδόγραμμα του Lomb-Scargle [Sca82] είναι ένας πολύ γνωστός αλγόριθμος για την ανίχνευση και τον χαρακτηρισμό της περιοδικότητας σε ανομοιόμορφες δειγματοληπτικές χρονοσειρές. Το περιοδόγραμμα Lomb-Scargle είναι μια μέθοδος που επιτρέπει τον αποτελεσματικό υπολογισμό ενός εκτιμητή φάσματος ισχύος τύπου Fourier από ανομοιόμορφα δειγματοληπτικά δεδομένα, με αποτέλεσμα ένα διαισθητικό μέσο προσδιορισμού της περιόδου ταλάντωσης. Το περιοδόγραμμα Lomb-Scargle ορίζεται ως εξής:

$$P_{LS}(\Omega) = \frac{1}{2} \left\{ \frac{[\sum_{n=0}^{N-1} x[n] \cos(\Omega(t_n - \tau))]^2}{\sum_{n=0}^{N-1} \cos^2(\Omega(t_n - \tau))} + \frac{[\sum_{n=0}^{N-1} x[n] \sin(\Omega(t_n - \tau))]^2}{\sum_{n=0}^{N-1} \sin^2(\Omega(t_n - \tau))} \right\} \quad (2.1.11)$$

όπου το  $\tau$  ορίζεται ως:

$$\tau = \frac{1}{2\Omega} \tan^{-1} \left( \frac{\sum_{n=0}^{N-1} \sin(2\Omega t_n)}{\sum_{n=0}^{N-1} \cos(2\Omega t_n)} \right) \quad (2.1.12)$$

το  $\Omega$  είναι η γωνιακή συχνότητα (rad/sec), το  $t_n$  η χρονική στιγμή (sec) δειγματοληψίας του σήματος, και το  $x[n]$  η τιμή του σήματος στο χρονικό σημείο  $t_n$ .

## 2.1.3 Μη Γραμμική Ανάλυση (Non-Linear Analysis)

### Fractal Signal Analysis

Ο όρος φράκταλ (Fractal) [Goo78] προέρχεται από τη λατινική λέξη fractum, που σημαίνει σπασμένο για να περιγράψει ‘ασυνήθιστα’ σχήματα που δεν μπορούν να μοντελοποιηθούν γεωμετρικά. Τα fractals είναι απείρως πολύπλοκα μοτίβα που είναι ίδια σε διαφορετικές κλίμακες (self-similarity). Δημιουργούνται επαναλαμβάνοντας μια απλή διαδικασία ξανά και ξανά σε έναν συνεχή βρόχο ανατροφοδότησης. Δεδομένου ότι οι παραδοσιακές γεωμετρικές μετρήσεις ήταν ανεπαρκείς στη μοντελοποίηση μιας τέτοιας συμπεριφοράς, εισήχθησαν αλγόριθμοι και μετρήσεις fractal για την αντιμετώπιση της ιδιότητας του self-similarity. Ωστόσο, εκτός από αυτά στα ιδανικά φράκταλ, πολλές φυσικές διεργασίες και σήματα παρουσιάζουν παρόμοιες ιδιότητες. Πολλά σήματα όπως και τα βιοϊατρικά σήματα [Eke+02] δείχνουν μια πολύπλοκη δομή σε όλες τις χρονικές κλίμακες.

### Fractal Dimension

Η διάσταση φράκταλ (fractal dimension) είναι ένα μέτρο του πόσο κατακερματισμένο είναι ένα φράκταλ αντικείμενο και μπορεί να γίνει κατανοητός ως ο χαρακτηρισμός της ομοιότητάς του με τον εαυτό του [MM82]. Ορίζεται ως ένας μη ακέραιος αριθμός που ποσοτικοποιεί την πυκνότητα των fractals στο μετρικό χώρο και είναι ένας τρόπος να προσδιορίσουμε πόσο πολύπλοκο είναι ένα fractal, προκειμένου να το συγκρίνετε με ένα άλλο. Η γεωμετρία ενός φράκταλ έχει διάφορες προσεγγίσεις για τον υπολογισμό της διάστασης φράκταλ ενός αντικειμένου. Για τη μέτρηση της διάστασης fractal μιας χρονοσειράς έχουν προταθεί πολλοί αλγόριθμοι όπως ο Higuchi Fractal Dimension (HFD) [Hig88] και ο Multi-scale Fractal Dimension (MFD) [Mar94].

**Higuchi Fractal Dimension** Η Higuchi Fractal Dimension (HFD) προτάθηκε από τους Higuchi et al. [Hig88] ως προσέγγιση για τη διάσταση μέτρησης πλαισίων (box-counting dimension) του γραφήματος μιας συνάρτησης ή μιας χρονολογικής σειράς με πραγματική αξία.



Υπάρχουν πολλές εφαρμογές της στην επιστήμη και τη μηχανική και έχει εφαρμοστεί σε θέματα όπως στη κλινική νευροφυσιολογία [KS16]. Για δοσμένη χρονοσειρά  $X : 1, \dots, N \rightarrow \mathbb{R}$  που αποτελείται από  $N$  σημεία δεδομένων και μια παράμετρο  $k_{max} \geq 2$  η Higuchi Fractal Dimension του  $X$  υπολογίζεται ως εξής: Για κάθε  $k \in \{1, \dots, k_{max}\}$  και  $m \in \{1, \dots, k\}$  ορίζουμε το μήκος  $L_m(k)$  ως:

$$L_m(k) = \frac{N-1}{\lfloor \frac{N-m}{k} \rfloor k^2} \sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |X_N(m+ik) - X_N(m+(i-1)k)| \quad (2.1.13)$$

Το μήκος  $L(k)$  ορίζεται ως ο μέσος όρος των  $k$  μηκών  $L_1(k), \dots, L_k(k)$ ,

$$L(k) = \frac{1}{k} \sum_{m=1}^k L_m(k) \quad (2.1.14)$$

Η κλίση της γραμμικής συνάρτησης που ταιριάζει καλύτερα στα σημεία δεδομένων  $\{(\log \frac{1}{k}, \log L(k))\}$  ορίζεται ως η φράκταλ διάσταση Higuchi της χρονοσειράς  $X$ .

### Multi-scale fractal Dimension

Οι μέθοδοι μέτρησης της fractal διάστασης που περιγράφηκαν μέχρι τώρα δεν μπορούν να αντεπεξέλθουν επαρκώς σε περίπλοκα σχήματα όπως αυτά του πραγματικού κόσμου αφού δεν έχουν την ίδια δομή σε διαφορετικές χρονικές κλίμακες. Επομένως, για να εξεταστεί η πολύκλιμακτη ανάλυση για τη μοντελοποίηση τέτοιων σημάτων αναπτύχθηκε ένας αποτελεσματικός αλγόριθμος για τη μέτρηση μιας πολλαπλής κλίμακας MFD, με βάση τη διάσταση Minkowski Bouligand [Mar94]. Μετρώντας λοιπόν το MFD, είμαστε σε θέση να εξετάσουμε την πολυπλοκότητα και τον κατακερματισμό των σημάτων σε πολλαπλές κλίμακες, δημιουργώντας έτσι ένα προφίλ τοπικών MFD σε κάθε χρονική τοποθεσία.

Ο αλγόριθμος που αναπτύχθηκε είναι γνωστός ως μέθοδος μορφολογικής κάλυψης (morphological covering method) και αποτελείται από τα πιο κάτω βήματα:

1. Δημιουργία του Minkowski cover χρησιμοποιώντας δισδιάστατες μορφολογικές διαστολές (dilations) του γραφήματος  $F$  του σήματος από εκδόσεις πολλαπλών κλιμάκων  $sB = \{sb : b \in B\}$  ενός κυρτού συμμετρικού επιπέδου συνόλου  $B$ , όπου το  $s \geq 0$  αντιπροσωπεύει την παράμετρο κλίμακας:

$$F \oplus sB = \{z + sb \in \mathbb{R}^2 : Z \in F, b \in B\} \quad (2.1.15)$$

2. Υπολογισμός του cover area  $A_B(s) = \text{area}(F \oplus sB)$ . Επομένως, το Fractal Dimension ορίζεται ως:

$$D = \lim_{s \rightarrow 0} \frac{\log[A_B(s)/s^2]}{\log[1/s]} \quad (2.1.16)$$

Ιδανικά το  $B$  είναι ένας δίσκος μονάδας. Ωστόσο, το  $D$  παραμένει αμετάβλητο όσο το  $B$  είναι συμπαγές, κυρτό και συμμετρικό. Στη διακριτή περίπτωση, επιλέγουμε ως  $B$  μια προσέγγιση στο δίσκο από ένα κυρτό συμμετρικό υποσύνολο ακτίνας μονάδας του  $\mathbb{Z}^2$ . Το παραπάνω όριο δεν θα αλλάξει αν προσεγγίσουμε το  $A_B(s)$  με το εμβαδόν της διαφοράς μεταξύ της μορφολογικής

διαστολής και της διάβρωσης (erosion) του διακριτού σήματος του  $N$  δείγματος  $F[n]$  με μια συνάρτηση  $G_s[n]$  που είναι ο επάνω φάκελος του διακριτού συνόλου  $sB$ :

$$A_B(s) = \sum_{n=0}^{N-1} ((F \oplus G_s) = (F \ominus G_s))[n] \quad (2.1.17)$$

για  $s = 1, \dots, s_{max} \leq \frac{N}{2}$ . Αυτό μειώνει σημαντικά την πολυπλοκότητα εισάγοντας μονοδιάστατες πράξεις σήματος, που είναι απλές μη γραμμικές συνελίξεις, αντί για πολύπλοκες πράξεις δισδιάστατου συνόλου. Περαιτέρω μείωση σε γραμμική πολυπλοκότητα επιτυγχάνεται με την εκτέλεση των παραπάνω πράξεων με scale-recursive.

Η Φράκταλ διάσταση (Fractal Dimension)  $D$  μπορεί να εκτιμηθεί με ελάχιστα τετράγωνα (least-squares) τα οποία προσαρμόζουν μια ευθεία γραμμή και μετρούν τη κλίση του plot  $\log[A_B(s)]$  έναντι του  $\log(s)$  διότι:

$$\log[A_B(s)] = (2 - D)\log(s) + constant \quad (2.1.18)$$

υποθέτοντας τον νόμο ισχύος  $A_B(s) \sim s^{2-D}$  ως  $s \rightarrow 0$ . Ωστόσο, τα σήματα μας, και τα περισσότερα σήματα του πραγματικού κόσμου, δεν έχουν την ίδια δομή σε κάθε κλίμακα, επομένως ο εκθέτης στην κυρίαρχη ισχύ  $s^{2-D}$  μπορεί να ποικίλλει. Επομένως, υπολογίζουμε την κλίση των δεδομένων σε ένα παράθυρο μικρής κλίμακας  $w$  που μετακινούνται κατά μήκος του άξονα κλίμακας  $s \{s, s + 1, \dots, s + w\}$ , δημιουργώντας ένα προφίλ τοπικών MFDs  $D(s, t)$  (ή φρακτόγραμμα) σε κάθε χρονική θέση  $t$ . Η τοπική κλίση είναι τώρα μια εκτίμηση  $2 - D$  και μας δίνει τη διάσταση φράκταλ. Το  $D$  κυμαίνεται μεταξύ 1 και 2 και όσο μεγαλύτερο γίνεται, τόσο μεγαλύτερος είναι ο γεωμετρικός κατακερματισμός του σήματος.

## Entropy

Η εντροπία είναι ένα μέτρο της τυχαιότητας ή της ‘αταξίας’ (disorder) ενός συστήματος. Οι προσεγγίσεις εντροπίας παρουσιάζουν μεγάλο ενδιαφέρον για την αξιολόγηση του βαθμού ανωμαλίας και πολυπλοκότητας των φυσικών, φυσιολογικών, κοινωνικών και οικονομικών συστημάτων.

Η εντροπία Shannon (ShEn), η conditional εντροπία (CE) και η information storage (IS) είναι οι πιο κοινές έννοιες που χρησιμοποιούνται στο πλαίσιο της ανάλυσης φυσιολογικών σημάτων. Το ShEn και το CE αντίστοιχα, δείχνουν την ποσότητα των πληροφοριών που μαθαίνονται και το ρυθμό παραγωγής πληροφοριών σε ένα σύστημα. Με βάση αυτές τις δύο προσεγγίσεις, έχουν προταθεί διάφορες τεχνικές που έχουν τις ρίζες τους στη θεωρία πληροφοριών, όπως η approximate εντροπία (ApEn), sample εντροπία (SampEn), corrected conditional εντροπία (CorCE), fuzzy εντροπία (FuzzyEn), permutation εντροπία (PermEn), distribution εντροπία (DistEn), dispersion διασποράς (DispEn), και fluctuation dispersion εντροπία (FDispEn).

## Sample Entropy

Η Εντροπία δειγματος (SampEn) χρησιμοποιείται για την αξιολόγηση της πολυπλοκότητας των σημάτων φυσιολογικής χρονοσειράς, και για τη διάγνωση νοσημάτων [RM00]. Επίσης, μπορεί να χρησιμοποιηθεί στη βιοϊατρική και εμβιομηχανική έρευνα, για παράδειγμα για την αξιολόγηση του ορθοστατικού ελέγχου [BKH21; Had+21]. Η εντροπία δείγματος είναι ένα μέτρο του ρυθμού παραγωγής πληροφοριών από το σύστημα.



Υποθέτοντας ότι έχουμε ένα σύνολο δεδομένων χρονοσειράς μήκους  $N = \{x_1, \dots, x_N\}$  με σταθερή διάρκεια  $\tau$ . Ορίζουμε ένα πρότυπο διάνυσμα μήκους  $m$ , όπως το  $X_m(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$  και την συνάρτηση απόστασης  $d[X_m(i), X_m(j)](ij)$ , η οποία σε αυτή τη περίπτωση είναι η απόσταση Chebyshev (αλλά θα μπορούσε να είναι οποιαδήποτε συνάρτηση απόστασης, συμπεριλαμβανομένης της Ευκλείδειας απόστασης). Επομένως, ορίζουμε την εντροπία του δείγματος ως εξής:

$$SampEn = -\log \frac{A}{B} \quad (2.1.19)$$

όπου,

$A$ : είναι ο αριθμός των ζευγών διανυσμάτων προτύπων που έχουν  $d[X_m + 1(i), X_m + 1(j)] < r$

$B$ : είναι ο αριθμός των ζευγών διανυσμάτων προτύπων που έχουν  $d[X_m(i), X_m(j)] < r$ .

Είναι σαφές από τον ορισμό ότι το  $A$  θα έχει πάντα μια τιμή μικρότερη ή ίση με το  $B$ . Επομένως, η τιμή  $SampEn(m, r, \tau)$  θα είναι πάντα είτε μηδέν είτε θετική. Μια μικρότερη τιμή του  $SampEn$  υποδηλώνει επίσης μεγαλύτερη αυτοομοιότητα (self-similarity) στο σύνολο δεδομένων ή λιγότερο θόρυβο.

### Poincare plot Analysis

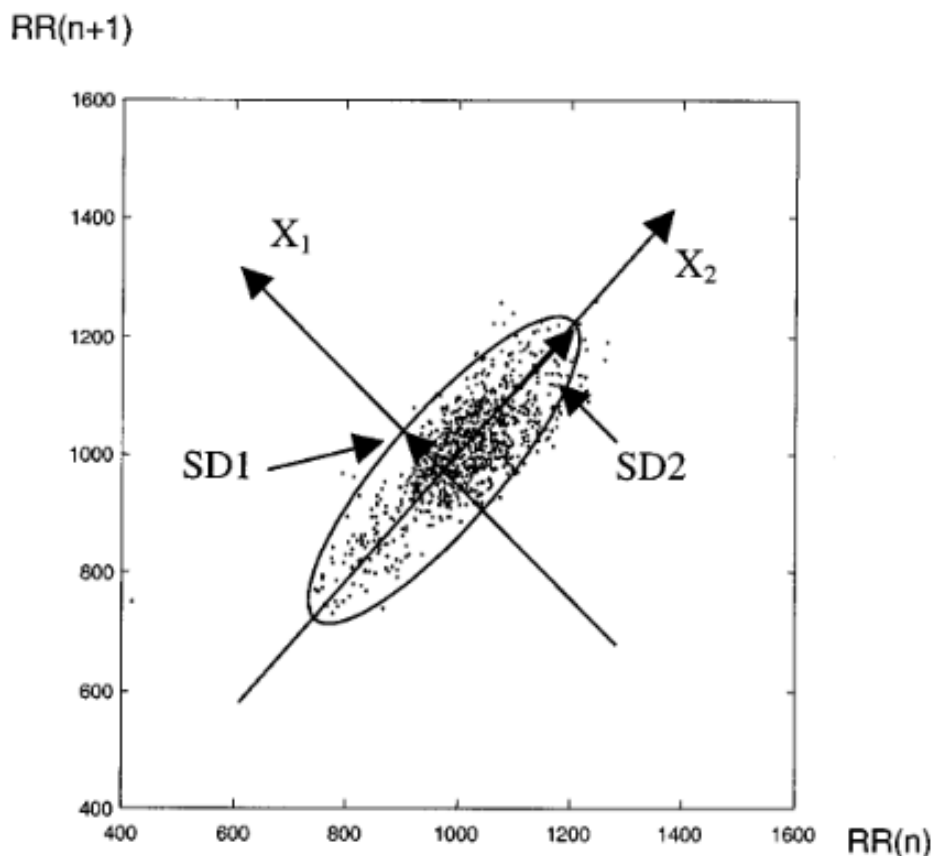
Το Poincare plot [BPK01] είναι ένα είδος γραφικής παράστασης επανάληψης όπου κάθε δείγμα μιας χρονοσειράς σχεδιάζεται σε σχέση με την προηγούμενη, και στη συνέχεια προσαρμόζεται μια έλλειψη σε αυτό το διάγραμμα διασποράς. Υπάρχουν τυπικοί περιγράφοι (descriptors) που χρησιμοποιούνται για τον ποσοτικό προσδιορισμό της γεωμετρίας των γραφικών παραστάσεων Poincare, δηλαδή τα  $SD1$  και  $SD2$  [BPK01; PG07]. Μπορούμε να τους αποκτήσουμε προσαρμόζοντας μια έλλειψη στο σχήμα. Οι descriptors  $SD1$  και  $SD2$  αντιπροσωπεύουν τον μικρό και τον κύριο ημιάξονα αυτής της προσαρμοσμένης έλλειψης. Οι Brennan et al. [BPK01] έδωσαν μια περιγραφή των  $SD1$  και  $SD2$  από την άποψη της γραμμικής στατιστικής. Το  $SD1$  είναι η τυπική απόκλιση των αποστάσεων των σημείων από τον άξονα 1 και καθορίζει το πλάτος της έλλειψης (βραχυπρόθεσμη μεταβλητότητα), το  $SD2$  ισούται με τις τυπικές αποκλίσεις από τον άξονα 2 και το μήκος της έλλειψης (μακροπρόθεσμη μεταβλητότητα) (Σχήμα 2.1.1).

Οι περιγραφείς (descriptors)  $SD1$  και  $SD2$  μπορούν να οριστούν ως:

$$SD1 = \frac{\sqrt{2}}{2} SD(x_n - x_{n+1}) \quad (2.1.20)$$

$$SD2 = \sqrt{2SD(x_n)^2} = \frac{1}{2} SD(x_n - x_{n+1})^2 \quad (2.1.21)$$

όπου  $SD$  είναι η τυπική απόκλιση (standard deviation) της χρονοσειράς.



Σχήμα 2.1.1: Παράδειγμα Poincaré plot που περιγράφει λεπτομερώς τη διαδικασία προσαρμογής της έλλειψης. Το σύστημα συντεταγμένων  $x_1$  και  $x_2$  είναι προσαρμοσμένος στις  $45^\circ$  ως προς τον κανονικό άξονα. Η τυπική απόκλιση της απόστασης των σημείων από κάθε άξονα καθορίζει το πλάτος ( $SD1$ ) και το μήκος ( $SD2$ ) της έλλειψης [BPK01].

Η μη γραμμική ανάλυση του HRV αναμένεται να παρέχει καλές πληροφορίες [TBC05]. Επομένως, χρησιμοποιείται το Poincaré plot με το πλάτος της έλλειψης ( $SD1$ ) να είναι ένα μέτρο του short-term HRV, ενώ το μήκος ( $SD2$ ) είναι ένα μέτρο του long-term HRV [Fil+20].

## 2.2 Μηχανική Μάθηση

Μηχανική Μάθηση ονομάζεται η ικανότητα ενός υπολογιστικού συστήματος να δημιουργεί μοντέλα ή πρότυπα από ένα σύνολο δεδομένων. Πολλοί επιστημονικοί κλάδοι επωφελούνται από τα επιτεύγματα στον τομέα της Μηχανικής Μάθησης όπως Εξόρυξη Δεδομένων, Πιθανότητες και Στατιστική, Θεωρία της Πληροφορίας, Αριθμητική Βελτιστοποίηση, Θεωρία της Πολυπλοκότητας, Θεωρία Ελέγχου (προσαρμοστική), Ψυχολογία (εξελικτική, γνωστική), Νευροβιολογία και Γλωσσολογία. Η μηχανική μάθηση, αποτελεί απαραίτητο εργαλείο στις μελέτες σχετικά με την ανίχνευση υποτροπών στο τομέα της ψυχικής υγείας. Ο τομέας της Μηχανικής Μάθησης χωρίζεται σε τρεις τρόπους μάθησης: επιβλεπόμενη (Supervised), μη επιβλεπόμενη (Unsupervised) και ενισχυτική (Reinforcement) μάθηση, που αναλύονται στις πιο κάτω παραγράφους.

**Επιβλεπόμενη Μάθηση:** Στην επιβλεπόμενη μάθηση (Supervised Learning), ανήκουν τα μοντέλα που εκπαιδεύονται χρησιμοποιώντας δεδομένα που έχουν επισημειωθεί με ετικέτες, δηλαδή χρησιμοποιούνται είσοδοι ταιριασμένες με τις αντίστοιχες εξόδους. Με αυτό τον τρόπο, επιχειρεί να προσδιορίσει μια συνάρτηση μεταξύ των εισόδων και εξόδων με στόχο τη γενίκευση της συνάρτησης αυτής για εισόδους με άγνωστη έξοδο.

**Μη επιβλεπόμενη Μάθηση:** Στην μη επιβλεπόμενη μάθηση (Unsupervised Learning), ανήκουν τα μοντέλα που εκπαιδεύονται χρησιμοποιώντας μόνο δεδομένα εισόδου, χωρίς γνώση των επιθυμητών εξόδων και ο αλγόριθμος πρέπει από μόνος του να ανακαλύψει συσχετίσεις, επαναλαμβανόμενα μοτίβα σε ένα σύνολο δεδομένων και να δημιουργήσει εκ νέου πρότυπα.

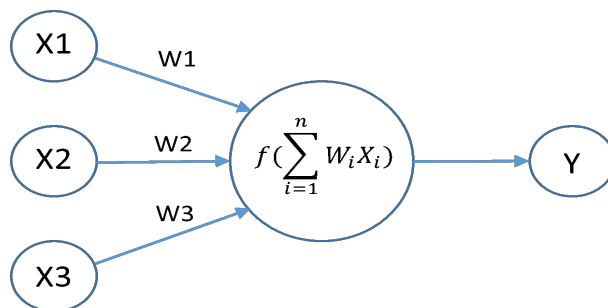
Στη δική μας μελέτη θα γίνει χρήση μη επιβλεπόμενης μάθησης εκπαιδεύοντας τα μοντέλα μας ώστε να μάθουν την ακολουθία εισόδου, να βρουν δηλαδή τη δομή των δεδομένων εισόδου και να τα δημιουργήσουν εκ νέου.

**Ενισχυτική Μάθηση:** Στην ενισχυτική μάθηση (Reinforcement Learning) ανήκουν τα μοντέλα τα οποία μαθαίνουν από την άμεση αλληλεπίδραση με το περιβάλλον. Το σύστημα έχει ως στόχο τη μεγιστοποίηση της ‘ανταμοιβής’ που λαμβάνει. Η διαφορά της ενισχυτικής μάθησης με τις υπόλοιπες μεθόδους της μηχανικής μάθησης είναι ότι αναπαριστά τον τρόπο με τον οποίο ο άνθρωπος μαθαίνει. Χρησιμοποιείται κυρίως σε προβλήματα Σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ και σε βελτιστοποίηση εργασιών σε εργοστασιακούς χώρους [SB18].

## 2.3 Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα είναι εμπνευσμένα από τη λειτουργία του ανθρώπινου εγκεφάλου με κύρια λειτουργία να μαθαίνουν να γενικεύουν. Επιχειρούν να χρησιμοποιήσουν κάποιες ‘οργανωτικές’ αρχές που πιστεύεται ότι χρησιμοποιούνται στον άνθρωπο [Bis94].

Ένα νευρωνικό δίκτυο αποτελείται από τους νευρώνες εισόδου, τους νευρώνες εξόδου και τους υπολογιστικούς νευρώνες. Οι νευρώνες εισόδου διοχετεύουν στους υπολογιστικούς νευρώνες την είσοδο του προβλήματος (π.χ., πρότυπα). Οι νευρώνες εξόδου, χρησιμοποιούνται για να παρουσιάσουν στο περιβάλλον το αποτέλεσμα του νευρωνικού δικτύου. Οι υπολογιστικοί νευρώνες ή κρυμμένοι νευρώνες πολλαπλασιάζουν κάθε είσοδο που δέχονται από τους νευρώνες εισόδου ή από άλλους νευρώνες επεξεργασίας με μια τιμή συσχετισμένη με αυτούς που ονομάζεται βάρος (weight). Το συνολικό αποτέλεσμα εισάγεται σε μια συνάρτηση που ονομάζουμε συνάρτηση ενεργοποίησης και είτε παρουσιάζεται στην έξοδο είτε περνάει σε κάποιον άλλο νευρώνα επεξεργασίας. Μαθηματικά μπορούμε να εκφράσουμε την έξοδο ενός νευρώνα:  $y = f[\sum_{i=1}^n x_i w_i]$ , όπου  $x_i$  είναι η είσοδος  $i$  από τις  $n$  εισόδους του προβλήματος,  $w_i$  είναι το βάρος διασύνδεσης με την  $i$  είσοδο και  $f(x)$  είναι η συνάρτηση ενεργοποίησης. Οι συχνότερες συναρτήσεις ενεργοποίησης σύμφωνα με τη βιβλιογραφία είναι η βηματική συνάρτηση (step function), γραμμική συνάρτηση (linear function), σιγμοειδής συνάρτηση (sigmoid function) και η υπερβολική εφαιπτομένη (hyperbolic functions) [Nwa+18].



Σχήμα 2.3.1: Απλό νευρωνικό δίκτυο [Ola15]

Τα πολυεπίπεδα νευρωνικά δίκτυα διαθέτουν τουλάχιστον ένα κρυφό στρώμα. Οι κόμβοι των διαφόρων επιπέδων μπορεί να είναι πλήρως συνδεδεμένοι (fully connected) ή μερικώς συνδεδεμένοι (partially connected). Επιπλέον, χαρακτηρίζονται με βάση τον τρόπο με τον οποίο είναι συνδεδεμένοι οι κόμβοι τους, δηλαδή αν είναι πρόσθιας τροφοδότησης (feed forward) ή οπίσθιας τροφοδότησης (feed backward ή recurrent). Στην πλειοψηφία των εφαρμογών χρησιμοποιούνται δίκτυα πρόσθιας τροφοδότησης [Bis94].

Η εκπαίδευση των νευρωνικών δικτύων είναι ο τρόπος με τον οποίο τροποποιούμε τις τιμές στα βάρη. Ανάλογα με την τεχνική εκπαίδευσης τα νευρωνικά δίκτυα διακρίνονται σε δίκτυα επιβλεπόμενης, μη επιβλεπόμενης ή ενισχυτικής μάθησης όπως αναφέραμε και στη Ενότητα 2.2. Η εκπαίδευση έχει ως βασικό στόχο να βρει ένα τρόπο αλλαγής των συνδετικών βαρών που θα έχει ως αποτέλεσμα την αλλαγή της γενικής συμπεριφοράς του δικτύου με την αύξηση της ικανότητας του δικτύου να παρέχει στο μέλλον μία επιθυμητή έξοδο μετά από μία δεδομένη είσοδο. Όταν η επιθυμητή έξοδος είναι εκ των προτέρων γνωστή λέμε ότι το δίκτυο μαθαίνει με επίβλεψη (supervised learning), αλλιώς μαθαίνει χωρίς επίβλεψη (unsupervised learning). Στα πλαίσια της παρούσας διπλωματικής εργασίας, χρησιμοποιήθηκε μη επιβλεπόμενη μάθηση για την εκπαίδευση των μοντέλων, αφού σκοπός μας είναι τα μοντέλα μας να μάθουν τη συμπεριφορά του κάθε ασθενή ανάλογα της περιόδου που διανύει.

Στη μάθηση χωρίς επίβλεψη (unsupervised learning), οι τιμές των βαρών και των πολώσεων (biases) τροποποιούνται σε σχέση με τις εισόδους. Αρχικά μοιάζει μη πρακτική μέθοδος, διότι τίθεται το ερώτημα πως μπορούμε να εκπαιδύσουμε ένα δίκτυο χωρίς να ξέρουμε τι πρόκειται να κάνει. Η απάντηση όμως είναι ότι οι περισσότεροι από αυτούς τους αλγόριθμους εκτελούν τη διαδικασία της ομαδοποίησης. Μαθαίνουν να κατηγοριοποιούν ένα πρότυπο εισόδου σε ένα πεπερασμένο αριθμό ομάδων [Bis94].

### Προβλήματα κατά την Εκπαίδευση ενός Δικτύου

Το βασικό πρόβλημα κατά την εκπαίδευση ενός δικτύου είναι η υπερπροσαρμογή. Υπερπροσαρμογή (Overfitting) [PS21] ονομάζουμε το φαινόμενο κατά το οποίο ένα νευρωνικό δίκτυο έχει ‘απομνημονεύσει’ τα παραδείγματα εκπαίδευσης, δηλαδή παρέχει ακριβείς τιμές εξόδου για τα παραδείγματα εκπαίδευσης, αλλά παρουσιάζει μειωμένη ικανότητα γενίκευσης. Όταν ένα μοντέλο δεν μπορεί να γενικευθεί καλά σε νέα δεδομένα, τότε αυτό σημαίνει ότι δεν θα μπορεί να εκτελέσει τη ταξινόμηση ή την πρόβλεψη στο πρόβλημα.

Για την αντιμετώπιση του φαινομένου της υπερπροσαρμογής συνήθως χρησιμοποιούνται τεχνικές όπως το πρόωρο σταμάτημα (early stopping) [Yin19] ή τη χρήση του Dropout Layer [Sri+14].

Πρόωρο σταμάτημα, ονομάζουμε τη μέθοδο όπου σταματάμε την εκπαίδευση του δικτύου σε ένα σημείο πριν από το τέλος όλων των εποχών, όπου σε αυτό το σημείο η περαιτέρω μείωση του σφάλματος εκπαίδευσης οδηγεί σε αύξηση του σφάλματος επικύρωσης. Η ιδέα της μεθόδου του dropout είναι η τυχαία απόρριψη ορισμένου αριθμού layers και σχετικών συνδέσεων των νευρωνικά δικτύων κατά τη διάρκεια της εκπαίδευσης. Αυτό έχει ως αποτέλεσμα να κάνει το layer να μοιάζει και να αντιμετωπίζεται σαν ένα layer με διαφορετικό αριθμό κόμβων και διαφορετική συνδεσιμότητα από το προηγούμενο layer.

## 2.4 Βαθιά Νευρωνικά Δίκτυα

### 2.4.1 Πλήρως Συνδεδεμένα Νευρωνικά Δίκτυα (FNN)

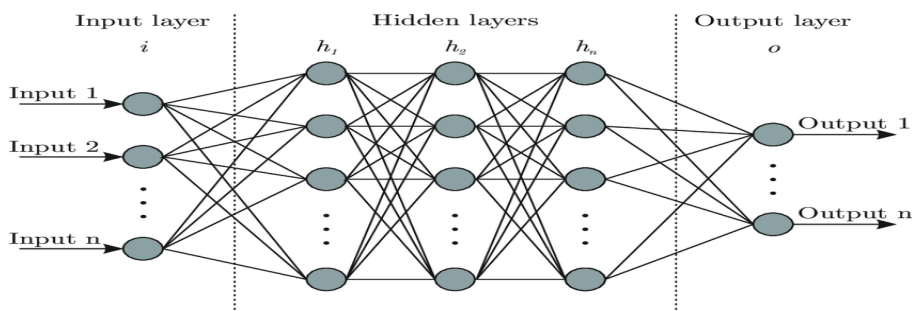
Τα πλήρως συνδεδεμένα (Fully Connected) νευρωνικά δίκτυα είναι απλά νευρωνικά δίκτυα πρόσθιας τροφοδοσίας (feed-forward). Κάθε νευρώνας των κρυφών στρωμάτων (hidden layers) συνδέεται με την έξοδο όλων των νευρώνων του προηγούμενου layer, με σκοπό να μάθει όλα τα χαρακτηριστικά της εισόδου.

Συγκεκριμένα, το δίκτυο όπως παρουσιάζεται στο Σχήμα 2.4.1 αποτελείται από τρία διαφορετικά στρώματα (layers), το στρώμα εισόδου (input layer), τα κρυφά στρώματα (hidden layers) και το στρώμα εξόδου (output layer), που αναλύονται παρακάτω.

**Input layer** Το στρώμα εισόδου είναι το πρώτο στρώμα το οποίο δέχεται την είσοδο.

**Hidden layers** Τα κρυφά στρώματα του δικτύου είναι τα στρώματα ανάμεσα στην είσοδο και έξοδο του δικτύου, στα οποία γίνεται η εξαγωγή όλων των απαραίτητων χαρακτηριστικών από το σήμα εισόδου προκειμένου να επιτευχθεί η κωδικοποίηση του. Ο αριθμός των κρυφών στρωμάτων ορίζεται ανάλογα με τις απαιτήσεις του προβλήματος. Συνήθως, όσο πιο απαιτητικό είναι το πρόβλημα για τον υπολογισμό των χαρακτηριστικών τόσο μεγαλύτερος θα είναι και ο αριθμός των hidden layers.

**Output layer** Το στρώμα εξόδου δίνει την τελική απόφαση του δικτύου για την λύση του προβλήματος με τα συγκεκριμένα δεδομένα εισόδου.



Σχήμα 2.4.1: Πλήρως συνδεδεμένο νευρωνικό δίκτυο (FNN) [Ola15]

### 2.4.2 Συνελικτικά Νευρωνικά Δίκτυα (CNNs)

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks - CNNs), περιγράφηκαν για πρώτη φορά από τον LeCun [LBH15]. Είναι μία υποκατηγορία νευρωνικών δικτύων τα οποία έχουν αποδειχθεί ότι είναι κατάλληλα για την επεξεργασία πληροφορίας που μπορεί να

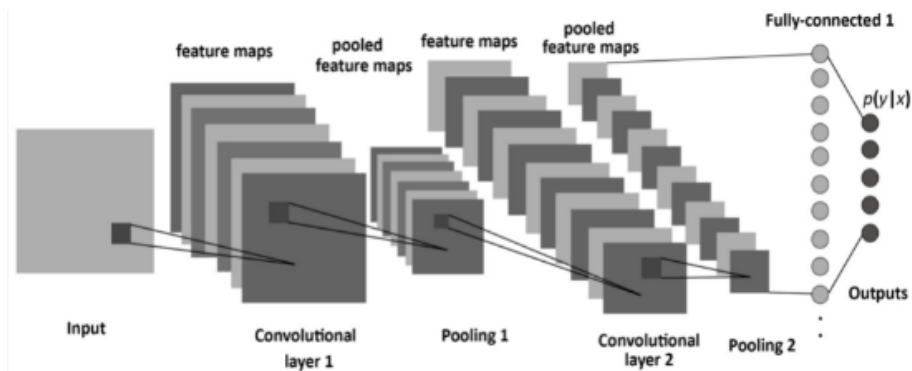
αναπαρασταθεί σε μορφή πλέγματος. Τα Συνελικτικά Νευρωνικά Δίκτυα έχουν εφαρμοστεί με επιτυχία σε πεδία όπως η αναγνώριση ομιλίας [AAM21], εικόνας [Gu+17] και βίντεο [Kar+14]. Τα δίκτυα αυτά αποτελούνται από νευρώνες οι οποίοι έχουν εκπαιδευσιμα weights και biases, παρόμοια με εκείνα των απλών Νευρωνικών Δικτύων που περιγράφηκαν αρχικά. Η διαφορά των συνελικτικών δικτύων σε σχέση με τα απλα δικτύα οφείλεται στην αρχιτεκτονική τους και στον τρόπο που προωθείται η πληροφορία εντός αυτών. Η αρχιτεκτονική τους είναι έτσι σχεδιασμένη ώστε πέραν από την αξιοποίηση μεμονωμένων τιμών εισόδου να αξιοποιείται και η χωρική τους εξάρτηση [LBH15].

Όπως δηλώνει και η ονομασία τους τα Συνελικτικά Νευρωνικά Δίκτυα, βασίζονται στη μαθηματική διαδικασία της συνέλιξης (convolution), για την επεξεργασία των δεδομένων και την εξαγωγή συμπερασμάτων. Η μαθηματική πράξη της δισδιάστατης συνέλιξης διακριτού χρόνου εισόδου  $x$  με ένα πυρήνα  $h$  στο σημείο  $(m, n)$  ορίζεται ως:

$$y[m, n] = h[m, n] * x[m, n] = \sum_j \sum_i h[i, j] * x[m - i, n - j] \quad (2.4.1)$$

όπου το εύρος των  $i, j$  εξαρτάται από το μέγεθος του κάθε πυρήνα  $h$ .

Η δομή ενός τέτοιου δικτύου απεικονίζεται στο Σχήμα 2.4.2. Όπως φαίνεται και στο σχήμα ένα CNN αποτελείται από έναν αριθμό από συνελικτικά στρώματα (convolutional layers) και συγκεντρωτικά στρώματα (pooling layers), τα οποία, προαιρετικώς, ακολουθούνται από πλήρως συνδεδεμένα στρώματα (fully connected layers). Η μεγαλύτερη διαφορά των Συνελικτικών Νευρωνικών δικτύων είναι στην είσοδο του δικτύου, αφού είναι πλέον δισδιάστατη ή και τρισδιάστατη, στη μορφή των παραμέτρων του δικτύου αλλά και στον τρόπο διασύνδεσης των νευρώνων διαφορετικών στρωμάτων. Το δίκτυο παίρνει σαν είσοδο ένα πολυδιάστατο πίνακα (δισδιάστατο ή τρισδιάστατο), στη συνέχεια υπολογίζει από αυτήν όλα τα χρήσιμα χαρακτηριστικά για την εξαγωγή ενός τελικού συμπεράσματος και δίνει την αντίστοιχη έξοδο. Για τον υπολογισμό των χρήσιμων χαρακτηριστικών η είσοδος περνάει ακολουθιακά από μια σειρά στρωμάτων όπως θα μπορούσε να είναι ένα συνελικτικό στρώμα, ένα συγκεντρωτικό στρώμα, ένα πλήρως συνδεδεμένο στρώμα ή ένα στρώμα απωλειών (loss layer). Τα στρώματα αυτά περιγράφονται λεπτομερώς στη συνέχεια.

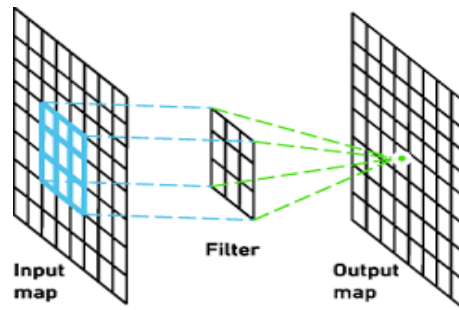


Σχήμα 2.4.2: Αρχιτεκτονική Convolutional Neural Network [AM17]



### Συνελικτικό στρώμα (Convolutional Layer)

Ένα συνελικτικό στρώμα, συνελίσει την είσοδο του με μία σειρά από φίλτρα ή πυρήνες (kernels), παράγοντας έτσι για κάθε φίλτρο έναν χάρτη ενεργοποίησης. Τα φίλτρα χρησιμοποιούνται για τον εντοπισμό συγκεκριμένων χαρακτηριστικών.



Σχήμα 2.4.3: Convolutional Layer [Yak+18]

Η έξοδος ενός συνελικτικού στρώματος δισδιάστατης συνέλιξης που δέχεται σαν είσοδο έναν τρισδιάστατο τανυστή διαστάσεων  $W_1 \times H_1 \times C_1$  (πλάτος, ύψος, κανάλια) και αποτελείται από φίλτρα διαστάσεων  $M \times N \times C_1$  (πλάτος, ύψος, κανάλια) το καθένα, έχει διαστάσεις  $W_2 \times H_2 \times C_2$  όπου:

$$W_2 = \frac{W_1 - M + 2Pad}{Stride} + 1, H_2 = \frac{H_1 - N_2Pad}{Stride}, C_2 = K \quad (2.4.2)$$

όπου Stride είναι το βήμα που χρησιμοποιείται κατά τον υπολογισμό της συνέλιξης, και Pad το εύρος του zero padding στα άκρα της εισόδου.

Με βάση τα παραπάνω και λαμβάνοντας υπόψη τη διάσταση του διανύσματος πόλωσης, ο αριθμός των υπό μάθηση παραμέτρων του συνελικτικού στρώματος είναι:

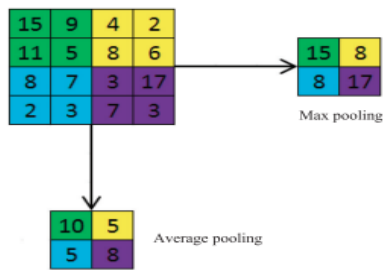
$$card(P) = K \cdot (M \cdot N \cdot C_1) + K \quad (2.4.3)$$

### Συγκεντρωτικό στρώμα (Pooling Layer)

Είναι τα στρώματα που χρησιμοποιούνται μεταξύ των Convolutional στρωμάτων προκειμένου να μειώσουν τις παραμέτρους του δικτύου. Οι βασικές πράξεις που εκτελούνται σε αυτό το στρώμα είναι:

- Max-Pooling - επιλογή μέγιστης τιμής ανά κελί.
- Average-Pooling - επιλογή της μέσης τιμής ανά κελί.

Το στρώμα υπερ-δειγματοληψίας (pooling layer) είναι ένα στρώμα δειγματοληψίας μεταξύ διαδοχικών επιπέδων συνέλιξης. Με το pooling layer επιτυγχάνουμε μείωση των διαστάσεων της εξόδου ενός συνελικτικού στρώματος και συνεπώς μειώνουμε τον αριθμό των παραμέτρων και των υπολογισμών στο δίκτυο. Αυτή η μείωση γίνεται με την χρήση ενός πυρήνα, όπου εφαρμόζεται μια συνάρτηση στα στοιχεία του πυρήνα.



Σχήμα 2.4.4: Pooling Layer [RW17]

Για ένα δίκτυο διδιάστατης συνέλιξης, αν η είσοδος είναι ένας ταυστής διαστάσεων  $W_1 \times H_1 \times C_1$  (πλάτος, ύψος, κανάλια), Stride το βήμα και  $M \times N$  η διάσταση των υποπεριοχών, τότε η έξοδος του συγκεντρωτικού στρώματος είναι ένας ταυστής με διαστάσεις  $W_2 \times H_2 \times C_2$  όπου:

$$W_2 = \frac{W_1 - M}{Stride} + 1, H_2 = \frac{H_1 - N}{Stride} + 1, C_2 = C_1. \quad (2.4.4)$$

### Fully Connected Layer

Τα fully connected layers είναι τα στρώματα που συνδέουν κάθε νευρώνα του τρέχοντος στρώματος με το στρώμα (Ενότητα 2.4.1). Ως είσοδος δίνεται ένας μονοδιάστατος πίνακας, ο οποίος αφού πλατυνθεί, στέλνεται σε ένα πλήρως συνδεδεμένο στρώμα για να κατηγοριοποιηθούν τα δεδομένα εισόδου.

Το βασικό πλεονέκτημα των συνελικτικών νευρωνικών δικτύων σε σχέση με τα πλήρως συνδεδεμένα επίπεδα είναι η μείωση των συνδέσεων από στρώμα σε στρώμα, γεγονός που προκαλεί μείωση των συνολικών βαρών του δικτύου. Αυτή η μείωση κάνει ταχύτερη την εκπαίδευση και μειώνει την πολυπλοκότητά του δικτύου.

### 2.4.3 Sequence Modeling: Αναδρομικά Νευρωνικά Δίκτυα (RNN)

Τα αναδρομικά νευρωνικά δίκτυα (RNN) είναι μια κατηγορία νευρωνικών δικτύων που χρησιμοποιεί διαδοχικά δεδομένα ή δεδομένα χρονοσειρών (time-series). Σε αντίθεση με τα παραδοσιακά βαθιά νευρωνικά δίκτυα, τα οποία υποθέτουν ότι οι εισοδοί και οι έξοδοι είναι ανεξάρτητες μεταξύ τους, τα RNNs θεωρούν ότι η έξοδος εξαρτάται από τα προηγούμενα στοιχεία της ακολουθίας. Διακρίνονται από τη 'μνήμη' τους καθώς λαμβάνουν πληροφορίες από προηγούμενες εισόδους για να επηρεάσουν την τρέχουσα είσοδο και έξοδο.

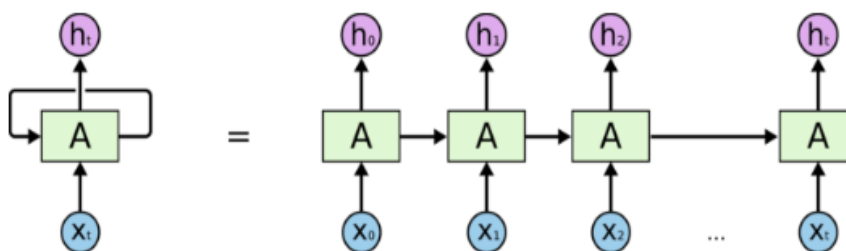
Όπως φαίνεται στο Σχήμα 2.4.5 στα RNNs οι συνδέσεις μεταξύ νευρώνων σχηματίζουν έναν κατευθυνόμενο κύκλο. Κάθε βήμα στο ξεδίπλωμα αναφέρεται ως χρονικό βήμα, με  $x_t$  την είσοδο στο χρονικό βήμα  $t$ . Το  $h_t$  είναι η κρυφή κατάσταση στο βήμα  $t$  και περιέχει πληροφορίες, που εξάγονται από όλα τα χρονικά βήματα έως το  $t$ . Η κρυφή κατάσταση  $h$  ενημερώνεται με πληροφορίες για τη νέα είσοδο μετά από κάθε βήμα όπως φαίνεται στην Εξίσωση 2.4.5:

$$h_t = f(Wx_t + Uh_{t-1}) \quad (2.4.5)$$



όπου τα  $U$  και  $W$  είναι τα διανύσματα βαρών πάνω από τις νέες εισόδους και την κρυφή κατάσταση, αντίστοιχα. Η συνάρτηση  $f$ , γνωστή ως συνάρτηση ενεργοποίησης, είναι συνήθως είτε η υπερβολική εφαπτομένη (Tanh) είτε η λογιστική συνάρτηση (Sigmoid) είτε η rectified γραμμική συνάρτηση (ReLU).

Στην απλούστερη μορφή ενός RNN, η κατάσταση κάθε νευρώνα σε ένα αναδρομικό στρώμα είναι απλώς η έξοδος του. Τα RNNs μπορούν να πάρουν μια αυθαίρετη ακολουθία μήκους ως είσοδο. Έτσι, για μια ακολουθία στοιχείων  $T$  διανυσμάτων  $m$ -διάστατων,  $x_1, x_2, \dots, x_t$ , η έξοδος ενός απλού RNN, που παραμετροποιείται από τους πίνακες  $W \in R_{m \times n}$  και  $U \in R_{n \times n}$  είναι μια ακολουθία  $n$ -διάστατων διανύσματα  $y_1, y_2, \dots, y_n$ , που δίνονται από την Εξίσωση 2.4.5, αν θεωρήσουμε ότι η κρυφή κατάσταση  $h$  είναι η έξοδος μας.

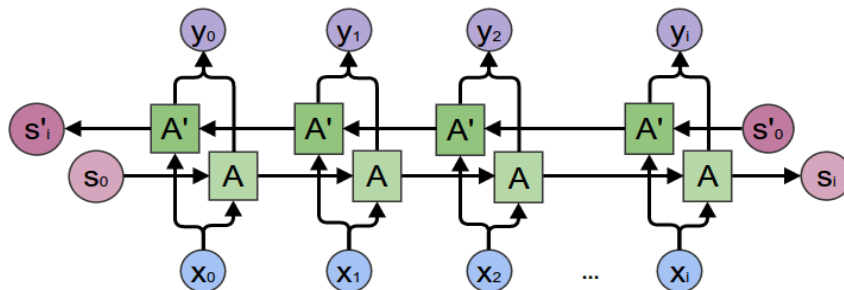


Σχήμα 2.4.5: Απλό RNN [Ola15]

Τα RNNs έχουν δύο σημαντικά προβλήματα, το πρόβλημα του vanishing gradient και του exploding gradient. Το exploding gradient είναι όταν ο αλγόριθμος, χωρίς ιδιαίτερο λόγο, αυξάνει εκθετικά τα gradients (τα οποία χρησιμοποιούνται για την ενημέρωση των βαρών) κατά την εκπαίδευση του νευρωνικού δικτύου. Το vanishing gradient είναι όταν ο αλγόριθμος μας δίνει τις gradients τιμές πολύ μικρές με αποτέλεσμα το μοντέλο να σταματά να μαθαίνει. Για την αντιμετώπιση των πιο πάνω αναπτυχθήκαν οι ιδέες των LSTM (Long short-term memory) και GRU (Gated Recurrent Unit) που αναλύονται στις πιο κάτω ενότητες.

### Αμφίδρομα αναδρομικά νευρωνικά δίκτυα (Bi-RNN)

Τα Αμφίδρομα αναδρομικά νευρωνικά δίκτυα, Σχήμα 2.4.6, είναι μια παραλλαγή της αρχιτεκτονικής των RNNs, αφού σε αντίθεση με τα μονοκατευθυντικά αναδρομικά νευρωνικά δίκτυα, που αντλούν δεδομένα μόνο από προηγούμενες εισόδους για να κάνουν προβλέψεις, τα Bi-RNN αντλούν μελλοντικά δεδομένα για να βελτιώσουν την ακρίβεια τους. Στην πραγματικότητα απλώς συνδυάζουν δύο ανεξάρτητες RNN, έτσι ώστε τα δίκτυα να έχουν τόσο προς τα πίσω όσο και προς τα εμπρός πληροφορίες για την ακολουθία σε κάθε χρονικό βήμα [SP97].

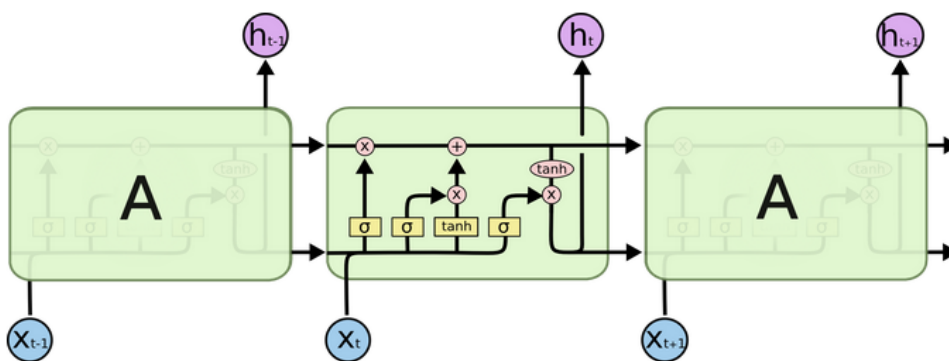


Σχήμα 2.4.6: Αμφίδρομο αναδρομικό νευρωνικό δίκτυο (Bi-RNN) [Ola15]

### Νευρωνικό Δίκτυο Μακράς-Βραχείας Μνήμης (LSTM)

Το νευρωνικό δίκτυο Μακράς-Βραχείας Μνήμης (LSTM) αποτελεί μία υποκατηγορία των RNNs, που προτάθηκε για πρώτη φορά από τους Sepp Hochreiter και Juergen Schmidhuber το 1997 ως λύση στο πρόβλημα του vanishing gradient [HS97]. Δηλαδή, εάν η προηγούμενη κατάσταση που επηρεάζει την τρέχουσα πρόβλεψη δεν βρίσκεται στο πρόσφατο παρελθόν, το μοντέλο RNN μπορεί να μην είναι σε θέση να προβλέψει με ακρίβεια την τρέχουσα κατάσταση. Για να αντιμετωπιστεί αυτό, τα LSTMs έχουν ‘κελιά’ στα κρυφά στρώματα του νευρωνικού δικτύου, τα οποία έχουν τρεις πύλες – μια πύλη εισόδου (input gate), μια πύλη εξόδου (output gate) και μια πύλη επιλεκτικής συγκράτησης (forget gate). Αυτές οι πύλες ελέγχουν τη ροή των πληροφοριών που απαιτούνται για την πρόβλεψη της εξόδου στο δίκτυο.

Όπως όλα τα RNNs, έτσι και τα LSTMs έχουν τη μορφή μιας αλυσίδας επαναλαμβανόμενων μονάδων από νευρωνικά δίκτυα, με τη διαφορά όμως ότι η δομή αυτών των μονάδων τους είναι αρκετά πιο περίπλοκη από αυτή των απλών RNNs, η οποία αποτελείται από ένα στρώμα με μία tanh συνάρτηση ενεργοποίησης. Το LSTM περιλαμβάνει αντί για ένα στρώμα νευρωνικού δικτύου, τέσσερα που αλληλεπιδρούν μεταξύ τους όπως φαίνεται στο Σχήμα 2.4.7.



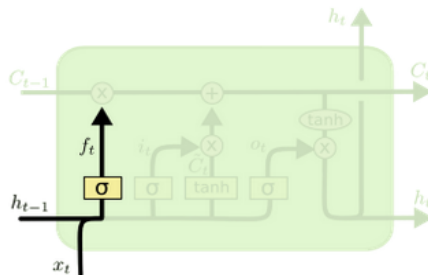
Σχήμα 2.4.7: Εσωτερική Δομή ενός LSTM layer. [Ola15]

Το κλειδί για τη λειτουργία του LSTM είναι η κατάσταση του κελιού μνήμης (memory cell). Η κατάσταση του κελιού  $C_t$  είναι σαν μια ζώνη που διατρέχει την αλυσίδα. Η γραμμή αυτή επεκτείνεται στην αλυσίδα με μικρές γραμμικές αλλαγές στο περιεχόμενό της. Το LSTM έχει την δυνατότητα να αφαιρεί ή να προσθέτει πληροφορία στο κελί κατάστασης (cell state), με

την χρήση δομών που λέγονται πύλες (gates) οι οποίες χρησιμοποιούν σιγμοειδείς συναρτήσεις ενεργοποίησης με τιμές στο εύρος  $[0, 1]$ .

**Forget gate** Η πύλη επιλεκτικής συγκράτησης (forget gate) όπου παρουσιάζεται στο Σχήμα 2.4.8, είναι υπεύθυνη για την αφαίρεση πληροφοριών από την κατάσταση κελιού. Οι πληροφορίες που δεν απαιτούνται πλέον για την κατανόηση του προβλήματος από το LSTM ή οι πληροφορίες που είναι μικρότερης σημασίας αφαιρούνται μέσω πολλαπλασιασμού με ένα φίλτρο. Αυτό απαιτείται για τη βελτιστοποίηση της απόδοσης του δικτύου LSTM. Αυτή η πύλη λαμβάνει δύο εισόδους,  $h_{t-1}$  και  $x_t$ . Η  $h_{t-1}$  είναι η κρυφή κατάσταση από το προηγούμενο κελί ή η έξοδος του προηγούμενου κελιού και η  $x_t$  είναι η είσοδος στο συγκεκριμένο χρονικό βήμα. Οι δεδομένες εισοδοί πολλαπλασιάζονται με τους πίνακες βάρους και προστίθεται ένα bias. Μετά, η σιγμοειδής συνάρτηση εφαρμόζεται σε αυτήν την τιμή και αποφασίζει εάν η τιμή εξόδου της προηγούμενης κατάστασης θα κρατηθεί ή όχι, όπως φαίνεται παρακάτω

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (2.4.6)$$



Σχήμα 2.4.8: Forget Gate [Ola15]

**Input gate** Η πύλη εισόδου (Input Gate) είναι υπεύθυνη για την προσθήκη πληροφοριών στο κελί μνήμης. Αυτή η προσθήκη πληροφοριών είναι βασικά μια διαδικασία τριών βημάτων όπως φαίνεται στο Σχήμα 2.4.9.

1. Ρύθμιση των τιμών που πρέπει να προστεθούν στο cell state με τη συμμετοχή μιας σιγμοειδούς συνάρτησης. Αυτό είναι βασικά πολύ παρόμοιο με την πύλη επιλεκτικής συγκράτησης και λειτουργεί ως φίλτρο για όλες τις πληροφορίες από τα  $h_{t-1}$  και  $x_t$ :

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.4.7)$$

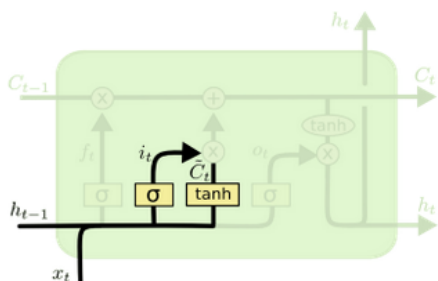
2. Δημιουργία διανύσματος που περιέχει όλες τις πιθανές τιμές που μπορούν να προστεθούν στο cell state. Αυτό γίνεται χρησιμοποιώντας τη συνάρτηση  $\tanh$ , η οποία εξάγει τιμές από  $-1$  έως  $+1$ :

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.4.8)$$

3. Πολλαπλασιασμός της τιμής του ρυθμιστικού φίλτρου (της σιγμοειδούς πύλης) με το δημιουργημένο διάνυσμα (τη συνάρτηση  $\tanh$ ) και στη συνέχεια προσθέτονται αυτές οι χρήσιμες πληροφορίες στο cell state μέσω της λειτουργίας προσθήκης:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (2.4.9)$$

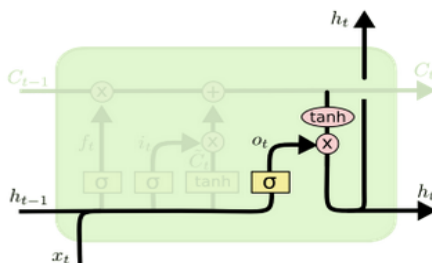
Μόλις ολοκληρωθεί αυτή η διαδικασία τριών βημάτων, διασφαλίζουμε ότι προστίθενται μόνο οι σημαντικές πληροφορίες στο cell state.



Σχήμα 2.4.9: Input Gate [Ola15]

**Output gate** Η λειτουργία μιας πύλης εξόδου (Output Gate) όπως φαίνεται στο Σχήμα 2.4.10, μπορεί και πάλι να αναλυθεί σε τρία βήματα:

1. Δημιουργία διανύσματος μετά την εφαρμογή της συνάρτησης  $\tanh$  στο cell state, κλιμακώνοντας έτσι τις τιμές στην περιοχή  $-1$  έως  $+1$ .
2. Δημιουργία φίλτρου χρησιμοποιώντας τις τιμές  $h_{t-1}$  και  $x_t$ , έτσι ώστε να μπορεί να ρυθμίσει τις τιμές που πρέπει να εξαχθούν από το διάνυσμα που δημιουργήθηκε παραπάνω. Αυτό το φίλτρο χρησιμοποιεί και πάλι μια σιγμοειδή συνάρτηση.
3. Πολλαπλασιασμός της τιμής αυτού του ρυθμιστικού φίλτρου με το διάνυσμα που δημιουργήθηκε στο βήμα 1 και στέλνοντάς το ως έξοδο και στην κρυφή κατάσταση του επόμενου κελιού.



Σχήμα 2.4.10: Output Gate [Ola15]

### Gated Recurrent Unit (GRU)

Η GRU είναι και αυτή μια παραλλαγή του RNN, είναι παρόμοια με τα LSTMs καθώς λειτουργεί επίσης για να αντιμετωπίσει το πρόβλημα βραχυπρόθεσμης μνήμης των μοντέλων RNNs αλλά είναι ελαφρώς διαφορετική αρχιτεκτονική. Συνδυάζει την forget και input gate σε μια ενιαία πύλη που ονομάζεται πύλη ενημέρωσης (update gate). Συγχωνεύει το cell state και την κρυφή κατάσταση μαζί και με λίγες αλλαγές καταλήγουν σε μια αρχιτεκτονική όπως στο Σχήμα 2.4.11. Δηλαδή, σε αντίθεση με το LSTM αντί για τρεις πύλες, έχει δύο - μια πύλη επαναφοράς και

μια πύλη ενημέρωσης. Παρόμοια με τις πύλες στα LSTM, οι πύλες επαναφοράς και ενημέρωσης ελέγχουν πόσες και ποιες πληροφορίες θα διατηρηθούν.

**Update gate** Η πύλη ενημέρωσης (update gate) υπολογίζει το  $z_t$  στο χρονικό βήμα  $t$  ακολουθώντας τα τρία βήματα που παρουσιάζονται παρακάτω:

1. Η είσοδος  $x_t$  πολλαπλασιάζεται με ένα βάρος  $W_{zx}$ .
2. Η προηγούμενη έξοδος  $h_{t-1}$  που περιέχει πληροφορίες από προηγούμενες μονάδες πολλαπλασιάζεται με ένα βάρος  $W_{zh}$ .
3. Και τα δύο προστίθενται μαζί και εφαρμόζεται μια σιγμοειδής συνάρτηση για συμπίεση της εξόδου μεταξύ του 0 και του 1.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (2.4.10)$$

Δεδομένης αυτής της πύλης, το ζήτημα του vanishing gradient εξαλείφεται, αφού το μοντέλο από μόνο του μαθαίνει πόσες από τις πληροφορίες του παρελθόντος θα περάσουν στο μέλλον.

**Reset gate** Η πύλη επαναφοράς (reset gate) υπολογίζει το  $r_t$  στο χρονικό βήμα  $t$  ακολουθώντας τα τρία ακόλουθα βήματα:

1. Η είσοδος  $x_t$  πολλαπλασιάζεται με ένα βάρος  $W_{rx}$ .
2. Η προηγούμενη έξοδος  $h_{t-1}$  που περιέχει πληροφορίες από προηγούμενες μονάδες πολλαπλασιάζεται με ένα βάρος  $W_{rh}$ .
3. Και τα δύο προστίθενται μαζί και εφαρμόζεται μια σιγμοειδής συνάρτηση για συμπίεση της εξόδου μεταξύ του 0 και του 1.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (2.4.11)$$

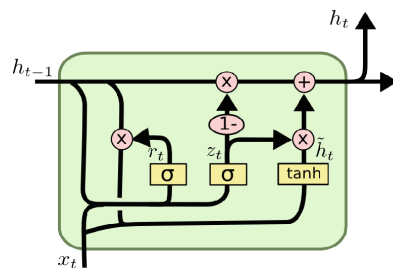
Η μαθηματική έκφραση 2.4.11 είναι πολύ παρόμοια με την παραπάνω πύλη και διαφέρει μόνο ως προς τα βάρη και τη χρήση της πύλης. Αυτή η πύλη έχει την αντίθετη λειτουργία συγκριτικά με την πύλη ενημέρωσης, καθώς χρησιμοποιείται από το μοντέλο για να αποφασίσει πόσες από τις προηγούμενες πληροφορίες θα ‘ξεχάσει’.

Στη συνέχεια υπολογίζεται το τρέχον περιεχόμενο μνήμης (current memory content) ως:

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (2.4.12)$$

Τέλος, πρέπει να υπολογιστεί το διάνυσμα  $h_t$ , Εξίσωση 2.4.13, που περιέχει πληροφορίες για τη τρέχουσα μονάδα και να τις μεταβιβάσει στο υπόλοιπο δίκτυο. Βασικό ρόλο σε αυτή τη διαδικασία έχει η πύλη ενημέρωσης, αφού αν το διάνυσμα  $z_t$  είναι κοντά στο 0, τότε ένα μεγάλο μέρος του τρέχοντος περιεχομένου θα αγνοηθεί καθώς είναι άσχετο με την πρόβλεψή μας. Ταυτόχρονα, δεδομένου ότι το  $z_t$  θα είναι κοντά στο 0 σε αυτό το χρονικό βήμα, το  $(1 - z_t)$  θα είναι κοντά στο 1 επιτρέποντας τη διατήρηση της πλειοψηφίας των πληροφοριών του παρελθόντος.

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (2.4.13)$$



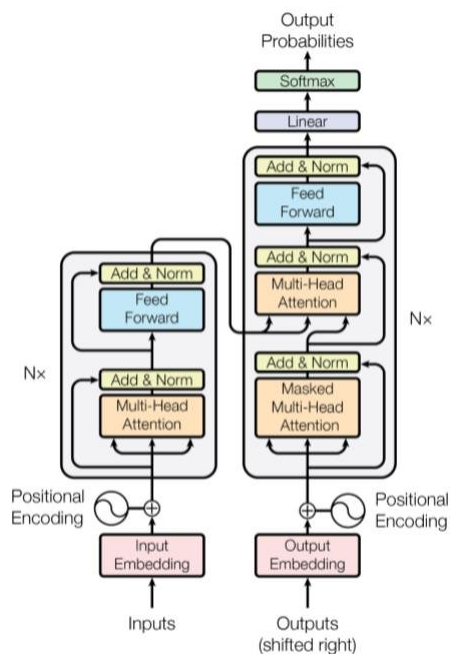
Σχήμα 2.4.11: Αρχιτεκτονική GRU [Ola15]

Το κύριο πλεονέκτημα της GRU είναι ο μειωμένος αριθμός παραμέτρων σε σύγκριση με το LSTM, γεγονός που έχει ως αποτέλεσμα ταχύτερη σύγκλιση και ένα πιο γενικευμένο μοντέλο.

#### 2.4.4 Transformer Network

Ο Transformer [Vas+17] είναι μια αρχιτεκτονική για τη μετατροπή μιας ακολουθίας σε μια άλλη με τη βοήθεια δύο μερών, ενός κωδικοποιητή (encoder) και ενός αποκωδικοποιητή (decoder). Διαφέρει από τα προηγούμενα αναδρομικά μοντέλα που περιγράφηκαν, αφού δεν συμπεριλαμβάνει επαναλαμβανόμενα δίκτυα όπως τις GRUs, τα LSTMs κ.λ.π. Ο Transformer χρησιμοποιεί μηχανισμούς προσοχής (attention-mechanism) για να αντλήσουν σφαιρικές εξαρτήσεις μεταξύ των εισόδων και των εξόδων.

Όπως φαίνεται στο Σχήμα 2.4.12 ο encoder βρίσκεται στα αριστερά και ο decoder βρίσκεται στα δεξιά. Ο encoder αντιστοιχίζει μια ακολουθία εισόδου σε μια ακολουθία αναπαραστάσεων σε έναν λανθάνοντα χώρο. Κατά συνέπεια, ο decoder δημιουργεί μια ακολουθία εξόδου ενός στοιχείου κάθε φορά συνδέοντας την κωδικοποιημένη αναπαράσταση με την έξοδο σε κάθε χρονικό βήμα. Τόσο ο encoder όσο και ο decoder αποτελούνται από modules που μπορούν να στοιβάζονται το ένα πάνω στο άλλο πολλές φορές, κάτι που περιγράφεται από το  $N_x$  στο σχήμα. Βλέπουμε ότι τα modules αποτελούνται κυρίως από επίπεδα Multi-Head Attention και Feed Forward.



Σχήμα 2.4.12: Αρχιτεκτονική Transformer Encoder-Decoder [Vas+17]

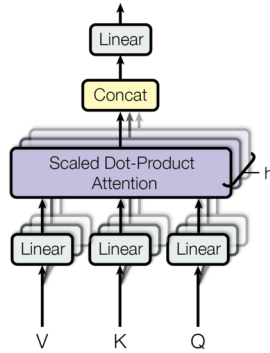
Η κύρια προϋπόθεση πίσω από την επιτυχία των μετασχηματιστών είναι η αυτοπροσοχή (self-attention). Το Self-attention είναι ένας μηχανισμός προσοχής που συσχετίζει διαφορετικές θέσεις μιας μεμονωμένης ακολουθίας προκειμένου να υπολογίσει μια αναπαράσταση της ακολουθίας, και περιγράφεται από την παρακάτω εξίσωση:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.4.14)$$

όπου το  $Q$  είναι ένας πίνακας που περιέχει το ερώτημα (διανυσματική αναπαράσταση μιας λέξης στην ακολουθία),  $K$  είναι όλα τα κλειδιά (διανυσματικές αναπαραστάσεις όλων των λέξεων της ακολουθίας),  $V$  είναι οι τιμές, οι οποίες είναι και πάλι οι διανυσματικές αναπαραστάσεις όλων των λέξεων στη σειρά και  $d_k$  είναι η διάσταση των ερωτημάτων και των κλειδιών.

Ο Transformer βελτιώνει το self-attention εισάγοντας την αυτοπροσοχή πολλών μηχανισμών (multi-head self attention), που φαίνεται στο Σχήμα 2.4.13. Ο μηχανισμός προσοχής επαναλαμβάνεται πολλές φορές με γραμμικές προβολές των  $Q$ ,  $K$  και  $V$ . Αυτό επιτρέπει στο σύστημα να μάθει από διαφορετικές αναπαραστάσεις των  $Q$ ,  $K$  και  $V$ , δηλαδή αν παρακολουθεί διαφορετικούς υπο-χώρους αναπαράστασης σε διαφορετικές θέσης. Αυτές οι γραμμικές αναπαραστάσεις γίνονται πολλαπλασιάζοντας τα  $Q$ ,  $K$  και  $V$  με τους πίνακες βάρους  $W$  που μαθαίνονται κατά τη διάρκεια της εκπαίδευσης. Μετά τα multi-head self-attention τόσο στον encoder όσο και στον decoder, ακολουθείται ένα στρώμα τροφοδοσίας (feed-forward), το οποίο εφαρμόζεται σε κάθε θέση ξεχωριστά και πανομοιότυπα.



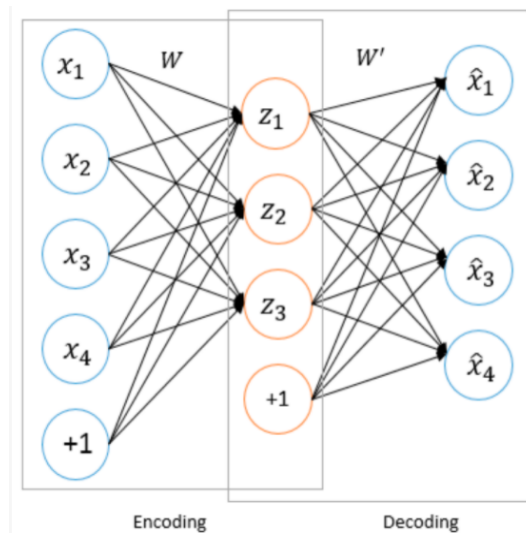


Σχήμα 2.4.13: Multi-head Self Attention [Vas+17]

### 2.4.5 Autoencoders

Η ιδέα πίσω από τους Autoencoders (AE) είναι η ανακατασκευή των δεδομένων εισόδου στην έξοδο με τη μικρότερη δυνατή παραμόρφωση [Bal12]. Διατυπωμένο με αυτόν τον τρόπο, μπορεί να μη φαίνεται χρήσιμο, αλλά η προσδοκία είναι ότι οι χρήσιμες ιδιότητες των εισόδων θα εξαχθούν κατά τη διάρκεια της εκπαίδευσης.

Ένας autoencoder αποτελείται αντίστοιχα από τρία στοιχεία: έναν κωδικοποιητή (encoder), έναν κώδικα (code) και τον αποκωδικοποιητή (decoder). Ο encoder συμπιέζει την είσοδο και δημιουργεί τον κώδικα, ο decoder αναδομεί την είσοδο με βάση τον κώδικα (Σχήμα 2.4.14). Η διαδικασία εκπαίδευσης είναι παρόμοια με την εκπαίδευση feed-forward μέσω backpropagation [GBC16]. Οι autoencoders από τη φύση τους έχουν απώλειες, πράγμα που σημαίνει ότι η έξοδος θα υποβαθμιστεί σε σύγκριση με την αρχική είσοδο. Ο στόχος της εκπαίδευσης είναι να ελαχιστοποιηθεί το σφάλμα ανακατασκευής.



Σχήμα 2.4.14: Αρχιτεκτονική Autoencoder [Alo+19]

Κάθε είσοδος του autoencoder αντιπροσωπεύεται από ένα διάνυσμα  $x \in R^n$  διάστασης  $n$  ίσης



με το μέγεθος εισόδου. Ο autoencoder παίρνει την είσοδο και την αντιστοιχίζει σε μια κρυφή αναπαράσταση μέσω μιας ντετερμινιστικής αντιστοίχισης, που ονομάζεται encoder και μπορεί να αναπαρασταθεί με την πιο κάτω συνάρτηση:

$$z = f_{\theta}(x) = \sigma(xW + b) \quad (2.4.15)$$

όπου  $\theta = W, b$ ,  $W \in R^{m \times n}$  είναι ο πίνακας βαρών,  $b \in R^m$  είναι το διάνυσμα bias,  $f_{\theta}(x)$  ο encoder και  $\sigma$  η συνάρτηση ενεργοποίησης.

Η κρυφή αναπαράσταση αντιστοιχίζεται με τον decoder έτσι ώστε να αναπαράγει το στρώμα εξόδου  $x' \in R^n$  ίδιας διάστασης  $n$  με την είσοδο. Αυτή η διαδικασία μπορεί να γραφτεί ως συνάρτηση:

$$x' = g_{\theta'}(h) = \sigma(W'z + b') \quad (2.4.16)$$

όπου  $\theta' = W', b'$ , και  $g_{\theta'}(h)$  ο decoder. Η συνάρτηση  $\sigma_f$  και  $\sigma_g$  είναι συνήθως μια μη γραμμική συνάρτηση ενεργοποίησης, π.χ., η υπερβολική συνάρτηση εφαπτομένης και η σιγμοειδής συνάρτηση [ZYX17].

### Convolutional Autoencoder

Ο convolutional autoencoder είναι μια ειδική περίπτωση ενός μοντέλου μάθησης χωρίς επίβλεψη, που έχει εκπαιδευτεί να αναπαράγει την εισόδο του στο στρώμα εξόδου. Η είσοδος περνά μέσα από έναν κωδικοποιητή (encoder), ο οποίος είναι ένα συνελικτικό δίκτυο που παράγει μια αναπαράσταση της εισόδου σε χαμηλές διαστάσεις. Ο αποκωδικοποιητής (decoder), που είναι ένα άλλο δείγμα συνελικτικού δικτύου, παίρνει αυτήν τη συμπιεσμένη είσοδο και αναδομεί την αρχική είσοδο. Τα συνελικτικά δίκτυα έχουν αναλυθεί στη Ενότητα 2.4.2.

Ο encoder χρησιμοποιείται για τη συμπίεση των δεδομένων και ο decoder χρησιμοποιείται για την αναπαραγωγή της αρχικής εισόδου. Επομένως, μπορούν να χρησιμοποιηθούν autoencoder για τη συμπίεση δεδομένων. Η λογική συμπίεσης είναι συγκεκριμένη, που σημαίνει ότι μαθαίνονται από τα δεδομένα και όχι από προκαθορισμένους αλγόριθμους συμπίεσης.

### Sequence-to-sequence Autoencoder

Ένας sequence-to-sequence Autoencoder, αποτελείται από 3 μέρη (Σχήμα 2.4.15): τον κωδικοποιητή (encoder), το διάνυσμα κωδικοποιητή (encoder vector) και τον αποκωδικοποιητή (decoder).

**Encoder** Αποτελείται από πολλές διαδοχικά επαναλαμβανόμενες μονάδες (π.χ., LSTM, GRU), όπου η καθεμία δέχεται ένα μεμονωμένο στοιχείο της ακολουθίας εισόδου, συλλέγει πληροφορίες για αυτό το στοιχείο και τις προωθεί προς τα εμπρός. Η είσοδος είναι ένα διάνυσμα  $x_t$  όπου  $t$  είναι η χρονική σειρά των στοιχείων που έχει το διάνυσμα μας. Οι κρυφές καταστάσεις  $h_t$  υπολογίζονται χρησιμοποιώντας τον τύπο:

$$h_t = f(W_{hh}h_{t-1} + W_{hx}x_t) \quad (2.4.17)$$

Η μαθηματική έκφραση 2.4.17 όπως παρατηρούμε αντιπροσωπεύει το αποτέλεσμα ενός συνθησιμένου επαναλαμβανόμενου νευρωνικού δικτύου. Όπως φαίνεται, απλώς εφαρμόζουμε τα κατάλληλα βάρη στην προηγούμενη κρυφή κατάσταση  $h_{(t-1)}$  και στο διάνυσμα εισόδου  $x_t$ .

**Encoder Vector** Αυτή είναι η τελική κρυφή κατάσταση που παράγεται από το τμήμα encoder του μοντέλου. Αυτό το διάνυσμα στοχεύει να αποθηκεύσει τις πληροφορίες για όλα τα στοιχεία εισόδου, προκειμένου να βοηθήσει τον decoder να κάνει ακριβείς προβλέψεις. Λειτουργεί ως η αρχική κρυφή κατάσταση του τμήματος decoder του μοντέλου.

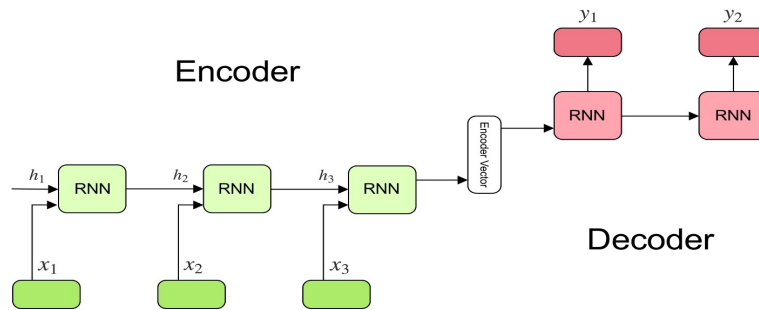
**Decoder** Αποτελείται από μια στοίβα πολλών επαναλαμβανόμενων μονάδων όπου η καθεμία προβλέπει μια έξοδο  $y_t$  σε ένα χρονικό βήμα  $t$ . Κάθε επαναλαμβανόμενη μονάδα δέχεται μια κρυφή κατάσταση από την προηγούμενη μονάδα και εξάγει τη δική της κρυφή κατάσταση. Η ακολουθία εξόδου είναι μια συλλογή όλων των στοιχείων της εισόδου. Κάθε στοιχείο αντιπροσωπεύεται ως  $y_t$  όπου  $t$  είναι η χρονική σειρά των στοιχείων που έχει το διάνυσμα μας. Οποιαδήποτε κρυφή κατάσταση  $h_t$  υπολογίζεται χρησιμοποιώντας τον τύπο:

$$h_t = f(W_{hh}h_{t-1}) \quad (2.4.18)$$

Η έξοδος  $y_t$  στο χρονικό βήμα  $t$  υπολογίζεται με τον τύπο:

$$y_t = \text{softmax}(W_S h_t) \quad (2.4.19)$$

όπου  $W_S$  το βάρος για τη δεδομένη χρονική κατάσταση.



Σχήμα 2.4.15: Αρχιτεκτονική Sequence-to-sequence Autoencoder [Kos19]

# Κεφάλαιο 3

## Δεδομένα και Προεπεξεργασία

### 3.1 Συλλογή Δεδομένων και Πειραματικό Πρωτόκολλο

Τα δεδομένα αποτελούνται από 24 ασθενείς που πάσχουν από ψυχωτικές διαταραχές. Από τους 24 ασθενείς, οι 12 πάσχουν από σχιζοφρένεια, 8 από διπολική διαταραχή I, 2 από σχιζοσυναισθηματική διαταραχή (Schizoaffective disorder), 1 με σύντομο ψυχωτικό επεισόδιο (Brief Psychotic Episode) και 1 με σχιζοφρενική διαταραχή (Schizophreniform disorder). Από τους 24 ασθενείς, μόνο οι 20 είχαν υποστεί κάποια ψυχωτική υποτροπή, επομένως σε αυτή τη διπλωματική εργασία έχουμε επεξεργαστεί τα δεδομένα μόνο των 20 ασθενών. Τα δεδομένα συλλέχθηκαν στο Πανεπιστήμιο Ψυχικής Υγείας, Ινστιτούτο Ερευνών Νευροεπιστημών και Ιατρικής Ακρίβειας ‘Κώστας Στεφανής’ (UMHRI) στην Αθήνα, Ελλάδα. Αφού ενημερώθηκαν οι συμμετέχοντες για τους στόχους του έργου, υπέγραψαν γραπτή συγκατάθεση και άδεια για τη συμμετοχή τους και τη χρήση των προσωπικών τους (ανώνυμων) δεδομένων, σύμφωνα με τις διατάξεις του Γενικού Κανονισμού (EU) 2016/679. Επιπρόσθετα, όλα τα πρωτόκολλα του έργου e-Prevention έχουν εγκριθεί από την Επιτροπή Δεοντολογίας του Ιδρύματος.

Ο Πίνακας 3.1 περιέχει πληροφορίες σχετικά με τα δημογραφικά στοιχεία ανά ψυχωτική διαταραχή, των 10 ασθενών, που μετά από την επεξεργασία δεδομένων θεωρήσαμε ότι είχαν επαρκή δεδομένα, τα οποία μπορούμε να εφαρμόσουμε στα μοντέλα μας, όπως θα αναλυθεί και στην Ενότητα 3.3. Στα δημογραφικά στοιχεία, περιλαμβάνουμε επίσης το BMI (δείκτης σώματος-μάζας).

	BP	SZ	Schizoaffective Disorder	Brief Psychotic Episode	Schizophreniform Disorder
<b>Demographics</b>					
Male/Female	3/1	1/1	1/1	0/1	1/0
Age (years)	27.8 ± 6.0	31.0 ± 17.0	33.5 ± 2.1	30.0	36.0
Education (years)	13.0 ± 2.0	14.0 ± 2.8	13.0 ± 1.3	16.0	16.0
Illness dur. (years)	7.25 ± 5.6	12.0 ± 15.6	8.0 ± 2.8	3.0	1.0
BMI	27.2 ± 5.6	33.0 ± 12.0	27.1 ± 0.1	–	–

Πίνακας 3.1: Δημογραφικά στοιχεία για τους ασθενείς με Σχιζοφρένεια (SZ), διπολική διαταραχή I (BP), Schizoaffective disorder, Brief Psychotic Episode και Schizophreniform disorder. Ο πίνακας περιλαμβάνει στοιχεία μόνο για τους 10 από τους 24 ασθενείς, αφού αυτά είναι και τα δεδομένα που χρησιμοποιήθηκαν για την συγκεκριμένη εργασία.

Οι ασθενείς υποβάλλονται σε μηνιαίες αξιολογήσεις από τους κλινικούς ιατρούς του έργου. Οι ιατροί σημειώνουν τις περιόδους του ασθενή που δεν έχει εμφανίσει συμπτώματα ψυχωτικής υποτροπής ως ‘Φυσιολογική’ (Normal-N) και ως ‘Υποτροπή’ (Relapse-R) όταν ο ασθενής έχει εμφανίσει συμπτώματα. Επίσης, ο κλινικός ιατρός χαρακτηρίζει τη σοβαρότητα της υποτροπής ως low, moderate ή severe. Επιπλέον, σύμφωνα με τις υποδείξεις των ιατρών θεωρήσαμε το διάστημα 21 ημερών πριν την υποτροπή μια ξεχωριστή κατηγορία και την ονομάσαμε ως περίοδος ‘Σχεδόν Υποτροπής’ (Near Relapse-NR).

Τα ανεπεξέργαστα δεδομένα των ασθενών καταγράφονται συνεχώς (24/7) (εκτός περίπου 2 ωρών που το ρολόι είναι σε φόρτιση), από ένα έξυπνο ρολόι Samsung Gear S3 Frontier που διαθέτουν όλοι οι ασθενείς και συλλέγονται σε μια πλατφόρμα που βασίζεται σε cloud [Mag+20]. Η περίοδος συλλογής των δεδομένων για κάθε ασθενή είναι διαφορετική, συνολικά όμως τα δεδομένα συλλέχθηκαν από το Νοέμβριο του 2019 μέχρι το Σεπτέμβρη του 2021. Σε αντίθεση με προηγούμενες έρευνες όπου τα δεδομένα ήταν για βραχυπρόθεσμες διάρκειες, στη δική μας εργασία τα δεδομένα ήταν μακροπρόθεσμα με διάρκεια 1 χρόνου και 10 μηνών. Τα δεδομένα περιέχουν μετρήσεις της γραμμικής επιτάχυνσης σε 3 άξονες (acc), της γωνιακής επιτάχυνσης σε 3 άξονες (gyr), τους καρδιακούς παλμούς κάθε λεπτό (hrm) και τα RR διαστήματα (χρονικά διαστήματα μεταξύ δύο διαδοχικών καρδιακών παλμών), τα βήματα και τη διανυόμενη απόσταση και το πρόγραμμα του ύπνου. Αυτά τα δεδομένα καταγράφονται από τους αισθητήρες του επιταχυνσιόμετρου, του γυροσκοπίου, του καρδιακού παλμού, του βηματομετρητή και του ύπνου, αντίστοιχα. Οι δύο πρώτες μετρήσεις δειγματοληπτούνται στα 20Hz, οι μετρήσεις από τον αισθητήρα του καρδιακού παλμού δειγματοληπτούνται στα 5Hz ενώ για τα βήματα και την διανυόμενη απόσταση καταγράφεται ο συνολικός αριθμός βημάτων ανά λεπτό, χρησιμοποιώντας το Tizen API [Mag+20] που έχει το έξυπνο ρολόι. Κατά τη συλλογή των δεδομένων, στο έργο e-Prevention υπάρχουν συνολικά 20 ασθενείς που αντιμετώπισαν τουλάχιστον μία ψυχωτική υποτροπή. Από τους 20 ασθενείς μετά από τη διαδικασία προεπεξεργασίας των δεδομένων μειώνεται σε 10, αφού θεωρήσαμε ότι δεν υπήρχαν αρκετά δεδομένα, για να μπορέσουμε να τα χρησιμοποιήσουμε για την μοντελοποίηση του προβλήματος. Σύμφωνα με τους ιατρούς οι σημειωμένες υποτροπές των 10 ασθενών (σε μέρες) παρουσιάζονται στον Πίνακα 3.2.

Παρατηρούμε ότι μόνο ο ασθενής #10 είχε severe υποτροπές και ότι ο ασθενής #2 είχε μόνο μια low υποτροπή για μικρό χρονικό διάστημα.

Αφού έγινε η συλλογή των δεδομένων και δημιουργήσαμε το dataset μας, χωρίσαμε τα δεδομένα μας σε τρεις περιόδους (Normal, Near Relapse και Relapse). Στον Πίνακα 3.3 παρουσιάζεται η διάρκεια σε μέρες της κάθε περιόδου για κάθε ασθενή.

Patients	Relapses (Days per Relapse)			
	Low	Moderate	Severe	Total
#1	75	21	-	96
#2	7	-	-	7
#3	30, 30, 61, 25	-	-	146
#4	-	48	-	48
#5	60	61	-	121
#6	6	28, 19, 80	-	133
#7	11, 41	40	-	92
#8	36	-	-	36
#9	31	-	-	31
#10	-	5	29, 60	94
Total	413	302	89	<b>804</b>

Πίνακας 3.2: Υποτροπές ασθενών ανά κατηγορία υποτροπής σε μέρες σύμφωνα με τις επισημειώσεις των ιατρών.

Από τους Πίνακες 3.2 και 3.3, παρατηρούμε ότι ενώ οι ιατροί έχουν σημειώσει ότι οι μέρες υποτροπής για τους ασθενείς είναι περίπου 804, στο dataset μας εμείς έχουμε σχεδόν τις μισές (416). Αυτό, οφείλεται πιθανόν στο γεγονός ότι πολλές φορές όταν οι ασθενείς πάθουν κάποια υποτροπή αφαιρούν το ρολόι τους και έτσι χάνονται τα δεδομένα, δηλαδή δεν είναι συνεπείς για τις ώρες που πρέπει να φοράνε το ρολόι.

Patients	Normal	Near Relapse	Relapse	Total
#1	243	15	48	306
#2	431	13	4	448
#3	279	44	78	401
#4	254	13	27	294
#5	220	9	50	279
#6	172	49	78	299
#7	249	38	68	355
#8	203	11	18	232
#9	55	7	12	74
#10	130	26	33	189
Total	2236	225	<b>416</b>	2877

Πίνακας 3.3: Αρχικό Dataset. Πίνακας με τις μέρες στο Dataset ανά κατηγορία πριν τη προεπεξεργασία δεδομένων.

## 3.2 Ανάλυση Δεδομένων

Το σύνολο δεδομένων που χρησιμοποιήθηκε αρχικά περιλαμβάνει κάποια από τα ανεπεξέργαστα (raw) δεδομένα. Τα raw δεδομένα που χρησιμοποιήσαμε αποτελούνται από τη γραμμική επιτάχυνση, τη γωνιακή επιτάχυνση, τους καρδιακούς παλμούς και το RR-interval του καρδιακού ρυθμού. Τα δεδομένα μας, ομαδοποιήθηκαν σε διαστήματα 5 λεπτών (5-min) και για

κάθε διάστημα υπολογίστηκε η μέση ενέργεια της γραμμικής επιτάχυνσης και η νόρμα της γωνιακής επιτάχυνσης (αφού μετριοούνται τριαξονικά), ο μέσος καρδιακός ρυθμός και το RR-interval του καρδιακού ρυθμού. Η εξαγωγή της μέση ενέργεια της γραμμικής επιτάχυνσης και της νόρμα της γωνιακής ταχύτητας αναλύεται πιο κάτω.

Στη συνέχεια σε αυτά τα raw δεδομένα προστέθηκαν κάποια επιπλέον χαρακτηριστικά αφού σε προηγούμενη στατιστική μελέτη [Fil+20] διαπιστώθηκε ότι διαφέρουν σημαντικά όταν ένα άτομο πάσχει από κάποια ψυχωτική διαταραχή. Επομένως, εκμεταλλευτήκαμε κάποια από τα χαρακτηριστικά που εξήχθησαν από τη μελέτη αυτή [Fil+20] με χρήση τεχνικών επεξεργασίας σήματος και τα προσθέσαμε στο δικό μας σύνολο δεδομένων. Για το δικό μας σύνολο δεδομένων (dataset), αρχικά μετασχηματίσαμε τα δεδομένα σε μια πιο πυκνή αναπαράσταση και τα ομαδοποιήσαμε σε διαστήματα 5 λεπτών. Συγκεντρωτικά, εξήχθησαν 10 χαρακτηριστικά για κάθε διάστημα:

- Ενέργεια: Η ενέργεια βραχέος χρόνου (STE) της γραμμικής επιτάχυνσης και της γωνιακή επιτάχυνσης (αφού μετριοούνται τριαξονικά). Επομένως, υπολογίστηκε η μέση ενέργεια της γραμμικής επιτάχυνσης και η νόρμα της γωνιακής επιτάχυνσης.
- Μέσος καρδιακός ρυθμός.
- Διάστημα RR του καρδιακού ρυθμού.
- Φασματικά χαρακτηριστικά (Spectral features): Δεδομένου ότι η μεταβλητότητα του καρδιακού ρυθμού (HRV) είναι εξ ορισμού ένα μη ομοιόμορφο σήμα, εκτελείται φασματική ανάλυση χρησιμοποιώντας το περιοδόγραμμα Lomb-Scargle και εξάγουμε για κάθε διάστημα τη σχετική ισχύ και τη κανονικοποιημένη ισχύ σε δύο ζώνες LF (χαμηλή συχνότητα, 0.04 – 0.15Hz) και HF (υψηλή συχνότητα, 0.15 – 0.40Hz).
- Poincare plot measure: Το διάγραμμα Poincare [BPK01] είναι ένας τύπος διαγράμματος ο οποίος χρησιμοποιείται για να ποσοτικοποιήσει την αυτό-ομοιότητα συνήθως σε περιόδους συναρτήσεων. Για το σχεδιασμό του, κάθε δείγμα μιας χρονοσειράς σχεδιάζεται σε σχέση με το προηγούμενο και στη συνέχεια τοποθετείται μια έλλειψη στο διάγραμμα της διασποράς. Το πλάτος της έλλειψης (SD1) αποτελεί ένδειξη του HRV σε βραχύ χρόνο, ενώ το μήκος της σε μεγαλύτερα χρονικά διαστήματα
- Το ποσοστό των σωστά προσδιορισμένων παλμών.
- Ημιτονοειδής και Συνημιτονοειδής αναπαράσταση των δευτερολέπτων. Τα χαρακτηριστικά αυτά μας βοηθάνε, για να μοντελοποιήσουμε τη χρονολογική σειρά των χαρακτηριστικών.

Στο Σχήμα 3.2.1 παρουσιάζονται τα 10 χαρακτηριστικά που εξήχθησαν και περιγράφηκαν παραπάνω.



Σχήμα 3.2.1: Παράδειγμα χαρακτηριστικών, που χρησιμοποιήθηκαν σε αυτή τη μελέτη, κατά τη διάρκεια της ημέρας ενός ασθενή.

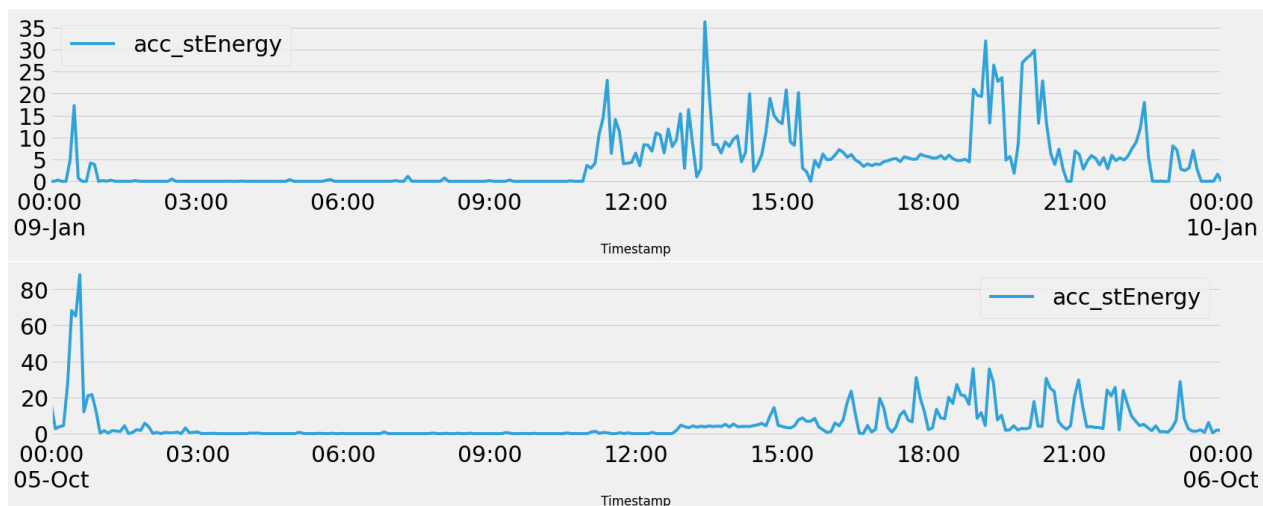
Η τεχνικές επεξεργασίας σήματος που χρησιμοποιήθηκαν αναλύθηκαν εκτενέστερα στην Ενότητα 2.1 Η μεταβλητότητα του καρδιακού ρυθμού (HRV) από το σήμα που δειγματολη-



πτείται στα 5Hz συνίστανται στην απόρριψη των ίδιων διαδοχικών τιμών, διασφαλίζοντας ότι για κάθε διάστημα μιας ώρας η λαμβανόμενη ακολουθία διαστημάτων RR αθροίζεται σε τουλάχιστον 54 λεπτά (ένα εμπειρικό όριο που αντιστοιχεί στο 90% έγκυρων καρδιακών δεδομένων). Αφαιρούνται επίσης διαστήματα RR μεγαλύτερα από 2000 ms και μικρότερα από 300 ms και αντικαθίστανται πιθανοί μη ανιχνευθέντες παλμοί χρησιμοποιώντας γραμμική παρεμβολή (linear interpolation). Δεν πραγματοποιείται κανένα είδους μείωση θορύβου στη γραμμική και γωνιακή επιτάχυνση, καθώς διαπιστώσαμε ότι για τα εξεταζόμενα χαρακτηριστικά η επίδραση του θορύβου είναι αμελητέα [Fil+20].

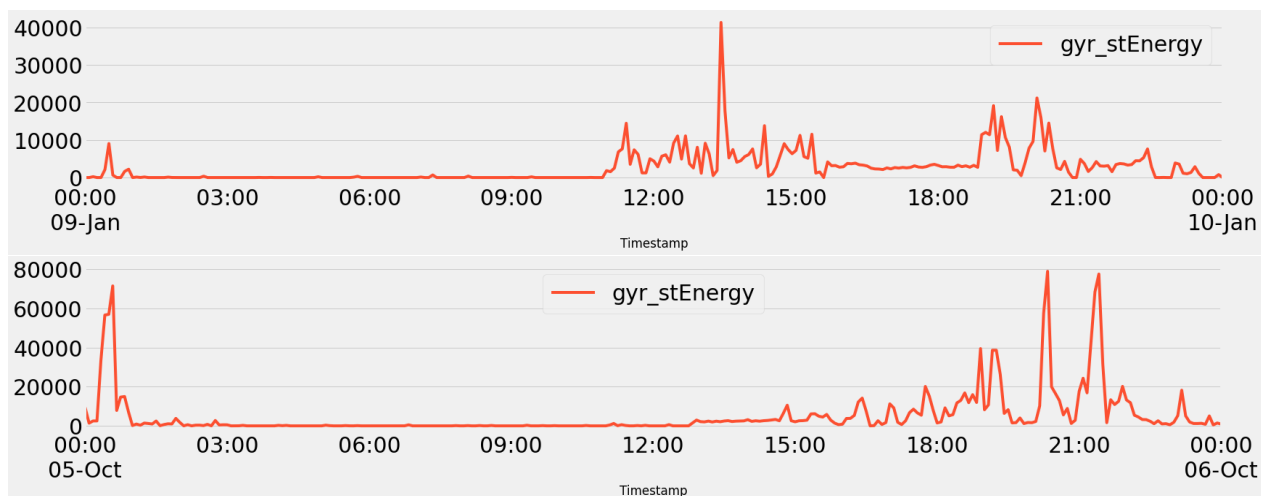
Το χαρακτηριστικό του ποσοστού των σωστά προσδιορισμένων παλμών προστέθηκε για να έχουμε ένα μέτρο του πόσο καλά φοράει ο ασθενής το ρολόι και τα δύο τελευταία χαρακτηριστικά που έχουν προστεθεί είναι για να μοντελοποιήσουμε τη χρονολογική σειρά των χαρακτηριστικών, αφού τα μοντέλα μας δεν λαμβάνουν ως χαρακτηριστικό το timestamp, και έτσι δεν μπορούν ουσιαστικά να καταλάβουν σε ποια περίοδο της ημέρας ανήκουν οι συγκεκριμένες μετρήσεις [Lon16]. Επιπλέον, επιλέχθηκε η τιμή των 5 λεπτών για τη διάρκεια του διαστήματος, αφού είναι αρκετά μικρή για να μπορούμε να καταγράψουμε μοτίβα μικροκλίμακας [Ret+] και επίσης μας επιτρέπει να έχουμε επαρκές μέγεθος δεδομένων για τις τεχνικές βαθιάς μάθησης. Επίσης, για να επιλέξουμε τη τιμή των 5-λεπτών για την διάρκεια διαστήματος έχουν γίνει πειράματα και για διάρκεια διαστήματος 10-λεπτών και 1-ώρας, τα οποία παρουσιάζονται στην Ενότητα 5.2.

Για να δείξουμε τις διαφορές των χαρακτηριστικών ανάμεσα σε μέρες υποτροπής και σε κανονικές μέρες, θα παραθέσουμε πιο κάτω κάποιες γραφικές από τα διάφορα χαρακτηριστικά. Στα Σχήματα 3.2.2 3.2.3, 3.2.4 3.2.5, 3.2.6 και 3.2.7 παρουσιάζονται τα χαρακτηριστικά της μέσης ενέργειας της γραμμικής επιτάχυνσης, της νόρμας της γωνιακής επιτάχυνσης, της μέσης τιμής του καρδιακού ρυθμού, του διαστήματος RR του καρδιακού ρυθμού, το SD1 και οι κανονικοποιημένες ισχύς στις ζώνες LF και HF της μεταβλητότητας του καρδιακού ρυθμού, αντίστοιχα για μια κανονική μέρα και μια μέρα severe υποτροπής του ασθενή #10. Για αυτά τα σχήματα οι καταγραφόμενες τιμές αντιπροσωπεύουν ενδείξεις εντός ενός 24ώρου.



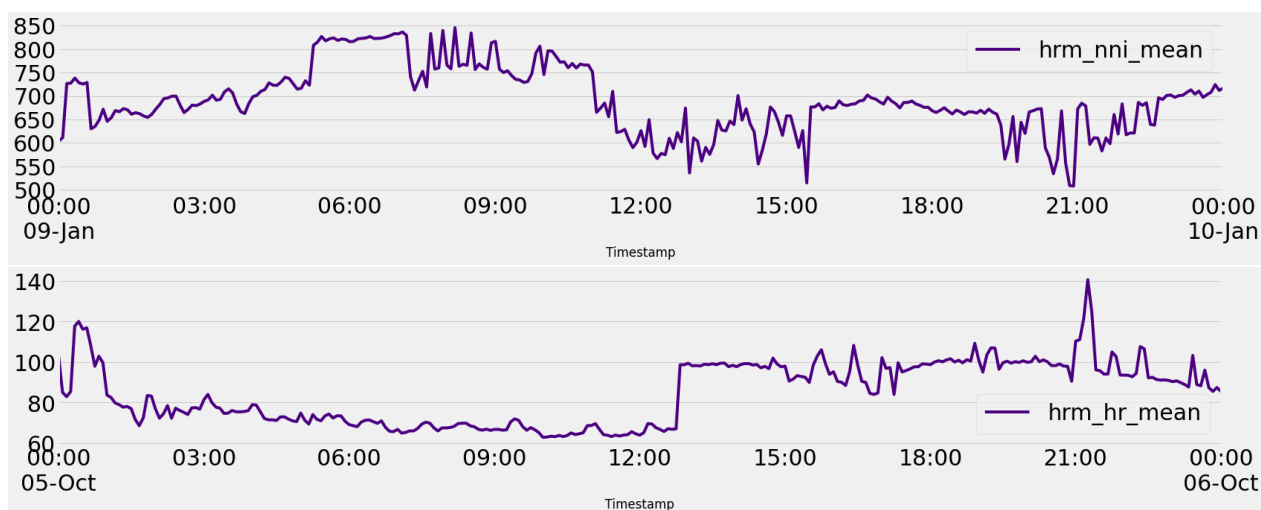
Σχήμα 3.2.2: Χαρακτηριστικό μέσης ενέργειας γραμμικής επιτάχυνσης του ασθενή #10. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γραμμικής επιτάχυνσης για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.





Σχήμα 3.2.3: Χαρακτηριστικό νόρμα γωνιακής επιτάχυνσης του ασθενή #10. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γωνιακής επιτάχυνσης για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.

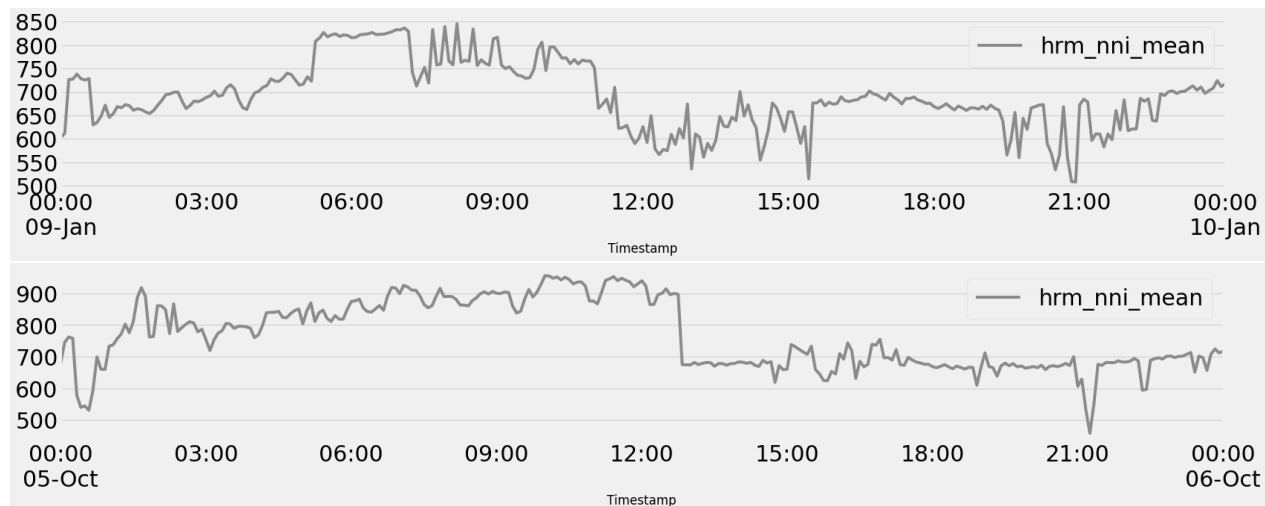
Στα χαρακτηριστικά της μέσης ενέργειας της γραμμικής επιτάχυνσης και της γωνιακής ταχύτητας, τα οποία σχετίζονται με την κινητικότητα του ασθενή, παρατηρείται μεγάλη μεταβλητότητα ως προς αυτή. Δηλαδή, παρατηρούμε από τον κάθετο άξονα ότι οι τιμές στη μέρα υποτροπής είναι μεγαλύτερες, δηλαδή η μέγιστη τιμή στην κανονική μέρα έχει περίπου τη μισή τιμή από την αντίστοιχη τιμή στη μέρα της υποτροπής. Επίσης, σε σχέση με την κανονική μέρα, στη μέρα υποτροπής παρατηρείται μια μεγαλύτερη περίοδος ακινησίας και ξαφνική έντονη κινητικότητα. Η περίοδος ακινησίας θα μπορούσε να θεωρηθεί ως περίοδος ύπνου, δηλαδή παρατηρείται ότι τη μέρα της υποτροπής ο ασθενής πιθανό να κοιμάται περισσότερες ώρες.



Σχήμα 3.2.4: Χαρακτηριστικό μέσης τιμής καρδιακού ρυθμού του ασθενή #10. Στην πάνω γραφική παρουσιάζεται ο καρδιακός ρυθμός για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.

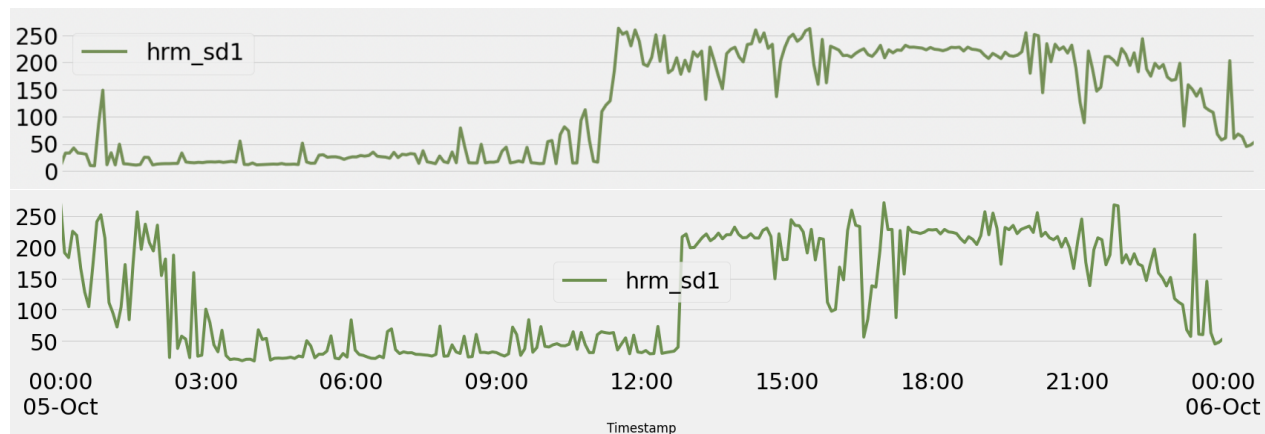
Στην πιο πάνω γραφική όπου περιγράφεται το χαρακτηριστικό της μέσης τιμής του καρδιακού

ρυθμού, αν και δεν παρατηρούνται έντονες διαφορές μεταξύ της κανονικής μέρας και της μέρας υποτροπής, αξίζει να σημειωθεί ότι υπάρχει μια απότομη αύξηση του καρδιακού ρυθμού στους 140 παλμούς, τη μέρα της υποτροπής. Διαισθητικά, αυτή η απότομη αύξηση μπορεί να συμβεί λόγω αυξημένους άγχους που να οφείλεται στην υποτροπή.



Σχήμα 3.2.5: Χαρακτηριστικό του διαστήματος RR του καρδιακού ρυθμού του ασθενή #10. Στην πάνω γραφική παρουσιάζεται το διάστημα RR του καρδιακού ρυθμού για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.

Στη γραφική του Σχήματος 3.2.5, αν και η μορφή των δύο χρονοσειρών είναι παρόμοια παρατηρούμε ότι οι τιμές του διαστήματος RR του καρδιακού ρυθμού τη μέρα της severe υποτροπής είναι μεγαλύτερες από τις αντίστοιχες της κανονικής μέρας.

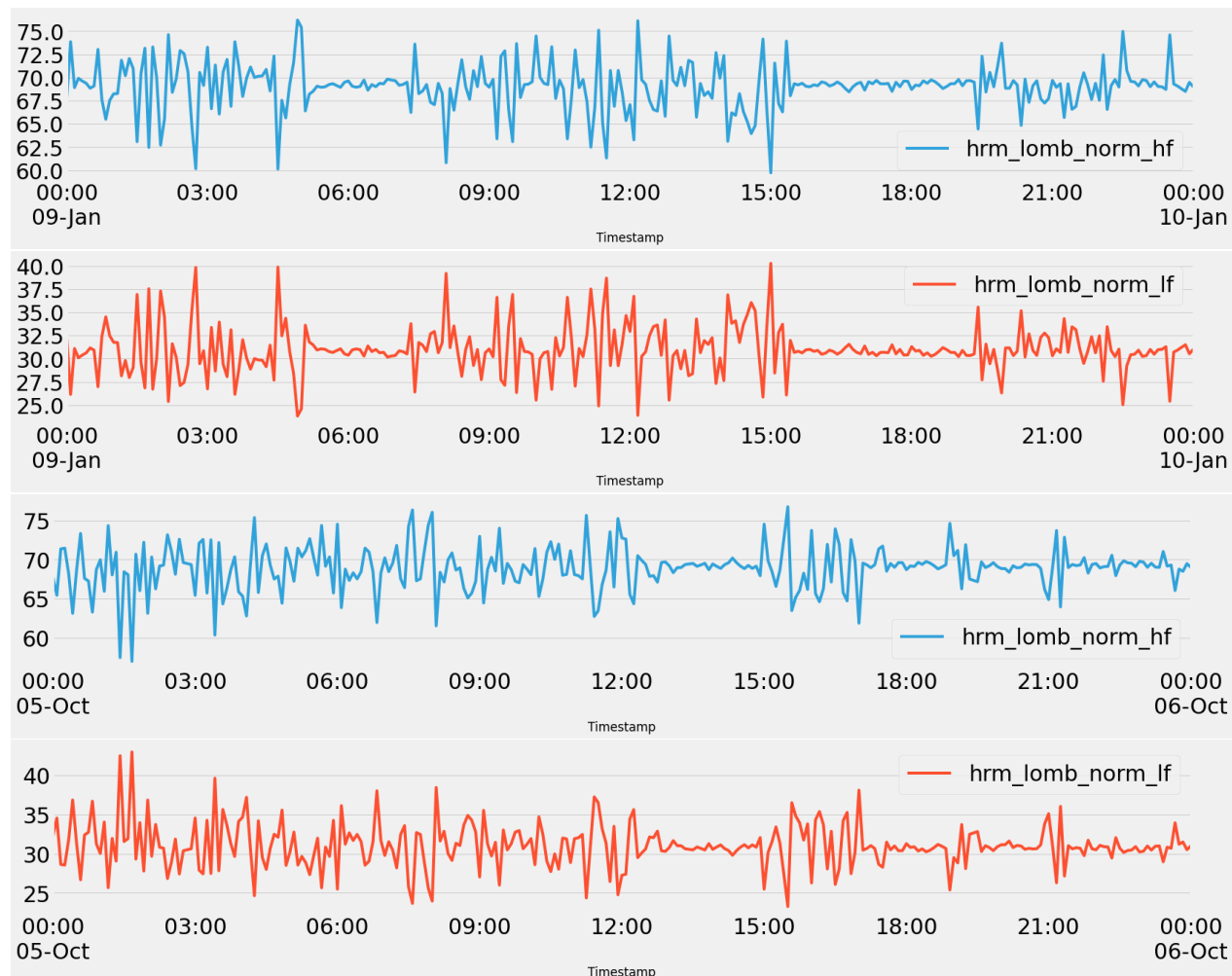


Σχήμα 3.2.6: Χαρακτηριστικό SD1 της μεταβλητότητας του καρδιακού ρυθμού του ασθενή #10. Στην πάνω γραφική παρουσιάζεται το SD1 για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.

Στο Σχήμα 3.2.6 παρουσιάζεται η γραφική παράσταση του χαρακτηριστικού SD1 της μεταβλητότητας του καρδιακού ρυθμού. Δεν μπορούμε να ερμηνεύσουμε πλήρως αυτές τις γραφικές

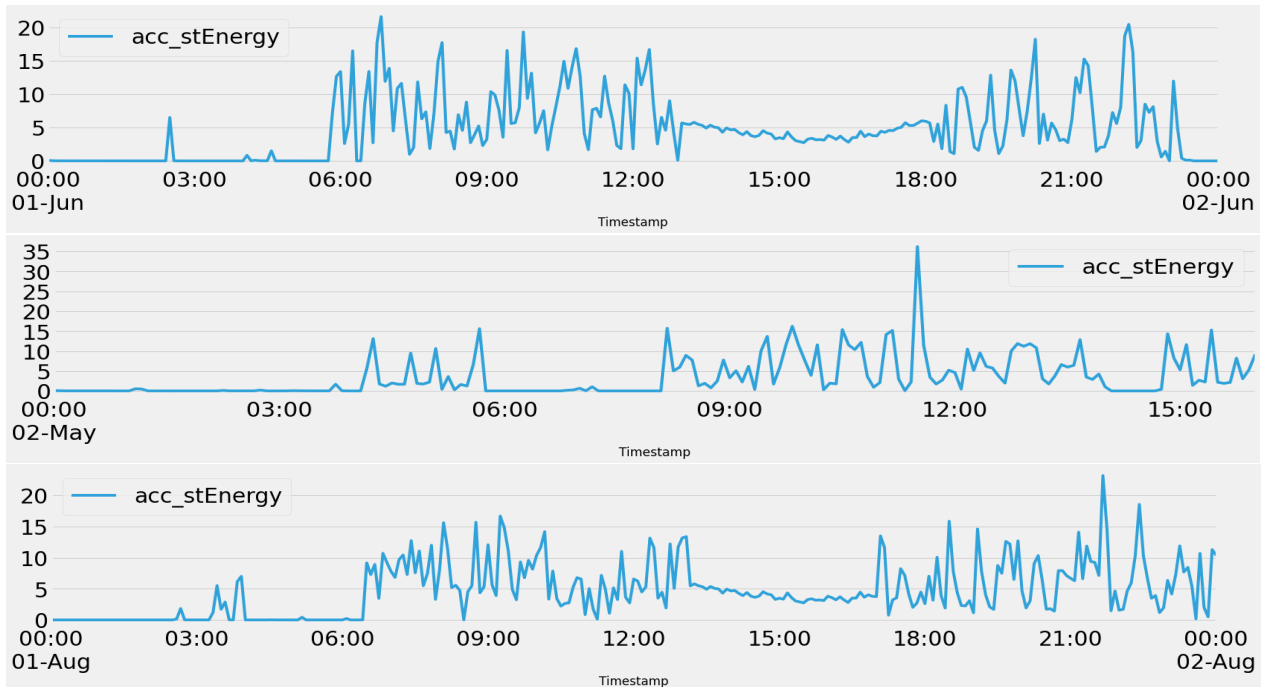
αφού δεν είμαστε ειδικοί, παρ'όλα αυτά παρατηρούμε έντονες εναλλαγές ή ταλαντώσεις κατά τη μέρα της severe υποτροπής συγκριτικά με την κανονική μέρα.

Στο Σχήμα 3.2.7 όπου παρουσιάζονται τα χαρακτηριστικά των κανονικοποιημένων LF και HF της μεταβλητότητας του καρδιακού ρυθμού, δεν παρατηρούμε έντονες διαφορές ανάμεσα στις δύο μέρες για το συγκεκριμένο χαρακτηριστικό.

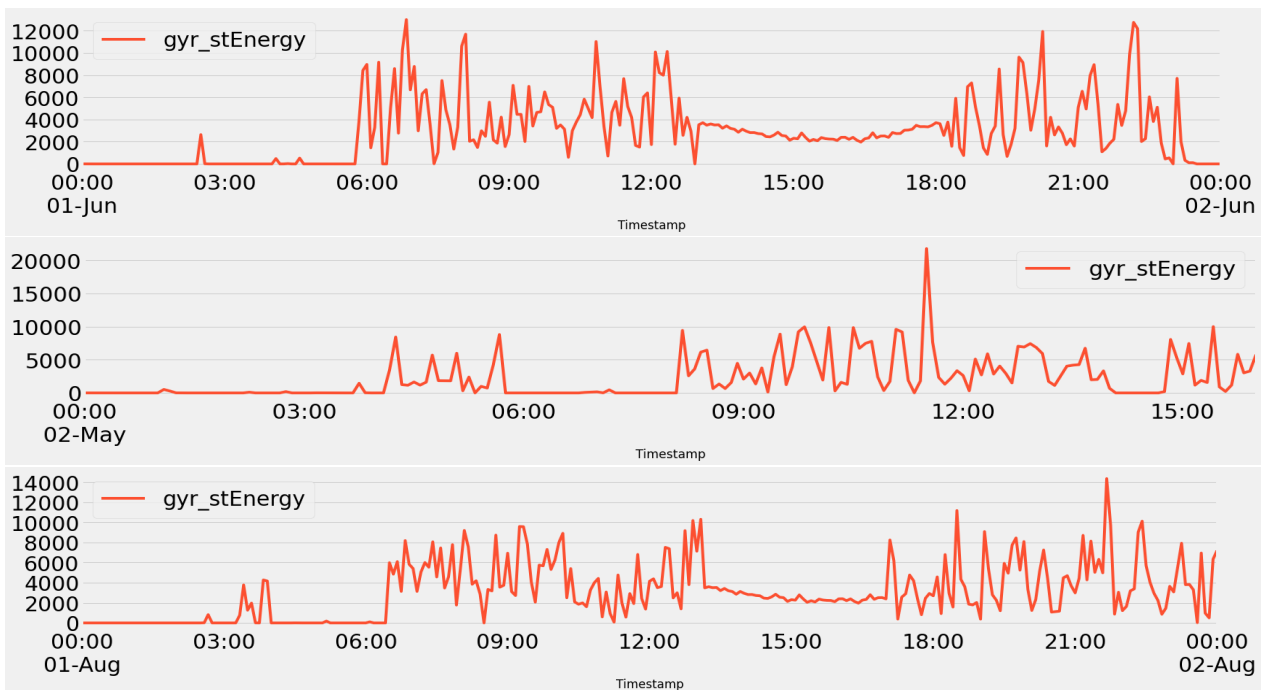


Σχήμα 3.2.7: Χαρακτηριστικό των κανονικοποιημένων ζώνων LF και HF της μεταβλητότητας του καρδιακού ρυθμού. Στην πάνω γραφική παρουσιάζονται οι κανονικοποιημένοι LF και HF του ασθενή #10 για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.

Στα Σχήματα 3.2.8 3.2.9, 3.2.10 3.2.11, 3.2.12 και 3.2.13 παρουσιάζονται τα χαρακτηριστικά της μέσης ενέργειας της γραμμικής επιτάχυνσης, της νόρμας της γωνιακής επιτάχυνσης, της μέσης τιμής του καρδιακού ρυθμού, του διαστήματος RR του καρδιακού ρυθμού, το SD1 και οι κανονικοποιημένες ισχύς στις ζώνες LF και HF της μεταβλητότητας του καρδιακού ρυθμού, αντίστοιχα για μια κανονική μέρα, μια μέρα low υποτροπής και μια μέρα moderate υποτροπής του ασθενή #7. Όπως προαναφέραμε οι τιμές των γραφικών παραστάσεων είναι εντός ενός 24ώρου.



Σχήμα 3.2.8: Χαρακτηριστικό μέσης ενέργειας γραμμικής επιτάχυνσης του ασθενή #7. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γραμμικής επιτάχυνσης για μια κανονική μέρα ενώ στην κάτω γραφική για μια μέρα severe υποτροπής.

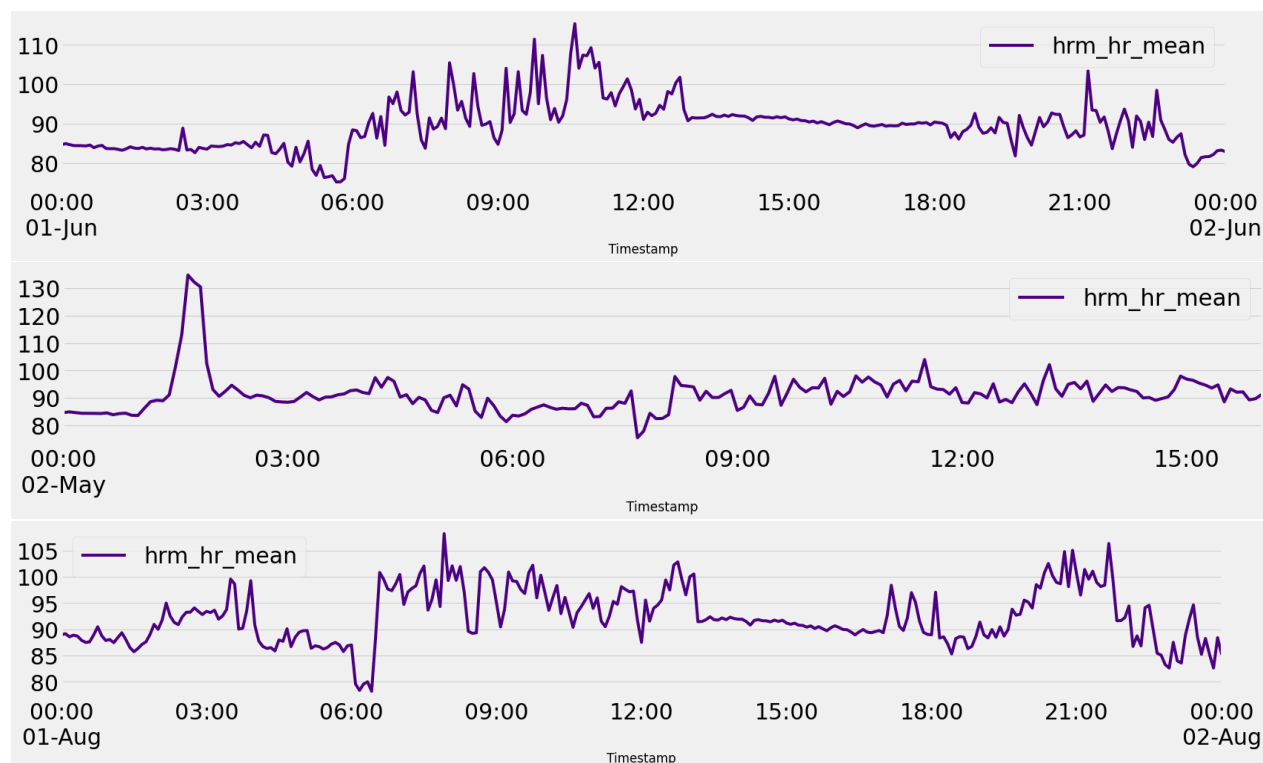


Σχήμα 3.2.9: Χαρακτηριστικό μέσης ενέργειας γωνιακής επιτάχυνσης του ασθενή #7. Στην πάνω γραφική παρουσιάζεται η μέση ενέργεια γωνιακής επιτάχυνσης για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής.

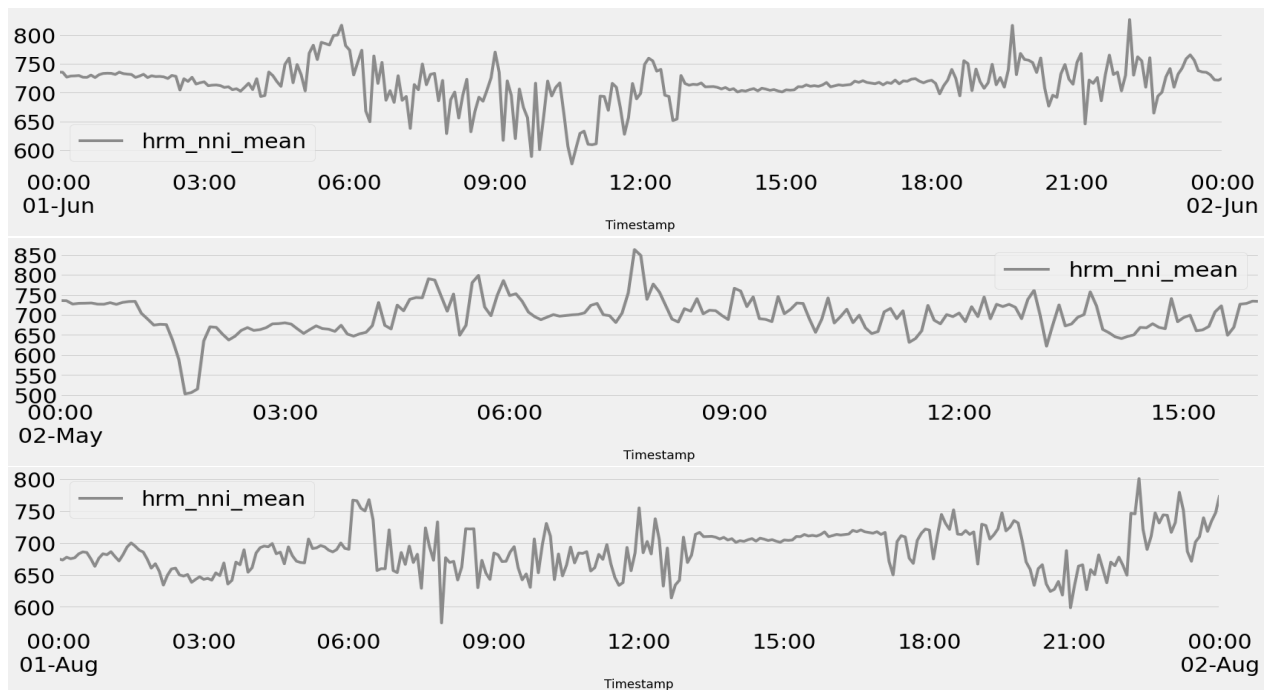
Και για τις δύο πιο πάνω γραφικές παραστάσεις, παρατηρούμε ότι η μορφή των αντίστοιχων χαρακτηριστικών για την κανονική μέρα και τη μέρα moderate υποτροπής είναι αρκετά παρόμοια, ενώ η μορφή τη μέρα της low υποτροπής είναι διαφορετική. Στη μέρα low υποτροπής παρατηρείται μεγαλύτερη ακινησία ή διάρκεια ύπνου συγκριτικά με τις υπόλοιπες μέρες (κανονικής και moderate υποτροπής). Στη low υποτροπή παρατηρείται στις ώρες εγρήγορσης πιο έντονη κινητικότητα, με μια μεγάλη απότομη αύξηση. Στη moderate υποτροπή από την γραφική παράσταση της γωνιακής επιτάχυνσης παρατηρούμε μεγαλύτερες τιμές της γωνιακής επιτάχυνσης για τη μέρα αυτή σε αντίθεση με την κανονική μέρα.

Στα Σχήματα 3.2.10 και 3.2.11, παρατηρούμε ότι στη low υποτροπή έχουμε μία απότομη αύξηση των παλμών, διαισθητικά μπορούμε να πούμε ότι υπάρχουν διάφοροι λόγοι για αυτή την απότομη αύξηση, όπως λόγω άγχους. Στη moderate υποτροπή παρατηρούμε γενικά πιο ήπιες τιμές στη μέση τιμή του καρδιακού ρυθμού από ότι και στις κανονικές μέρες.

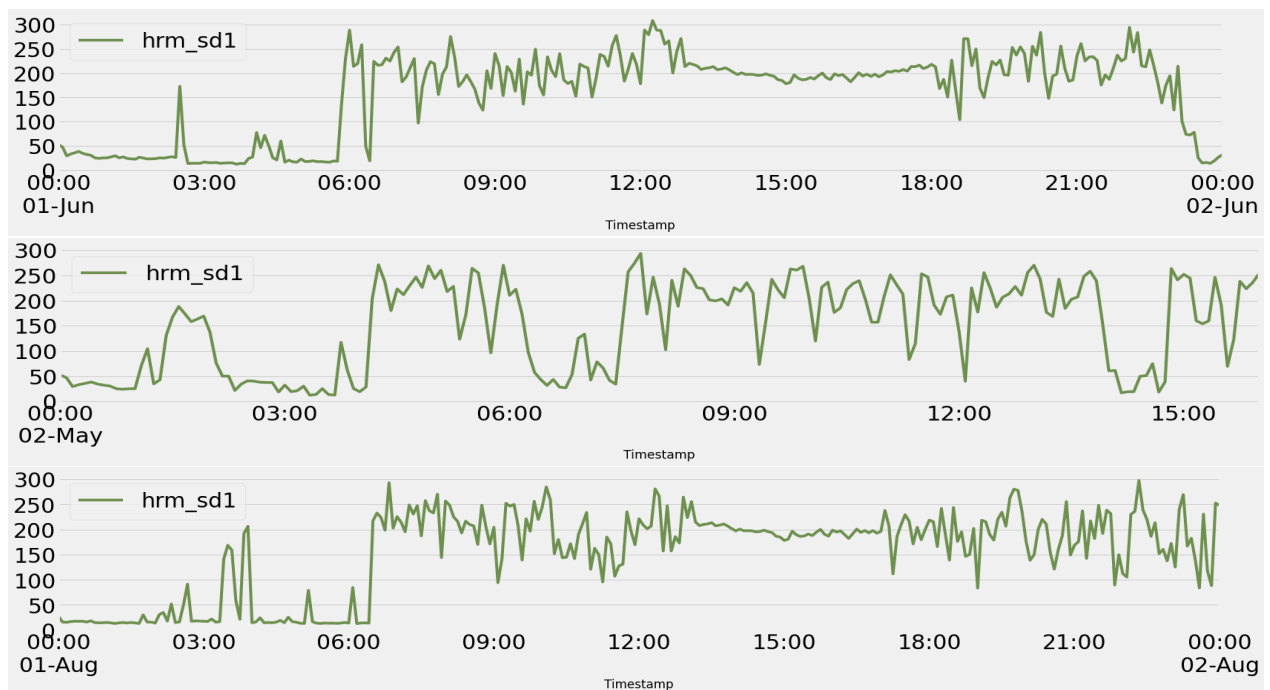
Παρόμοια με τις προηγούμενες μας παρατηρήσεις, στο Σχήμα 3.2.12 βλέπουμε ότι η μορφή της γραφικής για τη κανονική μέρα και τη moderate μέρα είναι αρκετά παρόμοιες καθώς για ένα χρονικό διάστημα στη μέση περίπου της μέρας φαίνεται να έχουμε σταθερές τιμές. Αντίθετα, τη μέρα της low υποτροπής, δεν παρατηρείται σταθερότητα στις τιμές, αφού ταλαντώνονται συνεχώς.



Σχήμα 3.2.10: Χαρακτηριστικό μέσης τιμές καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζεται ο καρδιακός ρυθμός για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής.

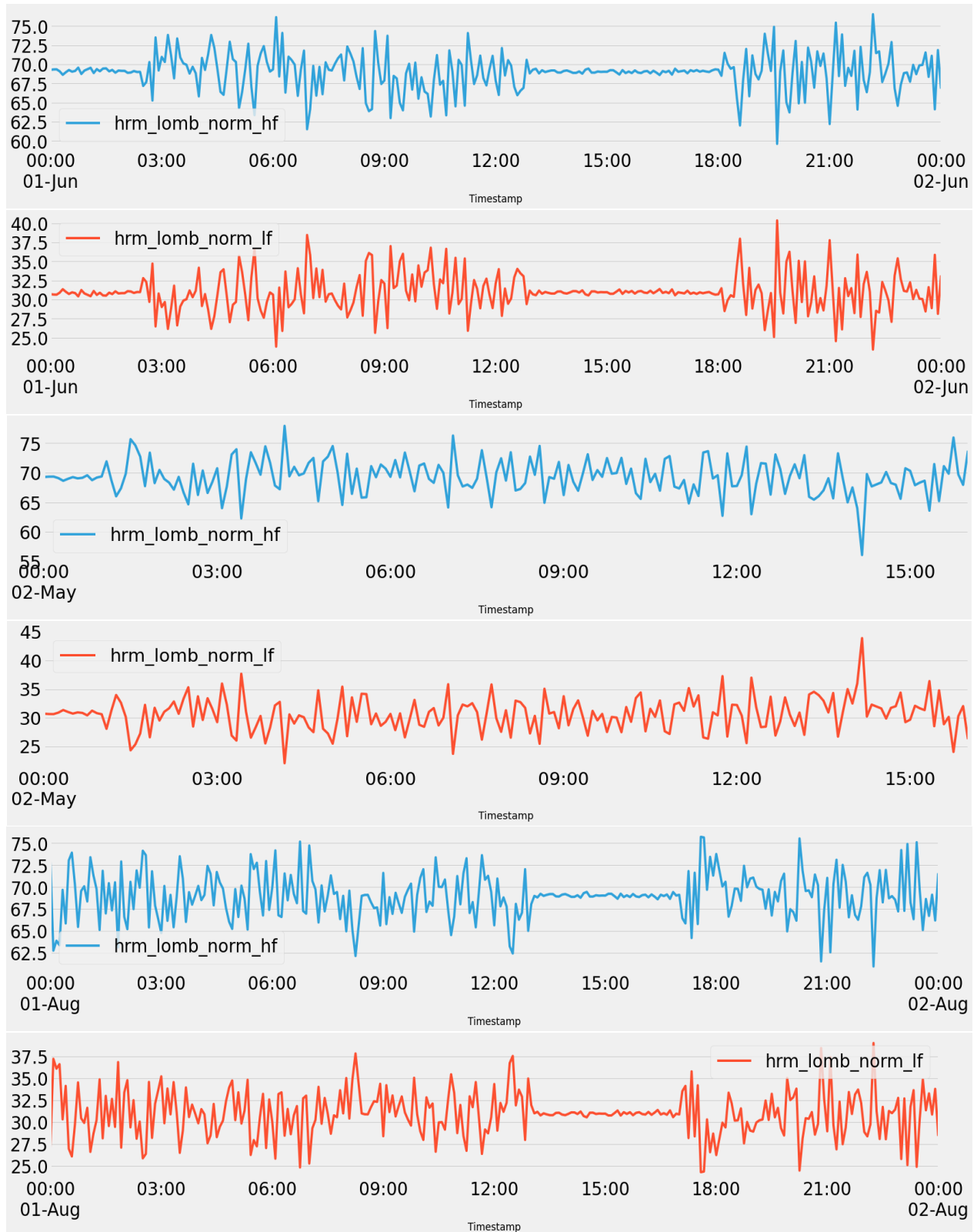


Σχήμα 3.2.11: Χαρακτηριστικό του διαστήματος RR του καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζεται το διάστημα RR του καρδιακού ρυθμού για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής.



Σχήμα 3.2.12: Χαρακτηριστικό SD1 της μεταβλητότητας του καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζεται το SD1 για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής.





Σχήμα 3.2.13: Χαρακτηριστικό των κανονικοποιημένων ζωνών LF και HF της μεταβλητότητας του καρδιακού ρυθμού του ασθενή #7. Στην πάνω γραφική παρουσιάζονται οι κανονικοποιημένες ζώνες LF και HF για μια κανονική μέρα, στην επομένη γραφική για μια μέρα low υποτροπής και στην κάτω γραφική για μια μέρα moderate υποτροπής.

Στο Σχήμα 3.2.13 βλέπουμε ότι η μορφή της γραφικής παράστασης του χαρακτηριστικού των κανονικοποιημένων ζωνών LF και HF για τη κανονική μέρα και τη moderate μέρα έχουν παρόμοια μορφή παρουσιάζοντας σταθερές τιμές χωρίς ταλαντώσεις λίγο μετά το μέσο της μέρας. Επίσης, έχουν και κάποιες διαφορές όπως στην αρχή του 24ώρου όπου ενώ στη κανονική μέρα έχουμε σταθερές τιμές, στη moderate μέρα υποτροπής έχουμε πιο πυκνές ταλαντώσεις. Αντίθετα, στη μέρα της low υποτροπής, δεν παρατηρείται σταθερότητα στις τιμές, αφού ταλαντώνονται συνεχώς.

Συμπερασματικά σύμφωνα με τα πιο πάνω ενδεικτικά σχήματα, στα χαρακτηριστικά που παρατηρούνται μεγαλύτερες διαφορές ανάμεσα στις μέρες όπου οι ασθενείς έχουν υποτροπή με τις κανονικές μέρες είναι στη γραμμική και γωνιακή επιτάχυνση, στα διαστήματα RR του καρδιακού ρυθμού και στις κανονικοποιημένες ζώνες LF και HF της μεταβλητότητας του καρδιακού ρυθμού. Ανάλογα με τη σοβαρότητα της υποτροπής, παρατηρούμε και πιο έντονες εναλλαγές στα πιο πάνω χαρακτηριστικά. Τα υπόλοιπα χαρακτηριστικά, όπως οι ημιτονοειδής και συνημιτονοειδής αναπαραστάσεις των δευτερολέπτων μας δίνουν πληροφορίες κατά την εκπαίδευση του μοντέλου.

### 3.3 Προεπεξεργασία Δεδομένων

Σε πολλές περιπτώσεις οι ασθενείς δεν φορούν το ρολόι συνεχόμενα, με αποτέλεσμα να έχουμε πολλά χαμένα δεδομένα για ένα χρονικό διάστημα. Επομένως, για την αντιμετώπιση των χαμένων δεδομένων αποφασίσαμε για το χρονικά διάστημα με διάρκεια έως και 10 συνεχόμενων ωρών να τα συμπληρώνουμε με τη μέθοδο των διαμέσων τιμών (median) του χαρακτηριστικού. Επίσης, εξετάστηκαν άλλες δύο μέθοδοι πλήρωσης, είτε με τις αντίστοιχες μέσες τιμές (mean), είτε με το μέσο όρο των υπάρχουσων τιμών (neighbor) μεταξύ των χρόνων που λείπουν. Εμπειρικά διαπιστώσαμε ότι η μέθοδος ‘median’ λειτουργεί καλύτερα στις περισσότερες περιπτώσεις. Τέλος, σε περιπτώσεις όπου έλειπαν δεδομένα για περισσότερες από 10 συνεχόμενες ώρες τα αγνοήσαμε.

Μετά την πιο πάνω προεπεξεργασία, τα δεδομένα κάθε ασθενούς θεωρούνται ως πολυμεταβλητές χρονοσειρές  $\mathbf{X}_{L \times d}$ , όπου  $L$  είναι το συνολικό μήκος της χρονοσειράς και  $d$  ο αριθμός των χαρακτηριστικών. Τα σημεία της χρονοσειράς είναι σε ανά πεντάλεπτα διαστήματα, σύμφωνα με την προηγούμενη προεπεξεργασία. Στη συνέχεια εφαρμόζουμε στα δεδομένα κάθε ασθενή ένα κυλιόμενο παράθυρο μήκους  $l$  και διασκελισμό 1, δημιουργώντας ένα σύνολο από  $N = L - l - 1$  υποακολουθίες. Αυτό μας δίνει ως αποτέλεσμα έναν τένσορα (tensor) μεγέθους  $M_{N \times l \times d}$ .

Όπως αναφέραμε πιο πάνω, έχουμε συνολικά 20 ασθενείς που αντιμετώπισαν τουλάχιστον μία ψυχωτική υποτροπή, όμως μετά τη διαδικασία προεπεξεργασίας των δεδομένων, όπου χωρίζουμε τα δεδομένα μας σε διαδοχικές ακολουθίες των 24 ωρών, ο αριθμός των υποτροπιάζοντων ατόμων με επαρκή δεδομένα μειώνεται σε δέκα. Ο Πίνακας 3.4 περιέχει πληροφορίες σχετικά με τον όγκο των δεδομένων που συλλέγονται μετά την προεπεξεργασία για κάθε ασθενή.

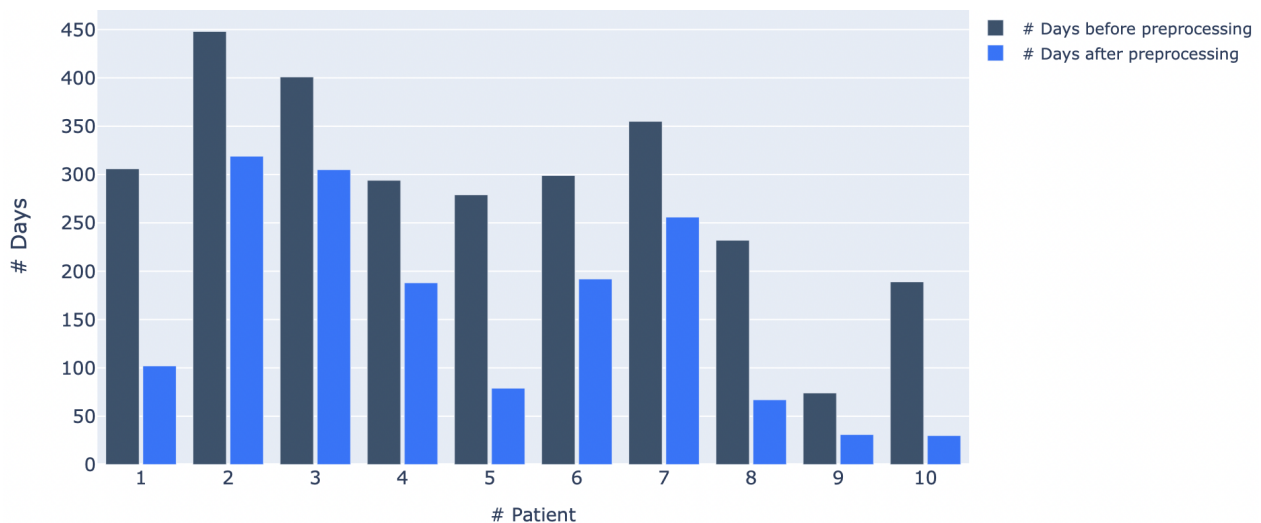
Παρατηρούμε ότι μετά τη προεπεξεργασία των δεδομένων, τα δεδομένα μειώνονται αισθητά. Αυτό συμβαίνει, αφού τα αρχικά μας δεδομένα είχαν πολλά κενά, τα οποία τις περισσότερες φορές ήταν μεγάλα, δηλαδή έλειπαν δεδομένα πάνω από 10 συνεχόμενες ώρες τα οποία δεν μπορούσαμε να τα συμπληρώσουμε διότι θα συμπληρώναμε ουσιαστικά πολλά ‘λανθασμένα’ δεδομένα.



Επίσης, βλέπουμε ότι οι 3 από τους 10 ασθενείς, οι οποίοι έχουν επισημειωθεί με έντονη γραφή στον Πίνακα 3.4, έχουν περισσότερες από μία υποτροπές, οι οποίες ανήκουν και σε διαφορετική κατηγορία σοβαρότητας. Επομένως, στη Ενότητα 5.2.3 θα μελετήσουμε αυτούς τους ασθενείς για να δούμε κατά πόσο επηρεάζει η σοβαρότητα της υποτροπής την ανίχνευση της.

Patients	Normal	Near Relapse	Relapses (Days per Relapse)			Total Relapses	Total
			Low	Moderate	Severe		
#1	90	1	–	11	–	11	102
#2	317	–	2	–	–	2	319
#3	262	13	30	–	–	30	305
#4	186	–	–	2	–	2	188
#5	<b>73</b>	-	<b>2H</b>	<b>6</b>	-	<b>6</b>	<b>79</b>
#6	<b>143</b>	<b>17</b>	<b>12H</b>	<b>32</b>	-	<b>32.5</b>	<b>192</b>
#7	<b>148</b>	<b>35</b>	<b>38</b>	<b>35</b>	-	<b>73</b>	<b>256</b>
#8	59	1	7	–	–	7	67
#9	29	–	2	–	–	2	31
#10	21	4	–	–	5	5	30
Total	1328	71	78	85	5	170	1569

Πίνακας 3.4: Πίνακας με τις μέρες στο Dataset ανά κατηγορία υποτροπής μετά τη προεπεξεργασία δεδομένων. Με έντονη γραφή επισημαίνονται οι ασθενείς οι οποίοι έχουν υποστεί περισσότερες από μία ψυχωτικές υποτροπές, με διαφορετική σοβαρότητα υποτροπής. Τα δεδομένα των ασθενών #5 και #6 από την περίοδο που είχαν υποστεί relapse χαμηλής σοβαρότητας, αποτελούνται μόνο από 2 και 12 ώρες, αντίστοιχα



Σχήμα 3.3.1: Συνολικές μέρες στο σύνολο δεδομένων για κάθε ασθενή πριν και μετά την προεπεξεργασία των δεδομένων.

Στο Σχήμα 3.3.1, παρουσιάζονται οι συνολικές μέρες στο σύνολο δεδομένων κάθε ασθενούς πριν και μετά την προεπεξεργασία δεδομένων. Με σκούρο μπλε χρώμα είναι ο αριθμός των

μερών πριν την προεπεξεργασία και με γαλάζιο χρώμα είναι ο αριθμός των μερών μετά την προεπεξεργασία. Είναι εμφανές ότι τα δεδομένα έχουν μειωθεί αρκετά μετά την προεπεξεργασία δεδομένων. Μπορούμε να χαρακτηρίσουμε κάποιους ασθενείς όπως για παράδειγμα τους %1, %5, %8 και %10 ότι δεν ήταν συνεπείς στο να φορούν το ρολόι τους. Παρ' όλα αυτά σε κάποιους από αυτούς τους ασθενείς έχουμε δει καλά αποτελέσματα, αυτό πιθανός να οφείλεται στο γεγονός ότι είχαν παρά πολλά δεδομένα και πιο σοβαρές υποτροπές σε σχέση με τους υπόλοιπους ασθενείς που δεν μας έχει επηρεάσει τόσο στο αποτέλεσμα.

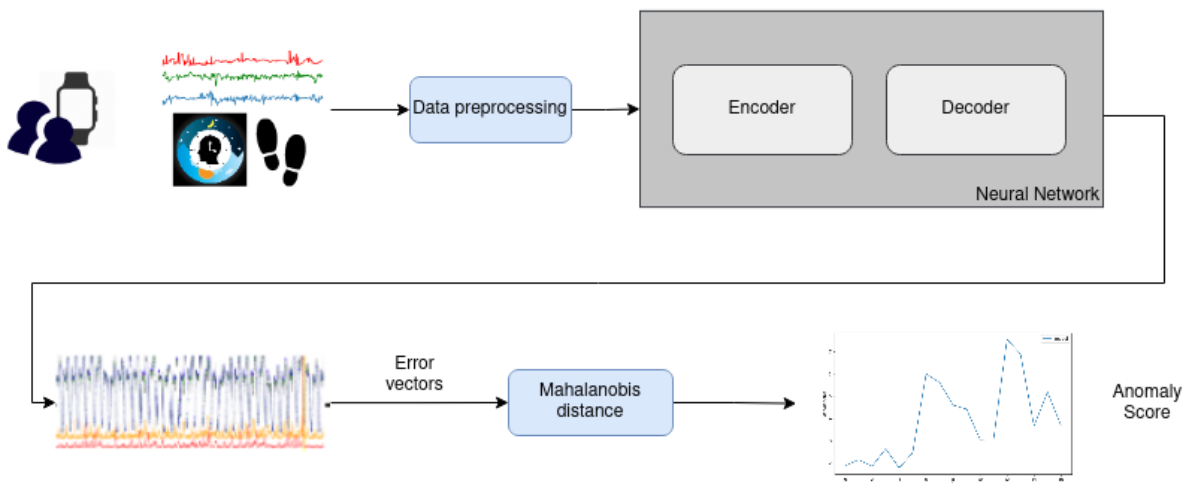
Τέλος, να σημειώσουμε ότι στην παρούσα εργασία τα δεδομένα κοντά στη υποτροπή, δηλαδή 21 μέρες πριν την υποτροπή δεν έχουν συμπεριληφθεί, αφού πιθανόν να περιέχουν δεδομένα που να θεωρηθούν ως υποτροπή λόγω του ότι είναι μια περίοδος κοντά σε αυτή.

# Κεφάλαιο 4

## Περιγραφή Μεθόδου

### 4.1 Περιγραφή Μεθόδου

Η κεντρική ιδέα αυτής της διπλωματικής εργασίας, για τον εντοπισμό υποτροπών σε άτομα με ψυχικές διαταραχές περιγράφεται από το Σχήμα 4.1.1. Αρχικά, συλλέγουμε τα δεδομένα από τους ασθενείς και τα προεπεξεργαζόμαστε όπως αναλύσαμε στο Κεφάλαιο 3. Στη συνέχεια, σχεδιάσαμε διάφορα νευρωνικά δίκτυα βασισμένα στη λογική κωδικοποιητή-αποκωδικοποιητή (encoder-decoder) με σκοπό την ανακατασκευή της ακολουθίας εισόδου, τα οποία θα αναλυθούν εκτενώς στην επομένη ενότητα. Η είσοδος των δικτύων είναι ακολουθίες πολυμεταβλητών χρονοσειρών, οι οποίες αποτελούνται από τα δέκα χαρακτηριστικά, τα οποία έχουμε προαναφέρει. Η έξοδος του νευρωνικού δικτύου είναι η ανακατασκευαζόμενη ακολουθία εισόδου. Χρησιμοποιώντας το σφάλμα ανακατασκευής, θα υπολογίσουμε την απόσταση Mahalanobis και με βάση αυτή θα γίνει η ανίχνευση της υποτροπής.



Σχήμα 4.1.1: Περιγραφή Μεθόδου

Σκοπός του νευρωνικού δικτύου είναι κατά την εκπαίδευση να μάθει τη ‘φυσιολογική’ συμπεριφορά για κάθε ασθενή, έτσι ώστε όταν αυτή η συμπεριφορά δεν είναι ‘φυσιολογική’ να μην

μπορεί να την ανακατασκευάσει πλήρως. Για να μάθει το νευρωνικό δίκτυο αυτή τη ‘φυσιολογική’ συμπεριφορά εκπαιδεύεται με normal δεδομένα, δηλαδή δεδομένα όπου ο ασθενής δεν είχε κάποια υποτροπή ούτε ήταν περίοδος κοντά σε υποτροπή. Επομένως, όταν δοθεί στο νευρωνικό μας δίκτυο είσοδος με δεδομένα που περιέχουν υποτροπές, δεν θα μπορεί να τα ανακατασκευάσει πλήρως με αποτέλεσμα το σφάλμα ανακατασκευής στα σημεία των υποτροπών να είναι μεγαλύτερο από τα φυσιολογικά. Δηλαδή, βασιζόμενη στο σφάλμα ανακατασκευής θα υπολογίσουμε την απόσταση mahalanobis (anomaly score) για κάθε σημείο και θα τη χρησιμοποιήσουμε για τον εντοπισμό των υποτροπών. Όταν το anomaly score είναι μεγαλύτερο από κάποιο κατώφλι θα θεωρείται το σημείο ως ανώμαλο, αλλιώς θα θεωρείται ως ομαλό.

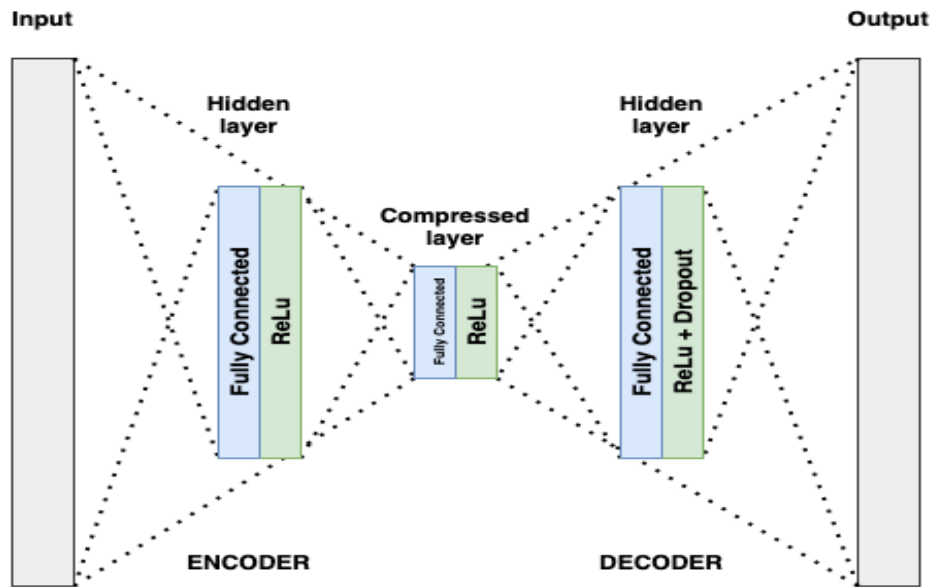
Τα μοντέλα μας θα εξεταστούν εξατομικευμένα για κάθε ασθενή άλλα και μη εξατομικευμένα (global), δηλαδή στα συνενωμένα δεδομένα όλων των ασθενών μαζί. Στα εξατομικευμένα μοντέλα, το μοντέλο θα εκπαιδευτεί και θα εξεταστεί με τα δεδομένα κάθε ασθενή ξεχωριστά. Στα global μοντέλα, το κάθε μοντέλο θα εκπαιδευτεί με όλα τα δεδομένα των ασθενών μαζί και θα εξεταστεί, αντίστοιχα. Τέλος, θα εξετάσουμε τον κάθε ασθενή εξατομικευμένα με βάση τα εκπαιδευμένα global μοντέλων.

## 4.2 Αρχιτεκτονικές Μοντέλων

### 4.2.1 Fully Connected Neural Network Autoencoder (FNN AE)

Η πρώτη αρχιτεκτονική που υλοποιήθηκε είναι εμπνευσμένη από το άρθρο του Adler et al., [Adl+20], όπου χρησιμοποιήθηκαν παθητικά δεδομένα από smartphones και είχαν υποσχόμενα αποτελέσματα για την πρόβλεψη ανωμαλιών. Συγκεκριμένα σε αυτή την εργασία, δημιουργήσαμε ένα πλήρως συνδεδεμένο μοντέλο autoencoder (FNN AE), Σχήμα 4.2.1 το οποίο αποτελείται από 2 πλήρως συνδεδεμένα στρώματα για τον κωδικοποιητή και αποκωδικοποιητή, αντίστοιχα. Ο κωδικοποιητής συμπιέζει την ακολουθία εισόδου σε χαμηλότερη διάσταση και έπειτα την εισάγουμε ως είσοδο στον αποκωδικοποιητή με στόχο να την αναδημιουργήσει.

Η είσοδος του νευρωνικού είναι το διάνυσμα  $\mathbf{X}_{l \times d}$ , δηλαδή όλα τα χαρακτηριστικά για  $l$  μήκος ακολουθίας, σε κάθε batch. Στο επόμενο επίπεδο, έχουμε το hidden layer του encoder που αποτελείται από  $h$  hidden layers. Στο επίπεδο συμπίεσης, τα hidden layers από  $h$  μειώνονται στα  $h/2$  έτσι ώστε η ακολουθία μας να συμπιεστεί σε μια χαμηλότερη διάσταση. Στο επίπεδο του decoder τα κρυφά επίπεδα είναι πάλι  $h$  για να γίνει η αναδημιουργία της ακολουθίας. Τέλος, στην έξοδο αναμένουμε την ανακατασκευασμένη ακολουθία εισόδου.

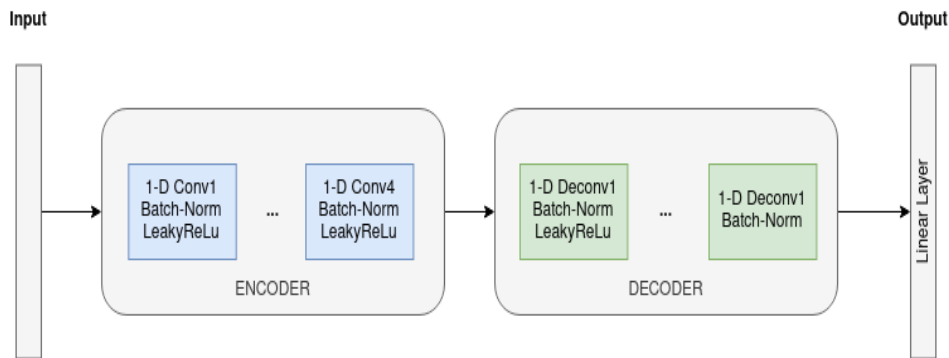


Σχήμα 4.2.1: Αρχιτεκτονική Fully Connected Neural Network Autoencoder

#### 4.2.2 Convolutional Neural Network Autoencoder (CNN AE)

Αρχιτεκτονικές βασισμένες σε 1D συνελίξεις, έχουν χρησιμοποιηθεί προηγουμένως σε σήματα ΗΚΓ [YSA18], στην ανίχνευση αρρυθμίας [Yil+18] ή ακόμη και στην πρόβλεψη της κατάστασης ύπνου [Tsi+16]. Στη δική μας μελέτη έχουμε υλοποίηση ένα 1D Convolutional Autoencoder, ο οποίος αποτελείται από δύο μέρη τον encoder και decoder όπως παρουσιάζεται στο Σχήμα 4.2.2. Ο encoder συμπιέζει τη ακολουθία εισόδου υψηλής διάστασης σε μια αναπαράσταση χαμηλότερης διάστασης και ο decoder προσπαθεί να ανακατασκευάσει την ακολουθία εισόδου από την έξοδο του encoder.

Η είσοδος του CNN AE είναι το διάνυσμα  $M_{N \times l \times d}$  που παράγεται όπως περιγράφουμε στην Ενότητα 3.3. Ο encoder και ο decoder αποτελούνται συνολικά από 4 συνελικτικά στρώματα ο καθένας. Κάθε συνελικτικό στρώμα (convolutional layer) ακολουθείται από ένα Batch Normalization layer, μια μη γραμμικότητα ReLu, εκτός από το τελευταίο layer του decoder που ακολουθείται μόνο από το Batch Normalization layer. Όλα τα convolutional layers, εκτός από το πρώτο layer του encoder και το τελευταίο layer του decoder που έχουν μήκος πυρήνα (kernel size) 10, έχουν έναν πυρήνα μήκους 3. Επίσης, στο τρίτο μπλοκ του decoder έχει προστεθεί ένα dropout layer για την αποφυγή του over-fitting. Στο τέλος του δικτύου εφαρμόζεται ένα πλήρως συνδεδεμένο στρώμα (fully connected layer).



Σχήμα 4.2.2: Αρχιτεκτονική Convolutional Neural Network Autoencoder

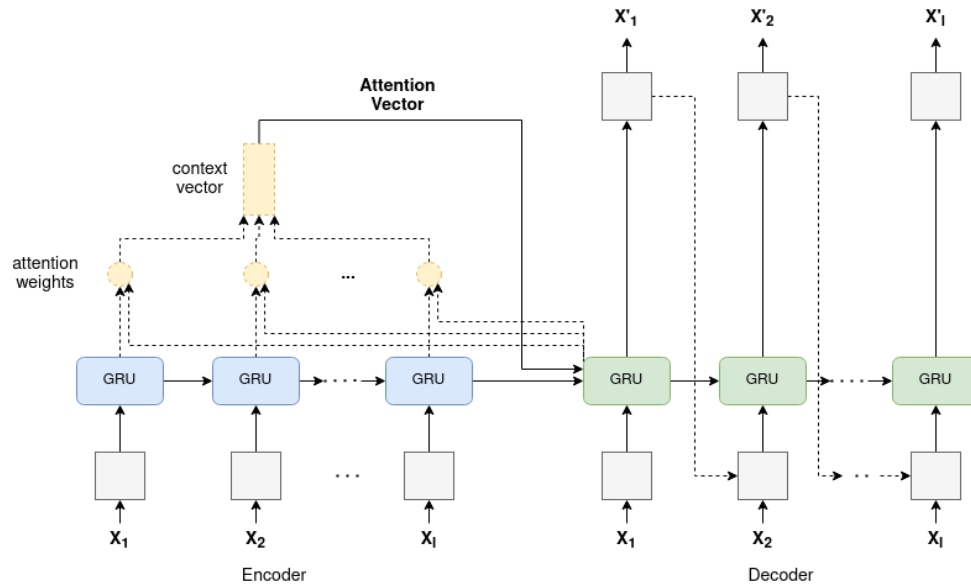
### 4.2.3 Recurrent Neural Network (RNN)

#### GRU Seq2Seq Autoencoder

Μετά από έρευνα, επιλέξαμε να σχεδιάσουμε ένα μοντέλο GRU Seq2Seq με Attention μηχανισμό Autoencoder (GRU AE). Παρόμοια αρχιτεκτονική έχει υλοποιηθεί στο άρθρο του Adler et al., [Adl+20]. Επιλέξαμε να χρησιμοποιήσουμε τη GRU αφού αντιμετωπίζει επιτυχώς το πρόβλημα του vanishing gradient και συγκλίνει γρηγορότερα σε σύγκριση με άλλες δημοφιλείς RNN αρχιτεκτονικές όπως τα LSTMs [Chu+14]. Επίσης, αντίθετα με τη περίπτωση που απλού RNN Autoencoder που χρησιμοποιεί τη τελική έξοδο από τον encoder ως είσοδο του decoder, επιλέξαμε να χρησιμοποιήσουμε έναν σταθμισμένο μέσο όρο όλων των κωδικοποιημένων εξόδων από όλα τα χρονικά βήματα. Ο μηχανισμός attention, μας επιτρέπει να βρούμε το βέλτιστο βάρος κάθε εξόδου του encoder για τον υπολογισμό των εισόδων του decoder σε ένα δεδομένο χρονικό βήμα. Αυτά τα βάρη ονομάζονται attention weights και ορίζονται για κάθε χρονικό βήμα. Το διάνυσμα προσοχής (attention vector) υπολογίζεται χρησιμοποιώντας την τελευταία κρυφή κατάσταση του encoder. Στη συνέχεια, τα attention weights πολλαπλασιάζονται με τις εξόδους του encoder. Η είσοδος του encoder είναι η ακολουθία  $M_{N \times l \times d}$ . Οι GRU του encoder και του decoder αποτελούνται από ένα επαναλαμβανόμενο στρώμα, 100 χαρακτηριστικών στη κρυφή κατάσταση και dropout ίσο με 0.1.

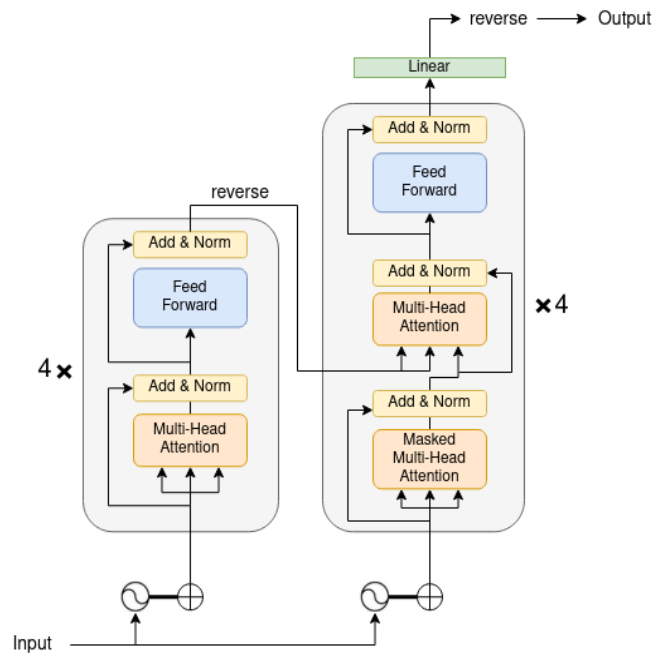
#### 4.2.4 Transformer

Η αρχιτεκτονική του Transformer που υλοποιήσαμε είναι όμοια με αυτή που περιγράψαμε στη θεωρία της Ενότητας 2.4.4. Στο μοντέλο Transformer, αρχικά εισάγεται η ακολουθία εισόδου  $M_{N \times l \times d}$  στο position encoding layer, όπου διατηρούνται πληροφορίες σχετικά με τη χρονολογική τους σειρά μέσω της εφαρμογής των ημιτονοειδών (sine) και συνημιτονοειδών (cosine) συναρτήσεων. Ο encoder μας αποτελείται από τέσσερα layers (encoder layers). Κάθε encoder layer αποτελείται από δύο sub-layers, συμπεριλαμβανομένου ενός head self attention layer και ενός πλήρως συνδεδεμένου δικτύου τροφοδοσίας με διάσταση 1024. Και τα δύο sub-layers ακολουθούνται από ένα normalization layer. Η έξοδος του encoder, αντιστρέφεται και εισάγεται στον decoder που αποτελείται από τέσσερα layers. Τα layers του decoder έχουν παρόμοια αρχιτεκτονική και τις ίδιες παραμέτρους με του encoder, με μόνη εξαίρεση ότι στο τέλος κάθε στρώματος decoder υπάρχει ένα άλλο sub-layer για το one-head attention στην έξοδο



Σχήμα 4.2.3: Αρχιτεκτονική GRU Seq2Seq Autoencoder with Attention

του decoder. Τέλος, προσθέτουμε ένα linear layer στην έξοδο του decoder, όπως φαίνεται στο Σχήμα 4.2.4.



Σχήμα 4.2.4: Αρχιτεκτονική Transformer



### 4.3 Εκπαίδευση

Στην ανάλυσή μας, διαχωρίζουμε τα Normal δεδομένα σε τρία σύνολα, δηλαδή σε train, validation και test set, με διαχωρισμό 60 – 20 – 20(%), αντίστοιχα. Κανονικοποιούμε τα δεδομένα εκπαίδευσης στην περιοχή  $[0, 1]$  και οι τιμές στα υπόλοιπα σύνολα μετασχηματίζονται κατάλληλα, εκτός από το χαρακτηριστικά χρόνου σε όλα τα σύνολα που είναι ήδη στο  $[-1, 1]$ .

Ονομάζοντας αυτή τη διαδικασία ως ‘fold’, την επαναλαμβάνουμε για 5 φορές και συλλέγουμε τα αποτελέσματα απόδοσης για κάθε fold, προσπαθώντας μία πιο robust επικύρωση της απόδοσης του μοντέλου. Αυτή η ροή επικύρωσης έχει δημιουργηθεί για κάθε ασθενή ξεχωριστά και για όλους τους ασθενείς μαζί.

Τα δεδομένα εισόδου του νευρωνικού δικτύου εξάγονται όπως περιγράφηκαν στην Ενότητα 3.3. Δηλαδή, είναι ακολουθίες  $X = \{x_1, x_2, \dots, x_L\}$  μήκους  $L$ , όπου  $L = 288$  στη περίπτωση των 5 λεπτών,  $L = 144$  στη περίπτωση των 10 λεπτών,  $L = 24$  στη περίπτωση των 1 ώρας και  $x = [F_1, F_2, \dots, F_i]$  το διάλυμα με  $i$  το πλήθος των χαρακτηριστικών και  $F_i$  : το χαρακτηριστικό  $i$ . Στη περίπτωση των raw δεδομένων το  $i = 4$  και στη περίπτωση ολόκληρου του συνόλου δεδομένων το  $i = 10$ .

Η δημιουργία εξατομικευμένων μοντέλων για ψυχωτικές ασθένειες αποτελεί μια λογική αφετηρία, όπως προτείνεται επίσης στο [Bar+18], αφού υπάρχει πολυπλοκότητα και η ετερογένεια σε όλες τις ψυχωτικές ασθένειες. Για τη δημιουργία των εξατομικευμένων μοντέλων εκπαιδύουμε και αξιολογούμε τα μοντέλα μας για τα δεδομένα κάθε ασθενή ξεχωριστά. Δηλαδή, στην περίπτωση μας θα εκπαιδεύσουμε 10 μοντέλα για κάθε αρχιτεκτονική. Στη συνέχεια, δημιουργήσαμε ένα γενικό (global) μοντέλο, δηλαδή εκπαιδύουμε και αξιολογούμε το μοντέλο μας για τα συνενωμένα δεδομένα εκπαίδευσης, επικύρωσης και αξιολόγησης όλων των ασθενών μαζί. Τέλος, έχοντας εκπαιδεύσει τα global μοντέλα μας (για τις διάφορες αρχιτεκτονικές), δηλαδή με όλα τα δεδομένα μαζί, τα αξιολογούμε για κάθε ασθενή ξεχωριστά. Τα αποτελέσματα για αυτά τα μοντέλα παρουσιάζονται στη Ενότητα 5.

Το framework που έχουν εφαρμοστεί τα βαθιά νευρωνικά δίκτυα είναι το Pytorch και οι παράμετροι της εκπαίδευσης των μοντέλων GRU Seq2Seq Autoencoder με Attention μηχανισμό και Transformer παρουσιάζονται στον Πίνακα 4.1 και για τα μοντέλα Fully Connected Autoencoder και Convolution Autoencoder παρουσιάζονται στον Πίνακα 4.2.

Οι παράμετροι για την εκπαίδευση των μοντέλων αποφασίστηκαν μετά από ενδελεχή πειραματισμό. Στην εκπαίδευση των μοντέλων γίνεται χρήση Early Stopping και dropout layer για την αποφυγή του over-fitting, το οποίο παρατηρήθηκε χωρίς την χρήση αυτών.

### 4.4 Υπολογισμός Anomaly score

Για τη διαδικασία ανίχνευσης ανωμαλιών ακολουθούμε παρόμοια προσέγγιση με το [Mal+15]. Αρχικά εκπαιδύουμε ένα μοντέλο χωρίς δεδομένα που περιέχουν υποτροπές (δηλαδή, ‘normal’ δεδομένα) και στη συνέχεια το αξιολογούμε σε ‘unseen’ δεδομένα που περιέχουν normal δεδομένα και δεδομένα με υποτροπές. Διαισθητικά, αναμένουμε να παρατηρήσουμε μικρά σφάλματα ανακατασκευής σε ‘unseen’ κανονικά δεδομένα και μεγαλύτερα σφάλματα στα δεδομένα υποτροπής.

Parameters	GRU AE
Hidden Units	100
Stacked Layers	1
Batch size	64
Activation Function	ReLU
Optimizer	<i>RMSprop</i>
Loss Function	MSE
Learning Rate	0.0001
Epochs	50
Dropout RNN	0.1
Early stopping (patience)	5

(a) Παράμετροι για το μοντέλο GRU Seq2Seq Autoencoder με Attention μηχανισμό.

Parameters	Transformer
Encoder Layers	4
Decoder Layers	4
Hidden Nodes	1024
Number of heads	1
Batch size	64
Optimizer	Adam
Loss Function	MSE
Learning Rate	0.001
Epochs	50
Dropout	0.6
Early stopping (patience)	1

(b) Παράμετροι για το μοντέλο Transformer.

Πίνακας 4.1: Οι παράμετροι για τα μοντέλα GRU Seq2Seq Autoencoder με Attention μηχανισμό και Transformer περιγράφονται από τους υποπίνακες **a** και **b**.

Parameters	FNN AE
Hidden Units	700
Batch size	64
Optimizer	Adam
Loss Function	MSE
Learning Rate	0.0001
Epochs	50
Dropout	0.1
Early stopping (patience)	2

(a) Παράμετροι για το μοντέλο Fully Connected Autoencoder.

Parameters	CNN AE
Batch size	64
Optimizer	RMSprop
Loss Function	MSE
Learning Rate	0.0001
Epochs	50
Dropout	0.1
Early stopping (patience)	5

(b) Παράμετροι για το μοντέλο Convolutional Autoencoder.

Πίνακας 4.2: Οι παράμετροι για τα μοντέλα Fully Connected Autoencoder και Convolutional Autoencoder περιγράφονται από τους υποπίνακες **a** και **b**.

Αρχικά, υπολογίζουμε στο validation set τα διανύσματα σφάλματος (error vectors) για κάθε χαρακτηριστικό ξεχωριστά σε όλα τα σημεία. Δηλαδή, τα διανύσματα σφάλματος  $e$  θα έχουν διαστάσεις ( $\#Points, \#Features$ ). Στη συνέχεια, χρησιμοποιούμε τα διανύσματα σφάλματος για να υπολογίσουμε το mean και το covariance, μιας πολυμεταβλητής κανονικής κατανομής  $N(\mu, \Sigma)$ . Έπειτα, υπολογίζουμε την απόσταση Mahalanobis που αναφέρεται ως ‘anomaly score’, μεταξύ των προβλεπόμενων σημείων στο test set και της Γκαουσιανής κατανομής που είχαμε υπολογίσει στο validation set. Επομένως, για κάθε σημείο  $x^{(i)}$  το ‘anomaly score’ υπολογίζεται ως [Mah36]:

$$a^{(i)} = \sqrt{(e^{(i)} - \mu)^T \Sigma^{-1} (e^{(i)} - \mu)} \quad (4.4.1)$$

Στην τρέχουσα μελέτη, τα anomaly scores υπολογίζονται με βάση τον μέσο όρο των anomaly scores των σημείων για μια ημέρα, παρόμοια με το [Adl+20] και η αξιολόγηση απόδοσης γίνεται σε αυτό το επίπεδο.

## 4.5 Ανίχνευση Ανωμαλιών

Η ανίχνευση μιας ανωμαλίας αποφασίζεται από ένα threshold που αποκτάται από ένα υποσύνολο των ‘unseen’ normal δεδομένων και των δεδομένων υποτροπής. Σύμφωνα με σχετική βιβλιογραφία [Mal+16; Adl+20], η μεγιστοποίηση μιας μέτρησης απόδοσης, όπως η βαθμολογία F-beta ή η αναλογία αληθινών θετικών και ψευδώς θετικών ποσοστών, δηλαδή TPR/FPR, μπορεί να χρησιμοποιηθεί για τον εντοπισμό κάποιου threshold.

Στη δική μας μελέτη, δεν εφαρμόζουμε κάποιο συγκεκριμένο threshold αφού χρησιμοποιούμε τις μετρικές PR και ROC AUC, οι οποίες για κάθε κατώφλι υπολογίζουν τις καμπύλες PR και ROC αντίστοιχα και στη συνέχεια παίρνουμε το AUC score για να δούμε πόσο αποδοτική είναι η μέθοδος που δημιουργήσαμε. Οι μετρικές PR και ROC AUC, περιγράφονται στην Ενότητα 5.1.

# Κεφάλαιο 5

## Πειράματα και Αξιολόγηση

### 5.1 Μετρικές Αξιολόγησης

Για την αξιολόγηση των αποτελεσμάτων χρησιμοποιούνται οι μετρικές ROC (Receiver Operating Characteristics Area Under Curve) και PR AUC (Precision-Recall Area Under Curve). Για να ορίσουμε τις πιο πάνω μετρικές πρέπει πρώτα να ορίσουμε τα ακόλουθα.

**True Positive Rate (TPR) ή Sensitivity ή Recall:** Το TPR μας δείχνει ποιο ποσοστό της θετικής κλάσης ταξινομήθηκε σωστά. Στην περίπτωση μας, θα προσδιορίσει το ποσοστό των υποτροπών που εντοπίστηκαν σωστά από το μοντέλο:

$$TPR = \frac{TP}{TP + FN} \quad (5.1.1)$$

**False Positive Rate (FPR):** Το FPR μας δείχνει ποιο ποσοστό της αρνητικής κλάσης ταξινομήθηκε λανθασμένα από το μοντέλο. Στην περίπτωση μας θα προσδιορίσει ποιο ποσοστό μη υποτροπών ταξινομήθηκε λανθασμένα ως υποτροπή:

$$FPR = \frac{FP}{TN + FP} \quad (5.1.2)$$

**Precision:** Η ακρίβεια (Precision) μας δείχνει πόσα σωστά έχει προβλέψει το μοντέλο μας σε σχέση με τον αριθμό των σωστών και των λανθασμένων προβλέψεων:

$$Precision = \frac{TP}{TP + FN} \quad (5.1.3)$$

όπου,

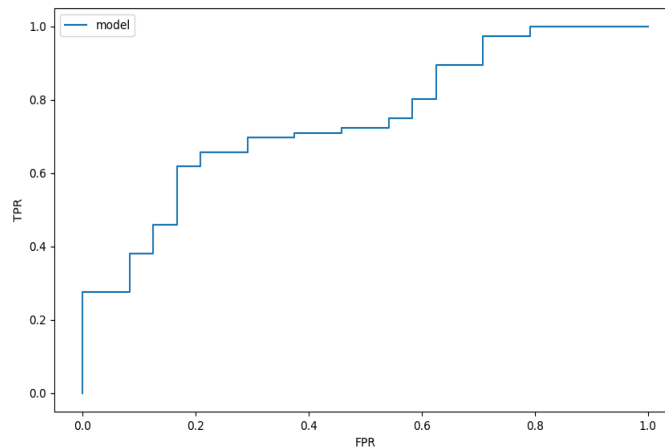
- True Positive (TP): Θεωρούμε την πρόβλεψη μας ότι είναι TP όταν το μοντέλο μας εντοπίσει σωστά την υποτροπή.
- True Negatives (TN): Θεωρούμε την πρόβλεψη μας ότι είναι TN όταν το μοντέλο μας εντοπίσει σωστά τη μη υποτροπή.

- False Positive (FP): Θεωρούμε την πρόβλεψη μας ότι είναι FP όταν το μοντέλο μας εντοπίσει ότι έχουμε υποτροπή ενώ στην πραγματικότητα δεν είχαμε.
- False Negative (FN): Θεωρούμε την πρόβλεψη μας ότι είναι FN όταν το μοντέλο μας εντοπίσει ότι δεν έχουμε υποτροπή ενώ στην πραγματικότητα είχαμε.

Επομένως, τώρα μπορούμε να ορίσουμε τα ROC, ROC AUC, PR και PR AUC.

**ROC:** Η καμπύλη ROC είναι ένα γράφημα που απεικονίζει την αντιστάθμιση μεταξύ TPR και του FPR. Δηλαδή, απεικονίζει το σύνολο των δυνατών ζευγών FPR και TPR που μπορεί να ληφθούν, ανάλογα με την επιλογή του σημείου απόφασης [Faw06]. Συγκεκριμένα για κάθε κατώφλι (threshold), υπολογίζουμε το TPR και το FPR και το σχεδιάζουμε σε ένα γράφημα. Όσο υψηλότερο είναι το TPR και χαμηλότερο το FPR για κάθε threshold τόσο καλύτερος είναι και ο ταξινομητής.

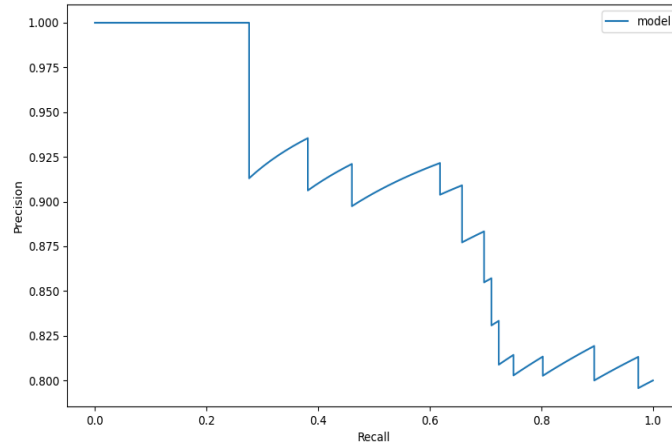
**ROC AUC:** ROC AUC είναι ο αριθμός που μας λέει πόσο καλή είναι η καμπύλη μας υπολογίζοντας την περιοχή κάτω από την καμπύλη ROC, δηλαδή το εμβαδόν κάτω από την καμπύλη. Όσο πιο πάνω αριστερά είναι η καμπύλη ROC τόσο μεγαλύτερη είναι η περιοχή και επομένως υψηλότερη βαθμολογία ROC AUC.



Σχήμα 5.1.1: Παράδειγμα ROC curve από τον ασθενή #6 του FNN AE μοντέλου, με ROC AUC Score 0.75. Παρατηρούμε ότι όσο μεταβάλλονται προοδευτικά οι πιθανές οριακές τιμές διαχωρισμού, τόσο μεγαλύτερος είναι ο λόγος του TPR προς το FPR, κάτι το οποίο είναι αναμενόμενο, αφού όσο μεγαλύτερο είναι το threshold τόσο μεγαλύτερο είναι και το ποσοστό των σωστά ταξινομημένων σημείων.

**PR:** Η καμπύλη PR είναι μια καμπύλη που συνδυάζει το precision και το recall. Για κάθε threshold, υπολογίζεται το precision και το recall και σχηματίζεται η καμπύλη.

**PR AUC:** PR AUC είναι ο αριθμός που μας λέει πόσο καλή είναι η καμπύλη μας υπολογίζοντας την περιοχή κάτω από την καμπύλη PR.



Σχήμα 5.1.2: Παράδειγμα Precision-Recall curve από το ασθενή #6 του FNN AE μοντέλου, με PR AUC Score 0.91. Παρατηρούμε ότι όσο αυξάνεται το recall, τόσο μειώνεται το precision και αντίστροφα. Με την καμπύλη αυτή, μπορούμε να οπτικοποιήσουμε τον τρόπο με τον οποίο η επιλογή του threshold επηρεάζει την απόδοση των μοντέλων μας και επίσης στο να επιλέξουμε το καλύτερο threshold. Στο συγκεκριμένο σχήμα παρατηρούμε ότι το FNN AE μοντέλο φτάνει σε Recall 0.8 χωρίς να προβλέψει ψευδής υποτροπές.

Επιπρόσθετες χρήσιμες μετρικές είναι οι ακόλουθες:

**F-score:** Το f-score είναι η μετρική που μετρά την ακρίβεια. Υπολογίζεται από την ακόλουθη φόρμουλα:

$$f_{beta} = (1 + \beta^2) \frac{Precision \cdot Recall}{(\beta^2 \cdot Precision) + Recall} \quad (5.1.4)$$

**Specificity (TNR):** Το TNR είναι η αναλογία των δειγμάτων που είναι αρνητικά χρησιμοποιώντας το εν λόγω τεστ που είναι πραγματικά αρνητικά. Δηλαδή, στη περίπτωση μας μετρά την ικανότητα του μοντέλου μας να παράγει σωστά το αποτέλεσμα όταν οι ασθενείς μας δεν είχαν κάποια υποτροπή. Διαισθητικά όταν έχουμε υψηλό specificity αναμένουμε ότι δεν θα παράγονται πολλά ψευδής θετικά αποτελέσματα (δηλ. λανθασμένος εντοπισμός υποτροπής) και θα εντοπίσει σωστά τις μέρες χωρίς υποτροπή. Εκφράζεται από τον παρακάτω τύπο:

$$Specificity = \frac{TN}{TN + FP} \quad (5.1.5)$$

**Inter-quantile range (IQR):** Το IQR ορίζεται ως η διαφορά μεταξύ του 75ου και του 25ου εκατοστημρίου των δεδομένων. Δηλαδή, χωρίζουμε το σύνολο δεδομένων σε τεταρτημόρια (qaurtiles) ή τέσσερα άρτια μέρη κατά σειρά κατάταξης. Αυτά τα τεταρτημόρια συμβολίζονται με  $Q_1$  (κάτω τεταρτημόριο),  $Q_2$  (διάμεσος) και  $Q_3$  (άνω τεταρτημόριο). Το  $Q_1$  αντιστοιχεί στο 25ο εκατοστημόριο και το  $Q_3$  αντιστοιχεί στο 75ο εκατοστημόριο, άρα το IQR υπολογίζεται ως εξής [Dek+05]:

$$IQR = Q_3 - Q_1 \quad (5.1.6)$$

## 5.2 Πειράματα και Αξιολόγηση

Στα πειράματα που ακολουθούν η αξιολόγηση έγινε χρησιμοποιώντας τις μετρικές PR και ROC AUC. Σε κάθε πίνακα παρουσιάζεται η μέση τιμή (median) και το IQR της μετρικής, για τα 4 μοντέλα, αφού τα αποτελέσματα συλλέχθηκαν από 5-folds.

### 5.2.1 Εξατομικευμένα Μοντέλα

Στο πείραμα αυτό εκπαιδεύτηκαν εξατομικευμένα μοντέλα για κάθε ασθενή. Δηλαδή, για κάθε ασθενή το μοντέλο εκπαιδεύεται και αξιολογείται με τα δεδομένα του συγκεκριμένου ασθενή, αντίστοιχα. Να σημειώσουμε ότι τα μοντέλα μας σε όλα τα πειράματα έχουν τις ίδιες υπερ-παραμέτρους, όπως περιγράφηκαν στην Ενότητα 4.2.

Αρχικά ελέγξαμε τα εξατομικευμένα μοντέλα μας, με είσοδο 10 χαρακτηριστικών για 3 διαφορετικά granularities (5-min, 10-min και 1-H). Τα δεδομένα μας χωρίστηκαν σε διαστήματα πεντάλεπτων, δεκάλεπτων και μιας ώρας. Στα αποτελέσματα για τα διαστήματα των πεντάλεπτων αναφέρονται οι Πίνακες 5.1 και 5.2, των δεκάλεπτων αναφέρονται οι Πίνακες 5.3 και 5.4 και για μιας ώρας αναφέρονται οι Πίνακες 5.5 και 5.6.

Στους Πίνακες 5.1 και 5.2, παρατηρούμε ότι το μοντέλο με την καλύτερη απόδοση είναι ο CNN AE, αφού έχει την ψηλότερη μέση τιμή συνολικά και για τους δέκα ασθενείς τόσο στο PR AUC όσο και στο ROC AUC. Με την μετρική PR AUC βλέπουμε ότι το CNN AE είχε τα καλύτερα αποτελέσματα σε τέσσερις από τους δέκα ασθενείς σε σχέση με τα άλλα μοντέλα. Ενώ, με την μετρική ROC AUC παρατηρούμε ότι το FNN AE είχε τα καλύτερα αποτελέσματα σε πέντε ασθενείς, σε αντίθεση με το CNN AE που τα καλύτερα αποτελέσματα τα είχε σε μόνο δύο ασθενείς. Η ψηλότερη απόδοση παρουσιάζεται στον ασθενή #1 με AUC 0.97 στο μοντέλο Transformer ενώ η χαμηλότερη απόδοση γενικότερα παρουσιάζεται στον ασθενή #2. Τα δεδομένα του ασθενή #1 αποτελούνται από 90 ‘κανονικές’ μέρες και από 11 μέρες μίας moderate υποτροπής. Το γεγονός ότι η σοβαρότητα της υποτροπής του ήταν moderate, διευκολύνει το μοντέλο μας να την εντοπίσει, καθώς οι διαφορές τις συμπεριφοράς σε περίοδο υποτροπής από την περίοδο μη υποτροπής είναι πιο έντονες, όπως φαίνεται και στο Σχήμα 5.2.1. Στο Σχήμα 5.2.1 παρουσιάζονται τα anomaly scores του test set για κάθε μέρα, για τον ασθενή #1 του μοντέλου CNN AE. Το anomaly score δεξιά της διαχωριστικής γραμμής, είναι μέρες στις οποίες ο ασθενής είχε διαγνωστεί με ψυχωτική υποτροπή, ενώ το anomaly score αριστερά της διαχωριστικής γραμμής είναι μέρες όπου ο ασθενής δεν είχε κάποιο σύμπτωμα υποτροπής (‘κανονικές’ μέρες). Παρατηρούμε ότι, τα anomaly scores στις μέρες υποτροπής είναι αρκετά υψηλότερα από τις κανονικές μέρες. Να σημειώσουμε ότι τα Points στον άξονα- $x$  είναι οι μέρες, οι οποίες δεν είναι συνεχόμενες, δηλαδή τα απεικονιζόμενα anomaly scores δεν είναι συνεχή. Στους ασθενείς #6 και #7 που είχαν και αυτή μια moderate υποτροπή και μία low υποτροπή, παρατηρούμε μέτρια αποτελέσματα. Αυτό, πιθανόν να συμβαίνει λόγω του ότι το μοντέλο μας δυσκολεύεται να ανιχνεύσει την low υποτροπή, με αποτέλεσμα η τελική συνολική απόδοση του να μην είναι τόσο καλή. Τα δεδομένα του ασθενή #2 αποτελούνται συνολικά από 319 μέρες και μόνο δύο από αυτές είναι ημέρες low υποτροπής, δηλαδή έχουμε ένα πάρα πολύ ανισόρροπο σύνολο δεδομένων (imbalanced dataset), με αποτέλεσμα να οδηγούμαστε σε μια τόσο χαμηλή απόδοση. Τέλος, να σημειώσουμε ότι η τελευταία στήλη (Random) από τους Πίνακες είναι ο Random Classifier (τυχαίος ταξινομητής), για το σύνολο δεδομένων που ελέγχθηκαν στους συγκεκριμένους ασθενείς. Παρατηρούμε ότι όλα μας τα αποτελέσματα είναι μεγαλύτερα από τον



random classifier.

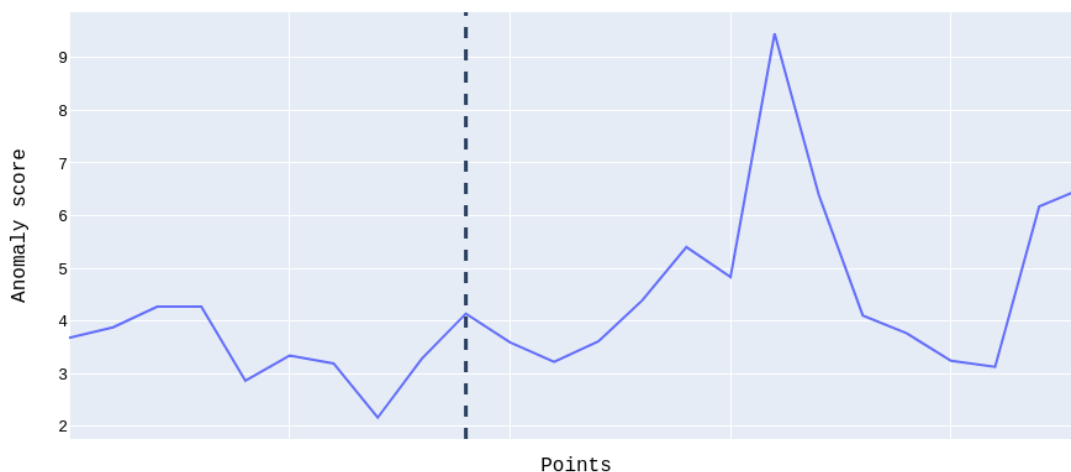
Patients	PR-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.94(0.93 – 0.97)	0.95(0.95 – 0.98)	<b>0.97(0.92 – 0.99)</b>	0.91(0.90 – 0.96)	0.91(0.90 – 0.99)
#2	<b>0.05(0.03 – 0.08)</b>	0.04(0.03 – 0.06)	0.02(0.02 – 0.02)	0.03(0.02 – 0.03)	0.03(0.03 – 0.04)
#3	<b>0.54(0.45 – 0.64)</b>	0.46(0.38 – 0.61)	0.43(0.41 – 0.59)	0.44(0.43 – 0.55)	0.53(0.52 – 0.64)
#4	0.26(0.21 – 0.36)	<b>0.34(0.28 – 0.42)</b>	0.18(0.12 – 0.25)	0.19(0.14 – 0.26)	0.18(0.14 – 0.18)
#5	<b>0.63(0.49 – 0.85)</b>	0.57(0.45 – 0.86)	0.60(0.47 – 0.72)	0.61(0.53 – 0.86)	0.63(0.50 – 0.83)
#6	0.70(0.70 – 0.78)	<b>0.72(0.67 – 0.79)</b>	0.63(0.56 – 0.76)	0.67(0.57 – 0.74)	0.68(0.60 – 0.78)
#7	0.82(0.78 – 0.82)	0.86(0.80 – 0.88)	<b>0.87(0.86 – 0.89)</b>	0.85(0.84 – 0.87)	0.86(0.84 – 0.87)
#8	0.83(0.82 – 0.87)	<b>0.87(0.80 – 0.92)</b>	0.65(0.60 – 0.84)	0.81(0.70 – 0.86)	0.85(0.81 – 0.89)
#9	0.79(0.77 – 1.00)	<b>0.80(0.54 – 1.00)</b>	0.45(0.39 – 0.77)	0.75(0.54 – 1.00)	0.68(0.48 – 1.00)
#10	0.97(0.95 – 0.97)	0.95(0.93 – 0.97)	0.94(0.81 – 0.95)	<b>0.97(0.97 – 0.97)</b>	0.97(0.95 – 1.00)
<b>Median</b>	0.75(0.59 – 0.82)	<b>0.76(0.52 – 0.87)</b>	0.61(0.44 – 0.76)	0.71(0.53 – 0.83)	0.68(0.58 – 0.85)

Πίνακας 5.1: Αποτελέσματα για **PR-AUC** σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Patients	ROC-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.94(0.91 – 0.97)	0.96(0.93 – 0.97)	<b>0.97(0.92 – 0.98)</b>	0.93(0.92 – 0.94)	0.91(0.90 – 0.97)
#2	<b>0.49(0.38 – 0.50)</b>	0.40(0.38 – 0.42)	0.22(0.18 – 0.23)	0.36(0.35 – 0.37)	0.28(0.23 – 0.30)
#3	<b>0.57(0.56 – 0.61)</b>	0.53(0.40 – 0.53)	0.49(0.48 – 0.53)	0.49(0.47 – 0.50)	0.52(0.52 – 0.54)
#4	<b>0.39(0.38 – 0.45)</b>	0.39(0.33 – 0.43)	0.35(0.28 – 0.43)	0.29(0.25 – 0.47)	0.22(0.18 – 0.24)
#5	0.44(0.30 – 0.68)	0.28(0.25 – 0.66)	<b>0.45(0.36 – 0.61)</b>	0.40(0.39 – 0.70)	0.42(0.35 – 0.68)
#6	<b>0.49(0.44 – 0.58)</b>	0.49(0.49 – 0.50)	0.39(0.35 – 0.44)	0.42(0.40 – 0.47)	0.48(0.46 – 0.48)
#7	0.56(0.56 – 0.60)	<b>0.69(0.63 – 0.74)</b>	0.69(0.66 – 0.72)	0.64(0.61 – 0.67)	0.62(0.61 – 0.62)
#8	0.72(0.64 – 0.76)	<b>0.78(0.74 – 0.81)</b>	0.64(0.47 – 0.67)	0.60(0.59 – 0.67)	0.72(0.70 – 0.74)
#9	<b>0.78(0.67 – 1.00)</b>	0.75(0.33 – 1.00)	0.28(0.25 – 0.33)	0.58(0.38 – 1.00)	0.42(0.25 – 1.00)
#10	0.91(0.88 – 0.91)	0.88(0.88 – 0.88)	0.81(0.78 – 0.84)	<b>0.94(0.88 – 0.94)</b>	0.91(0.88 – 1.00)
<b>Median</b>	0.57(0.49 – 0.75)	<b>0.61(0.44 – 0.76)</b>	0.47(0.37 – 0.67)	0.54(0.41 – 0.62)	0.50(0.42 – 0.67)

Πίνακας 5.2: Αποτελέσματα για **ROC-AUC** σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Στους Πίνακες 5.3 και 5.4 παρατηρούμε ότι το μοντέλο FNN AE έχει την καλύτερη απόδοση. Το FNN AE συγκριτικά με τα άλλα μοντέλα, στους έξι από τους δέκα ασθενείς έχει το καλύτερο αποτέλεσμα. Παρ' όλα αυτά η ψηλότερη απόδοση παρουσιάζεται στον ασθενή #1 με PR και ROC AUC 0.96 και 0.97 αντίστοιχα, στο μοντέλο GRU AE. Ο ασθενής #2 εξακολουθεί να έχει τη χαμηλότερη απόδοση. Επίσης, να σημειώσουμε ότι όλα μας τα αποτελέσματα είναι μεγαλύτερα από τον random classifier, εκτός στο PR AUC στους ασθενείς #9 και #10 και κατά συνέπεια του συνολικού median που είναι μικρότερο. Αυτό, πιθανό να οφείλεται στο γεγονός ότι τα δεδομένα σε αυτούς τους ασθενείς στο διάστημα 10 λεπτών μειώνονται αρκετά με αποτέλεσμα τα μοντέλα μας να μην μαθαίνουν επαρκώς. Γενικά, η απόδοση των μοντέλων μας για το granularity των 10 λεπτών είναι χειρότερη από αυτή των 5 λεπτών, με τα καλύτερα median PR και ROC AUC στα διαστήματα 5 λεπτών να είναι 0.76 και 0.61 αντίστοιχα, ενώ



Σχήμα 5.2.1: Γραφική παράσταση του Anomaly Score του ασθενή #1 στο μοντέλο CNN στο διάστημα 5 λεπτών. Το anomaly score δεξιά της διαχωριστικής γραμμής, είναι οι μέρες στις οποίες ο ασθενής είχε διαγνωστεί με ψυχωτική υποτροπή, ενώ το anomaly score αριστερά της διαχωριστικής γραμμής είναι οι μέρες όπου ο ασθενής δεν είχε κάποιο σύμπτωμα υποτροπής ('κανονικές' μέρες). Παρατηρούμε ότι τα anomaly scores στις μέρες υποτροπής είναι αρκετά υψηλότερα από τις κανονικές μέρες. Να σημειώσουμε ότι τα Points στον άξονα- $x$  είναι οι μέρες, οι οποίες δεν είναι συνεχόμενες, δηλαδή τα απεικονιζόμενα anomaly scores δεν είναι συνεχή.

τα καλύτερα Median PR και ROC AUC στα διαστήματα 10 λεπτών να είναι 0.63 και 0.57, αντίστοιχα. Ένας από τους λόγους που συμβαίνει αυτό είναι το γεγονός ότι έχουμε λιγότερα δεδομένα στα διαστήματα των 10 λεπτών από ότι στα των διαστήματα 5 λεπτών. Επίσης, λόγω του aggregation που γίνεται μπορεί να χάνονται σημαντικές πληροφορίες από τα δεδομένα.

Στους Πίνακες 5.5 και 5.6, παρατηρούμε ότι το μοντέλο για τη μετρική PR AUC καλύτερη απόδοση έχει ο GRU AE ενώ για τη μετρική ROC AUC καλύτερη απόδοση έχει ο Transformer. Σε γενικές γραμμές όμως δεν παρατηρούμε μεμονωμένα να ξεχωρίζει κάποιο μοντέλο αλλά παρατηρούμε χειρότερα αποτελέσματα από το granularity των 5 λεπτών, με τα καλύτερα median PR και ROC AUC στα διαστήματα 5 λεπτών να είναι 0.76 και 0.61 αντίστοιχα, ενώ τα καλύτερα Median PR και ROC AUC στα διαστήματα της 1 ώρας να είναι 0.72 και 0.54. Αυτό πιθανόν να συμβαίνει για τους ίδιους λόγους που έχουμε αναφέρει και πιο πάνω, δηλαδή είτε να χάνονται σημαντικές πληροφορίες μέσω του aggregation είτε να μειώνονται αρκετά τα δεδομένα. Επίσης, να σημειώσουμε ότι όλα μας τα αποτελέσματα είναι μεγαλύτερα από τον random classifier.

Συμπερασματικά, παρατηρούμε ότι όταν τα δεδομένα μας έχουν μικρότερο granularity τα μοντέλα μας έχουν καλύτερη απόδοση, όπως φαίνεται στα Σχήματα 5.2.2 και 5.2.3. Όπως παρατηρούμε, η απόδοση των μοντέλων FNN και CNN AE είναι αρκετά ψηλότερη όταν τα δεδομένα μας έχουν διάρκειας διαστήματος 5 λεπτών παρά όταν έχουν διάρκεια διαστήματος 10 λεπτών ή 1 ώρας. Για τα μοντέλα Transformer και GRU AE, παρατηρούμε ότι γενικά οι αποδόσεις είναι αρκετά κοντά. Επομένως, τα γενικότερα υψηλά αποτελέσματα τα έχουμε σε δεδομένα διάρκειας διαστήματος πεντάλεπτων. Αυτό μπορεί να συμβαίνει είτε επειδή τα δεδομένα

Patients	PR-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.84(0.77 – 0.93)	0.88(0.75 – 0.94)	0.90(0.77 – 0.98)	<b>0.96(0.86 – 1.00)</b>	0.82(0.78 – 0.90)
#2	<b>0.04(0.03 – 0.04)</b>	0.03(0.03 – 0.04)	0.03(0.02 – 0.03)	0.04(0.02 – 0.04)	0.03(0.02 – 0.03)
#3	<b>0.52(0.45 – 0.52)</b>	0.46(0.44 – 0.48)	0.51(0.48 – 0.52)	0.47(0.47 – 0.50)	0.46(0.44 – 0.48)
#4	<b>0.33(0.26 – 0.37)</b>	0.25(0.20 – 0.33)	0.19(0.17 – 0.24)	0.17(0.16 – 0.28)	0.17(0.16 – 0.25)
#5	<b>0.66(0.59 – 0.78)</b>	0.55(0.52 – 0.75)	0.55(0.47 – 0.64)	0.58(0.53 – 0.80)	0.63(0.59 – 0.75)
#6	<b>0.60(0.50 – 0.61)</b>	0.55(0.50 – 0.62)	0.49(0.46 – 0.60)	0.50(0.47 – 0.64)	0.54(0.50 – 0.66)
#7	0.84(0.82 – 0.88)	0.84(0.77 – 0.84)	<b>0.86(0.85 – 0.88)</b>	0.86(0.84 – 0.86)	0.84(0.82 – 0.88)
#8	<b>0.83(0.78 – 0.86)</b>	0.81(0.77 – 0.82)	0.72(0.68 – 0.80)	0.77(0.75 – 0.77)	0.76(0.70 – 0.81)
#9	0.58(0.54 – 1.00)	0.68(0.68 – 0.83)	0.68(0.68 – 0.80)	0.69(0.48 – 0.95)	0.79(0.64 – 0.95)
#10	0.93(0.92 – 0.94)	0.91(0.91 – 0.95)	0.88(0.85 – 0.92)	0.95(0.92 – 0.95)	0.97(0.96 – 0.97)
<b>Median</b>	0.63(0.55 – 0.83)	0.62(0.51 – 0.83)	0.62(0.50 – 0.79)	0.63(0.48 – 0.81)	0.70(0.50 – 0.81)

Πίνακας 5.3: Αποτελέσματα για **PR-AUC** σε δεδομένα διαστήματος 10 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Patients	ROC-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.85(0.73 – 0.88)	0.81(0.72 – 0.93)	0.93(0.86 – 0.99)	<b>0.97(0.91 – 0.99)</b>	0.90(0.87 – 0.93)
#2	<b>0.38(0.38 – 0.41)</b>	0.34(0.24 – 0.42)	0.09(0.08 – 0.14)	0.29(0.29 – 0.33)	0.15(0.14 – 0.16)
#3	<b>0.63(0.49 – 0.69)</b>	0.49(0.47 – 0.58)	0.55(0.50 – 0.56)	0.51(0.46 – 0.54)	0.49(0.49 – 0.52)
#4	0.41(0.40 – 0.44)	<b>0.42(0.33 – 0.44)</b>	0.33(0.29 – 0.42)	0.29(0.26 – 0.40)	0.22(0.17 – 0.22)
#5	<b>0.52(0.50 – 0.56)</b>	0.49(0.40 – 0.50)	0.30(0.25 – 0.48)	0.52(0.44 – 0.55)	0.50(0.41 – 0.51)
#6	0.30(0.29 – 0.53)	<b>0.43(0.31 – 0.46)</b>	0.29(0.29 – 0.40)	0.31(0.31 – 0.33)	0.43(0.36 – 0.46)
#7	<b>0.61(0.52 – 0.65)</b>	0.60(0.57 – 0.62)	0.59(0.57 – 0.63)	0.61(0.56 – 0.62)	0.59(0.57 – 0.68)
#8	<b>0.71(0.67 – 0.76)</b>	0.71(0.68 – 0.73)	0.64(0.58 – 0.65)	0.60(0.53 – 0.62)	0.67(0.61 – 0.68)
#9	0.44(0.42 – 1.00)	<b>0.63(0.33 – 0.67)</b>	0.42(0.41 – 0.42)	0.50(0.31 – 0.75)	0.58(0.41 – 0.75)
#10	<b>0.81(0.78 – 0.81)</b>	0.78(0.78 – 0.80)	0.78(0.39 – 0.78)	0.78(0.78 – 0.83)	0.83(0.83 – 0.89)
<b>Median</b>	<b>0.57(0.43 – 0.67)</b>	0.54(0.46 – 0.67)	0.48(0.31 – 0.61)	0.51(0.40 – 0.60)	0.54(0.46 – 0.63)

Πίνακας 5.4: Αποτελέσματα για **ROC-AUC** σε δεδομένα διαστήματος 10 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

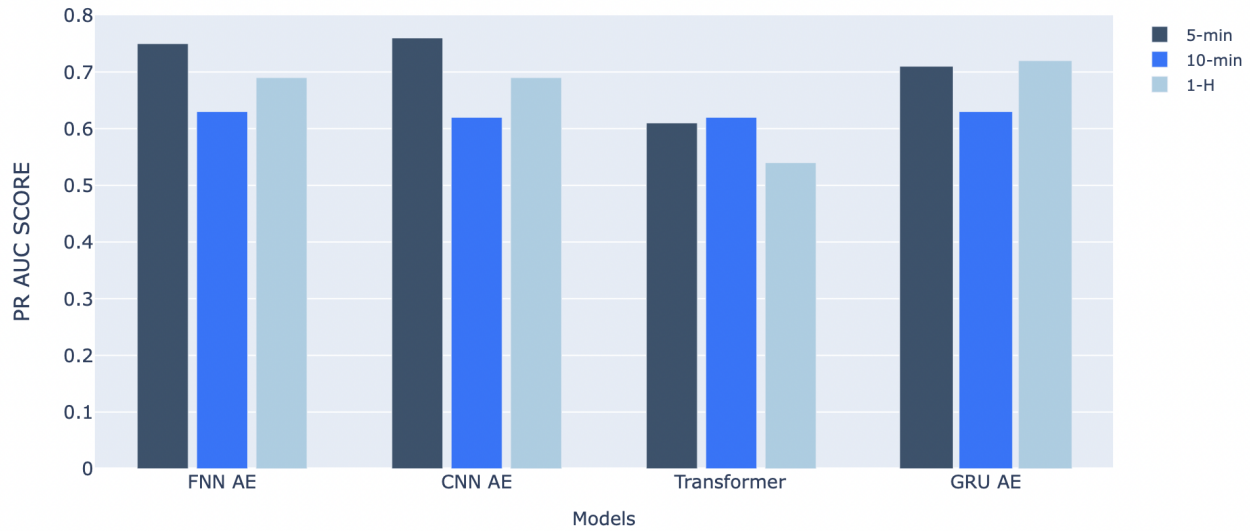
Patients	PR-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.81(0.78 – 0.83)	<b>0.89(0.88 – 0.92)</b>	0.53(0.44 – 0.60)	0.85(0.83 – 0.86)	0.73(0.73 – 0.78)
#2	0.06(0.04 – 0.06)	0.06(0.05 – 0.07)	<b>0.07(0.06 – 0.13)</b>	0.05(0.05 – 0.06)	0.06(0.06 – 0.14)
#3	0.47(0.44 – 0.48)	0.46(0.43 – 0.46)	<b>0.55(0.50 – 0.56)</b>	0.46(0.38 – 0.49)	0.46(0.42 – 0.47)
#4	<b>0.34(0.31 – 0.36)</b>	0.21(0.21 – 0.23)	0.26(0.20 – 0.33)	0.25(0.24 – 0.31)	0.22(0.21 – 0.25)
#5	0.73(0.67 – 0.86)	0.65(0.64 – 0.79)	0.68(0.66 – 0.83)	<b>0.76(0.71 – 0.81)</b>	0.69(0.62 – 0.72)
#6	0.65(0.62 – 0.78)	<b>0.81(0.64 – 0.82)</b>	0.65(0.60 – 0.72)	0.69(0.64 – 0.76)	0.69(0.64 – 0.74)
#7	0.79(0.71 – 0.83)	0.72(0.67 – 0.73)	0.75(0.68 – 0.80)	0.75(0.75 – 0.80)	0.81(0.70 – 0.83)
#8	<b>0.82(0.64 – 0.82)</b>	0.76(0.66 – 0.79)	0.52(0.50 – 0.60)	0.75(0.72 – 0.83)	0.79(0.70 – 0.82)
#9	<b>0.52(0.41 – 0.60)</b>	0.39(0.37 – 0.44)	0.46(0.42 – 0.54)	0.45(0.42 – 0.55)	0.42(0.36 – 0.51)
#10	0.93(0.91 – 0.94)	<b>0.95(0.95 – 0.96)</b>	0.86(0.86 – 0.88)	0.93(0.93 – 0.93)	0.91(0.89 – 0.91)
<b>Median</b>	0.69(0.50 – 0.80)	0.69(0.43 – 0.78)	0.54(0.49 – 0.66)	<b>0.72(0.45 – 0.75)</b>	0.69(0.44 – 0.76)

Πίνακας 5.5: Αποτελέσματα για **PR-AUC** σε δεδομένα διαστήματος μιας ώρας για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

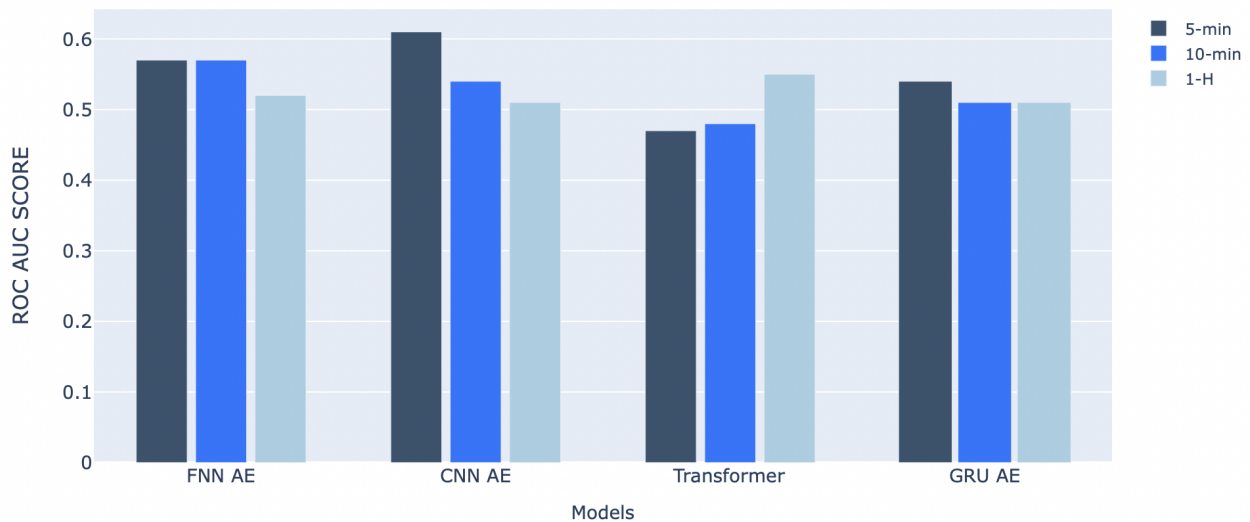
Patients	ROC-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.76(0.76 – 0.82)	<b>0.92(0.87 – 0.92)</b>	0.58(0.53 – 0.65)	0.85(0.81 – 0.91)	0.79(0.73 – 0.81)
#2	0.25(0.23 – 0.26)	0.31(0.22 – 0.45)	<b>0.44(0.41 – 0.56)</b>	0.28(0.22 – 0.29)	0.30(0.29 – 0.46)
#3	0.52(0.48 – 0.59)	0.53(0.52 – 0.57)	<b>0.59(0.56 – 0.61)</b>	0.54(0.51 – 0.55)	0.50(0.49 – 0.51)
#4	0.44(0.39 – 0.46)	0.44(0.38 – 0.47)	0.36(0.35 – 0.39)	<b>0.45(0.43 – 0.48)</b>	0.41(0.38 – 0.41)
#5	0.51(0.44 – 0.62)	0.43(0.41 – 0.65)	<b>0.58(0.50 – 0.68)</b>	0.49(0.49 – 0.52)	0.45(0.34 – 0.63)
#6	0.28(0.27 – 0.44)	<b>0.50(0.33 – 0.57)</b>	0.27(0.26 – 0.40)	0.36(0.34 – 0.43)	0.38(0.35 – 0.39)
#7	<b>0.58(0.57 – 0.58)</b>	0.51(0.40 – 0.51)	0.54(0.53 – 0.54)	0.55(0.52 – 0.56)	0.57(0.52 – 0.68)
#8	0.63(0.57 – 0.71)	0.62(0.57 – 0.67)	0.56(0.55 – 0.59)	<b>0.65(0.61 – 0.72)</b>	0.61(0.57 – 0.68)
#9	0.38(0.29 – 0.47)	0.25(0.19 – 0.31)	<b>0.46(0.28 – 0.58)</b>	0.43(0.37 – 0.53)	0.27(0.19 – 0.42)
#10	0.73(0.70 – 0.73)	<b>0.75(0.74 – 0.80)</b>	0.60(0.53 – 0.60)	0.68(0.65 – 0.73)	0.60(0.60 – 0.63)
<b>Median</b>	0.52(0.41 – 0.60)	0.51(0.43 – 0.58)	<b>0.55(0.45 – 0.58)</b>	0.51(0.44 – 0.60)	0.47(0.39 – 0.59)

Πίνακας 5.6: Αποτελέσματα για **ROC-AUC** σε δεδομένα διαστήματος μιας ώρας για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

μειώνονται καθώς αυξάνεται το granularity είτε επειδή χάνονται σημαντικές πληροφορίες μέσω του aggregation.



Σχήμα 5.2.2: Median PR AUC για τα μοντέλα FNN AE, CNN AE, Transformer και GRU AE στα διαφορετικά granularities (5 λεπτών, 10 λεπτών και 1 ώρας).



Σχήμα 5.2.3: Median ROC AUC για τα μοντέλα FNN AE, CNN AE, Transformer και GRU AE στα διαφορετικά granularities (5 λεπτών, 10 λεπτών και 1 ώρας).

Τα σφάλματα ανακατασκευής (reconstruction error) για τα μοντέλα μας στο πείραμα των 5 λεπτών, το οποίο έχει την καλύτερη απόδοση σε σχέση με τα πειράματα στα υπόλοιπα granularities παρουσιάζεται στον Πίνακα 5.7.

Patients	Normal	NR	Relapse
#1	0.000455	0.000432	<b>0.000674</b>
#2	<b>0.000364</b>	0.000285	0.000311
#3	<b>0.000411</b>	0.000347	0.000392
#4	<b>0.000403</b>	0.000264	0.000305
#5	<b>0.000570</b>	0.000338	0.000530
#6	0.000392	<b>0.000411</b>	0.000387
#7	0.000490	0.000511	<b>0.000554</b>
#8	0.000460	0.000630	<b>0.000651</b>
#9	0.0	-	<b>3.0</b>
#10	0.000740	0.001051	<b>0.001625</b>
Median	0.000433	0.000411	<b>0.000542</b>

(a) Σφάλμα Ανακατασκευής - Median Square Error στο εξατομικευμένο μοντέλο FNN AE για τις τρεις περιόδους δεδομένων (Normal, Near Relapse και Relapse).

Patients	Normal	NR	Relapse
#1	0.000611	0.000748	<b>0.001081</b>
#2	<b>0.000582</b>	0.000318	0.000507
#3	0.000489	0.000428	<b>0.000532</b>
#4	<b>0.000493</b>	0.000325	0.000369
#5	<b>0.001058</b>	0.000689	0.001003
#6	<b>0.000619</b>	0.000654	0.000591
#7	0.000850	0.000874	<b>0.000889</b>
#8	0.001145	0.001574	<b>0.001818</b>
#9	0.0	-	<b>3.0</b>
#10	0.004276	0.004414	<b>0.004547</b>
Median	0.000615	0.000689	<b>0.000946</b>

(b) Σφάλμα Ανακατασκευής - Median Square Error στο εξατομικευμένα μοντέλο CNN AE για τις τρεις περιόδους δεδομένων (Normal, Near Relapse και Relapse).

Patients	Normal	NR	Relapse
#1	0.000638	0.000807	<b>0.000954</b>
#2	<b>0.000824</b>	0.000563	0.000694
#3	0.000667	0.000683	<b>0.000682</b>
#4	<b>0.000618</b>	0.000392	0.000408
#5	0.000914	0.000790	<b>0.001062</b>
#6	0.000771	<b>0.000817</b>	0.000775
#7	0.001375	<b>0.001417</b>	0.001380
#8	0.000826	<b>0.001096</b>	0.000983
#9	0.142857	-	<b>0.333333</b>
#10	0.002481	0.002363	<b>0.002806</b>
Median	0.000825	0.000807	<b>0.000969</b>

(c) Σφάλμα Ανακατασκευής - Median Square Error στο εξατομικευμένο μοντέλο Transformer για τις τρεις περιόδους δεδομένων (Normal, Near Relapse και Relapse).

Patients	Normal	NR	Relapse
#1	0.002939	0.004076	<b>0.005821</b>
#2	<b>0.002386</b>	0.001441	0.001619
#3	0.002676	0.002424	<b>0.002733</b>
#4	<b>0.002532</b>	0.001110	0.001477
#5	0.002901	0.001683	<b>0.003176</b>
#6	0.002482	<b>0.002803</b>	0.002620
#7	0.004250	<b>0.004435</b>	0.004286
#8	0.002896	0.002912	<b>0.003670</b>
#9	0.0	-	<b>1.0</b>
#10	0.004067	0.006027	<b>0.006602</b>
Median	0.002786	0.002803	<b>0.003423</b>

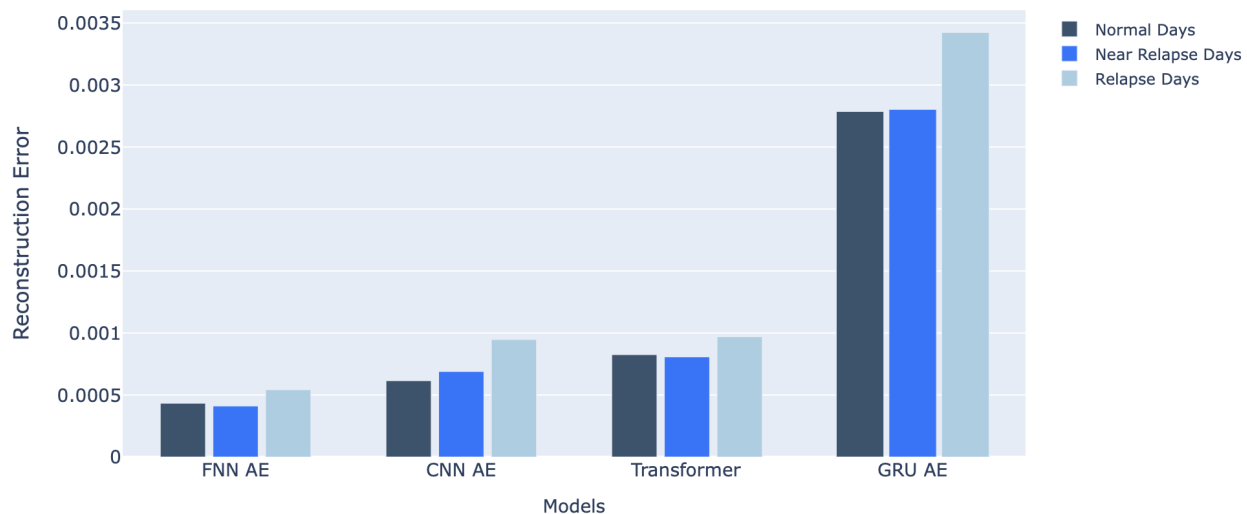
(d) Σφάλμα Ανακατασκευής - Median Square Error στο εξατομικευμένα μοντέλο GRU AE για τις τρεις περιόδους δεδομένων (Normal, Near Relapse και Relapse).

Πίνακας 5.7: Τα Σφάλμα Ανακατασκευής των εξατομικευμένων μοντέλων μας περιγράφονται από τους Υποπίνακες **a**, **b**, **c** και **d**.

Παρατηρούμε ότι στους περισσότερους ασθενείς υπάρχει μία κλιμακωτή άνοδος των σφαλμάτων ανακατασκευής ανάμεσα στη Normal, Near Relapse και Relapse περίοδο. Το αποτέλεσμα αυτό είναι αναμενόμενο, αφού τα μοντέλα μας εκπαιδεύονται για ένα κομμάτι της Normal περιόδου, επομένως είναι λογικό το μοντέλο να μάθει τις συγκεκριμένες ‘κανονικές’ συμπεριφορές και το σφάλμα ανακατασκευής σε άγνωστες συμπεριφορές να είναι μεγαλύτερο. Συγκεκριμένα,



στους ασθενείς οι οποίοι έχουν καλή απόδοση όπως ο #1, #7, #8, #9 και #10 παρατηρείται και μεγαλύτερο reconstruction error για την περίοδο υποτροπής σε σχέση με την κανονική περίοδο. Επίσης, παρατηρούμε ότι ο ασθενής #2, ο οποίος έχει και τη χαμηλότερη απόδοση, το reconstruction error του σε όλα τα μοντέλα είναι μεγαλύτερο στα κανονικά δεδομένα από ότι στα δεδομένα υποτροπής. Από εδώ φαίνεται ότι τα μοντέλα μας δεν μπορούν να ξεχωρίσουν κάποια διαφορετική συμπεριφορά στα δεδομένα υποτροπής από ότι στα κανονικά, για τον συγκεκριμένο ασθενή. Αξίζει να σημειωθεί ότι, ενώ ο ασθενής #3 στο μοντέλο FNN AE έχει μεγαλύτερο σφάλμα ανακατασκευής στα κανονικά δεδομένα από ότι στα δεδομένα υποτροπής, έχει την καλύτερη απόδοση στις μετρικές PR και ROC AUC. Αυτό μπορεί να οφείλεται στο γεγονός ότι δεν υπολογίζουμε το κατώφλι πάνω στα σφάλματα ανακατασκευής αλλά πάνω στην απόσταση mahalanobis, η οποία λαμβάνει υπόψη το mean και το covariance που υπολογίστηκε πάνω στο validation set, όπως αναλύσαμε στην Ενότητα 4.4. Στη δική μας έρευνα έχουμε αφαιρέσει τα Near Relapse δεδομένα, αφού λόγω της περίεργης συμπεριφοράς που μπορεί να έχει ο ασθενής καθώς θα υποτροπιάσει λίγο καιρό αργότερα, το μοντέλο μας πιθανόν να μπερδεύεται και να τα θεωρούσε ως υποτροπή. Αυτή η θεωρία επαληθεύεται για κάποιους ασθενείς, αφού μπορούμε να παρατηρήσουμε σε κάποιους υποπίνακες ότι τα Near Relapse δεδομένα έχουν μεγαλύτερο σφάλμα ανακατασκευής και από τα δεδομένα υποτροπής και από τα κανονικά δεδομένα. Γενικότερα, όπως φαίνεται και στο Σχήμα 5.2.4 τα median σφάλματα ανακατασκευής των μοντέλων μας, ακολουθούν μία ανοδική κατάταξη από τις normal μέρες μέχρι τις μέρες υποτροπής, όπως ήταν αναμενόμενο για τους λόγους που περιγράψαμε πιο πάνω.



Σχήμα 5.2.4: Σφάλματα Ανακατασκευής (median) για τα εξατομικευμένα μοντέλα με δεδομένα διαστήματος διάρκειας 5 λεπτών.

Λόγω των ανώτερων πειραμάτων, καταλήξαμε στο συμπέρασμα ότι το καλύτερο granularity είναι τα διαστήματα διάρκειας 5 λεπτών. Επομένως, όλα τα πειράματα που θα ακολουθήσουν θα αναλυθούν για το διάστημα διάρκειας μόνο των 5 λεπτών.

### Raw Δεδομένα

Σε αυτό το πείραμα ελέγχουμε τα εξατομικευμένα μοντέλα με είσοδο των raw δεδομένων, δηλαδή τεσσάρων χαρακτηριστικών (accelerometer, gyroscope, heart rate, heart rate variability) για



το διάστημα των πέντε λεπτών. Τα αποτελέσματα παρουσιάζονται στους Πίνακες 5.8 και 5.9.

Στους Πίνακες 5.8 και 5.9, παρατηρούμε ότι με την μετρική PR AUC έχει την καλύτερη απόδοση το μοντέλο Transformer, ενώ με τη μετρική ROC AUC έχει την καλύτερη απόδοση το μοντέλο GRU AE. Παρ' όλα αυτά ο ασθενής #1 έχει το μεγαλύτερο σκορ και για τις δύο μετρικές ενώ το χειρότερο σκορ το έχει ο ασθενής #2. Για τους ίδιους λόγους που έχουμε αναφέρει και στα προηγούμενα πειράματα οι δύο αυτή ασθενείς έχουν τα καλύτερα και χειρότερα αποτελέσματα, αντίστοιχα. Γενικά, όλα μας τα αποτελέσματα είναι μεγαλύτερα από τον random classifier.

Patients	PR-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.64(0.56 – 0.70)	0.70(0.56 – 0.79)	<b>0.91(0.85-0.92)</b>	0.67(0.59 – 0.80)	0.89(0.83 – 0.89)
#2	0.03(0.03 – 0.05)	0.03(0.03 – 0.04)	0.03(0.02 – 0.06)	<b>0.05(0.04-0.1)</b>	0.03(0.03 – 0.05)
#3	0.68(0.66 – 0.71)	<b>0.71(0.70-0.72)</b>	0.68(0.66 – 0.70)	0.68(0.66 – 0.70)	0.68(0.67 – 0.68)
#4	0.18(0.16 – 0.19)	0.18(0.17 – 0.21)	0.16(0.15 – 0.17)	<b>0.22(0.17-0.22)</b>	0.15(0.15 – 0.16)
#5	0.55(0.50 – 0.60)	<b>0.56(0.55-0.60)</b>	0.51(0.46 – 0.55)	0.49(0.48 – 0.48)	0.50(0.47 – 0.55)
#6	0.63(0.57 – 0.64)	0.62(0.58 – 0.63)	<b>0.67(0.67-0.68)</b>	0.67(0.60 – 0.68)	0.67(0.66 – 0.67)
#7	<b>0.78(0.72-0.80)</b>	0.75(0.75 – 0.79)	0.66(0.56 – 0.66)	0.70(0.68 – 0.74)	0.70(0.61 – 0.70)
#8	0.41(0.39 – 0.43)	0.43(0.41 – 0.51)	<b>0.64(0.49-0.64)</b>	0.46(0.46 – 0.56)	0.62(0.51 – 0.65)
#9	0.71(0.68 – 1.00)	0.85(0.58 – 1.00)	0.65(0.48 – 1.00)	<b>0.95(0.95-1.00)</b>	0.61(0.45 – 0.95)
#10	0.61(0.54 – 0.68)	<b>0.61(0.59-0.70)</b>	0.60(0.50 – 0.71)	0.60(0.56 – 0.66)	0.55(0.55 – 0.70)
<b>Median</b>	0.62(0.46 – 0.66)	0.61(0.49 – 0.70)	<b>0.64(0.55-0.66)</b>	0.63(0.48 – 0.68)	0.61(0.53 – 0.67)

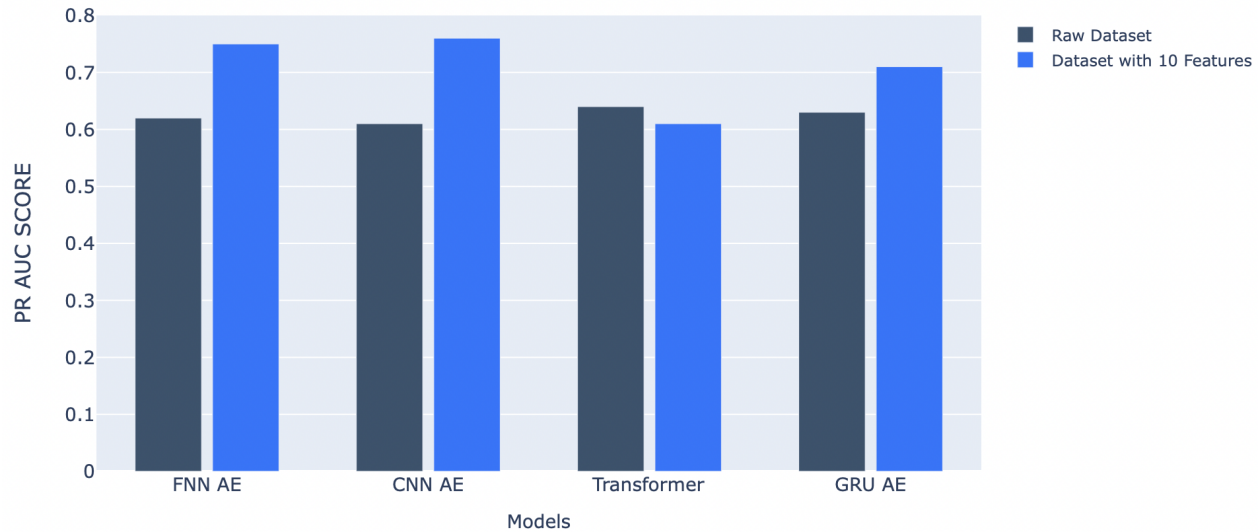
Πίνακας 5.8: Αποτελέσματα για **PR-AUC** σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα, τα οποία εκπαιδεύτηκαν και αξιολογήθηκαν στα raw δεδομένα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Patients	ROC-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.49(0.42 – 0.54)	0.55(0.46 – 0.57)	<b>0.77(0.74-0.82)</b>	0.59(0.48 – 0.68)	0.71(0.70 – 0.76)
#2	0.32(0.32 – 0.41)	0.22(0.21 – 0.35)	0.35(0.21 – 0.44)	<b>0.53(0.40-0.53)</b>	0.28(0.26 – 0.34)
#3	0.50(0.48 – 0.56)	0.56(0.55 – 0.59)	0.57(0.50 – 0.58)	<b>0.58(0.53-0.59)</b>	0.58(0.51 – 0.61)
#4	0.33(0.31 – 0.34)	0.32(0.30 – 0.33)	0.25(0.24 – 0.38)	<b>0.42(0.32-0.44)</b>	0.23(0.22 – 0.26)
#5	0.43(0.40 – 0.48)	<b>0.44(0.42-0.45)</b>	0.30(0.29 – 0.33)	0.37(0.25 – 0.39)	0.29(0.28 – 0.35)
#6	0.34(0.32 – 0.37)	0.33(0.32 – 0.41)	0.44(0.44 – 0.47)	0.43(0.41 – 0.44)	0.47(0.44 – 0.53)
#7	0.64(0.62 – 0.70)	<b>0.65(0.63-0.72)</b>	0.47(0.46 – 0.60)	0.58(0.57 – 0.59)	0.54(0.50 – 0.62)
#8	0.48(0.45 – 0.53)	0.53(0.48 – 0.53)	0.63(0.62 – 0.66)	0.61(0.60 – 0.61)	0.66(0.63 – 0.67)
#9	0.50(0.42 – 1.00)	0.75(0.42 – 1.00)	0.58(0.031 – 1.00)	<b>0.92(0.75-1.00)</b>	0.50(0.34 – 0.75)
#10	0.63(0.60 – 0.64)	0.70(0.68 – 0.72)	<b>0.70(0.69-0.75)</b>	0.67(0.59 – 0.68)	0.62(0.57 – 0.67)
<b>Median</b>	0.48(0.40 – 0.50)	0.54(0.68 – 0.61)	0.52(0.39 – 0.61)	<b>0.58(0.48-0.60)</b>	0.52(0.38 – 0.60)

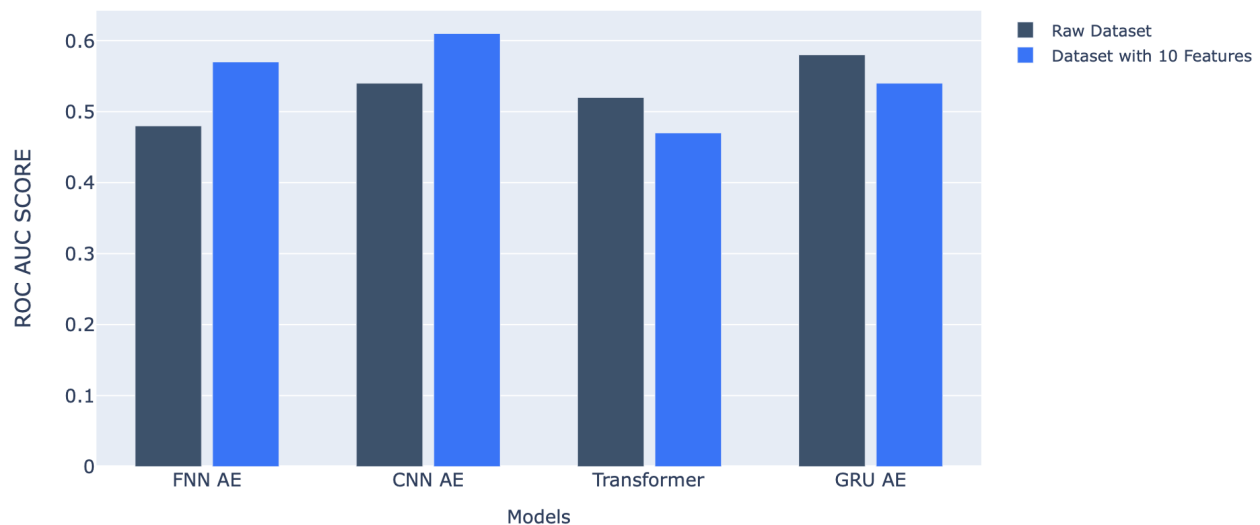
Πίνακας 5.9: Αποτελέσματα για **ROC-AUC** σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα, τα οποία εκπαιδεύτηκαν και αξιολογήθηκαν στα raw δεδομένα. Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Αξίζει να σημειωθεί, ότι προσθέτοντας χαρακτηριστικά και υλοποιώντας μεθόδους επεξεργασίας σήματος στα raw δεδομένα οδηγηθήκαμε σε πολύ καλύτερες αποδόσεις, όπως φαίνεται και στα

Σχήματα 5.2.5 και 5.2.6 και επαληθεύσαμε μέσω των πειραματισμών μας. Συγκριτικά, τα median των καλύτερων μετρικών PR και ROC AUC στα raw δεδομένα είναι 0.64 και 0.58, αντίστοιχα, ενώ στα δεδομένα με τα δέκα χαρακτηριστικά τα καλύτερα αποτελέσματα στις μετρικές PR και ROC AUC είναι 0.76 και 0.61 αντίστοιχα. Δηλαδή, για τα καλύτερα αποτελέσματα κάθε πειράματος παρατηρούμε αύξηση στο PR AUC κατά 12% και στο ROC AUC κατά 3%, όταν προστέθηκαν χαρακτηριστικά. Όπως παρατηρούμε και στα Σχήματα 5.2.5 και 5.2.6 τα μοντέλα CNN και FNN AE για τα δεδομένα με τα 10 χαρακτηριστικά, έχουν τα καλύτερα αποτελέσματα με αρκετή διαφορά από τα υπόλοιπα.



Σχήμα 5.2.5: Median PR AUC score για τα εξατομικευμένα μοντέλα μας, χρησιμοποιώντας τα Raw δεδομένα (στήλη με σκούρο μπλε χρώμα) και τα δεδομένα με τα 10 χαρακτηριστικά (γαλάζια στήλη).



Σχήμα 5.2.6: Median ROC AUC score για τα εξατομικευμένα μοντέλα μας, χρησιμοποιώντας τα Raw δεδομένα (στήλη με σκούρο μπλε χρώμα) και τα δεδομένα με τα 10 χαρακτηριστικά (γαλάζια στήλη).

Καταλήγουμε στο συμπέρασμα ότι όταν προσθέσαμε χαρακτηριστικά στα raw δεδομένα η απόδοση των μοντέλων έχει αυξηθεί αισθητά. Επομένως, όλα τα υπόλοιπα πειράματα θα συνεχιστούν εφαρμόζοντας τα δεδομένα με τα 10 χαρακτηριστικά, τα οποία εξήχθησαν με μεθόδους επεξεργασίας σήματος.

## 5.2.2 Global Μοντέλα

Παρόλο που σε επίπεδο πληθυσμού η ανίχνευση υποτροπών είναι μια πρόκληση λόγω τις πολυπλοκότητας και της διαφορετικότητας των ψυχωτικών ασθενών, θα παρουσιάσουμε πειράματα τα οποία σχετίζονται με global μοντέλα, έτσι ώστε να συγκρίνουμε την απόδοση τους με τα εξατομικευμένα. Για αυτό το σκοπό εκπαιδεύουμε ένα γενικό (global) μοντέλο, δηλαδή η εκπαίδευση γίνεται με συνενωμένα τα δεδομένα εκπαίδευσης και των 10 ασθενών. Ακολούθως, δοκιμάζουμε αυτό το εκπαιδευμένο μοντέλο στα δεδομένα αξιολόγησης κάθε ασθενή ξεχωριστά και στα συνενωμένα δεδομένα αξιολόγησης όλων των ασθενών μαζί. Τα αποτελέσματα του πειράματος παρουσιάζονται στους Πίνακες 5.10 και 5.11.

Patients	PR-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.83(0.79 – 0.87)	<b>0.99(0.98-1.00)</b>	0.87(0.60 – 0.93)	0.87(0.84 – 0.91)	0.91(0.90 – 0.99)
#2	0.04(0.03 – 0.05)	0.05(0.03 – 0.04)	<b>0.07(0.05-0.08)</b>	0.04(0.03 – 0.05)	0.03(0.03 – 0.04)
#3	0.50(0.43 – 0.66)	0.46(0.44 – 0.67)	0.47(0.47 – 0.64)	0.46(0.43 – 0.60)	0.53(0.52 – 0.64)
#4	0.15(0.15 – 0.17)	0.18(0.18 – 0.27)	<b>0.21(0.13-0.28)</b>	0.17(0.14 – 0.20)	0.18(0.14 – 0.18)
#5	<b>0.63(0.49-0.78)</b>	0.61(0.48 – 0.85)	0.58(0.53 – 0.80)	0.62(0.47 – 0.72)	0.63(0.50 – 0.83)
#6	0.70(0.63 – 0.74)	0.68(0.62 – 0.68)	<b>0.72(0.59-0.72)</b>	0.68(0.66 – 0.73)	0.68(0.60 – 0.78)
#7	0.84(0.82 – 0.88)	<b>0.89(0.78-0.90)</b>	0.88(0.87 – 0.88)	0.86(0.84 – 0.89)	0.86(0.84 – 0.87)
#8	0.85(0.71 – 0.86)	<b>0.87(0.81-0.93)</b>	0.83(0.74 – 0.88)	0.78(0.70 – 0.86)	0.85(0.81 – 0.89)
#9	0.89(0.71 – 0.89)	0.75(0.68 – 0.79)	0.80(0.54 – 0.92)	<b>0.92(0.82-1.00)</b>	0.68(0.48 – 1.00)
#10	0.96(0.95 – 0.98)	0.95(0.91 – 0.97)	0.91(0.91 – 0.96)	0.91(0.88 – 0.95)	0.97(0.95 – 1.00)
<b>Median</b>	<b>0.77(0.57-0.84)</b>	0.71(0.53 – 0.88)	0.76(0.52 – 0.85)	0.73(0.54 – 0.87)	0.68(0.58 – 0.85)
<b>Global</b>	0.48(0.48 – 0.49)	0.49(0.47 – 0.50)	0.47(0.46 – 0.48)	<b>0.52(0.50-0.52)</b>	0.50(0.48 – 0.50)

Πίνακας 5.10: Αποτελέσματα για PR-AUC σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα global μοντέλα που αξιολογήθηκαν ξεχωριστά για κάθε ασθενή αλλά και για όλους τους ασθενείς μαζί (επισημαίνεται ως Global). Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Στους Πίνακες 5.10 και 5.11 παρατηρούμε ότι το FNN AE μοντέλο έχει την καλύτερη απόδοση με PR και ROC AUC 0.77 και 0.62, αντίστοιχα. Η ψηλότερη απόδοση παρουσιάζεται στον ασθενή #1 με PR AUC 0.99 και ROC AUC 0.98 στο μοντέλο CNN AE. Στον ασθενή #1, παρατηρούμε μία πολύ μικρή αύξηση της απόδοσης του σε σχέση με το εξατομικευμένο μοντέλο, η οποία για τις μετρικές PR και ROC AUC ήταν στο 0.97. Ο λόγος αυτής της μικρής αύξησης, πιθανόν να οφείλεται στο γεγονός ότι η εκπαίδευση των μοντέλων μας έγινε με πολύ περισσότερα κανονικά δεδομένα, από ότι στα εξατομικευμένα μοντέλα, με αποτέλεσμα τα global μοντέλα να μάθουν περισσότερες φυσιολογικές συμπεριφορές και να μπορούν να βρουν ευκολότερα τις διαφορές με τις μη φυσιολογικές συμπεριφορές. Η χειρότερη απόδοση παρουσιάζεται στον ασθενή #2, η οποία είναι λίγο βελτιωμένη σε σχέση με τη μετρική PR AUC του εξατομικευμένου μοντέλου, δεν παύει όμως να ισχύει ότι δεν μπορούμε να βελτιώσουμε αυτή την απόδοση,

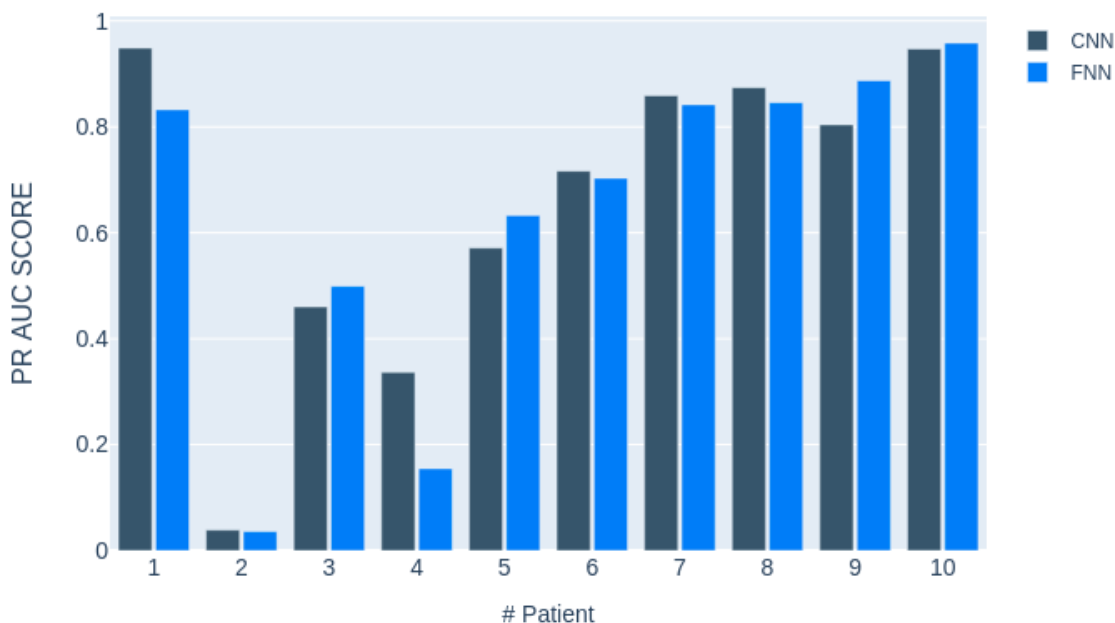
με τόσα λίγα δεδομένα και τόσο imbalanced dataset. Παρ' όλα αυτά όλα τα αποτελέσματα είναι μεγαλύτερα από τον random classifier. Τέλος, το global μοντέλο παρατηρούμε ότι έχει αρκετά χαμηλότερη απόδοση από τα μοντέλα που αξιολογούνται μεμονωμένα. Η χαμηλή αυτή απόδοση πιθανόν να οφείλεται σε διάφορους λόγους. Ένας από αυτούς μπορεί να είναι το γεγονός ότι στα δεδομένα του test set μας, τα οποία είναι τα συνενωμένα όλα τα test set των ασθενών, οι περισσότερες υποτροπές είναι χαμηλής σοβαρότητας, με αποτέλεσμα τα μοντέλα μας να δυσκολεύονται να την ανιχνεύσουν και η συνολική απόδοση τους να πέφτει.

Patients	ROC-AUC				
	FNN AE	CNN AE	Transformer	GRU AE	Random
#1	0.88(0.83 – 0.88)	<b>0.98(0.98-1.00)</b>	0.85(0.66 – 0.91)	0.90(0.88 – 0.92)	0.91(0.90 – 0.97)
#2	0.37(0.34 – 0.40)	0.28(0.27 – 0.36)	<b>0.49(0.48-0.50)</b>	0.34(0.27 – 0.40)	0.28(0.23 – 0.30)
#3	<b>0.57(0.51-0.59)</b>	0.57(0.51 – 0.59)	0.51(0.46 – 0.57)	0.50(0.46 – 0.51)	0.52(0.52 – 0.54)
#4	0.22(0.20 – 0.25)	<b>0.34(0.25-0.36)</b>	0.29(0.25 – 0.35)	0.27(0.24 – 0.39)	0.22(0.18 – 0.24)
#5	<b>0.42(0.34-0.66)</b>	0.40(0.30 – 0.66)	0.27(0.24 – 0.36)	0.42(0.30 – 0.55)	0.42(0.35 – 0.68)
#6	<b>0.49(0.45-0.51)</b>	0.39(0.35 – 0.49)	0.33(0.32 – 0.37)	0.48(0.35 – 0.52)	0.48(0.46 – 0.48)
#7	0.67(0.61 – 0.71)	0.68(0.58 – 0.77)	<b>0.73(0.70-0.75)</b>	0.69(0.64 – 0.72)	0.62(0.61 – 0.62)
#8	0.71(0.70 – 0.74)	<b>0.78(0.78-0.81)</b>	0.75(0.66 – 0.75)	0.64(0.60 – 0.64)	0.72(0.70 – 0.74)
#9	0.69(0.50 – 0.83)	0.58(0.28 – 0.58)	0.53(0.33 – 0.83)	<b>0.83(0.67-1.00)</b>	0.42(0.25 – 1.00)
#10	0.88(0.88 – 0.90)	0.81(0.81 – 0.88)	0.69(0.69 – 0.90)	0.84(0.69 – 0.88)	0.91(0.88 – 1.00)
<b>Median</b>	<b>0.62(0.45-0.70)</b>	0.58(0.40 – 0.73)	0.52(0.41 – 0.71)	0.57(0.45 – 0.76)	0.50(0.42 – 0.67)
<b>Global</b>	0.47(0.44 – 0.50)	0.51(0.51 – 0.54)	0.45(0.45 – 0.50)	<b>0.53(0.51-0.56)</b>	0.50(0.50 – 0.54)

Πίνακας 5.11: Αποτελέσματα για **ROC-AUC** σε δεδομένα διαστήματος 5 λεπτών για τα μοντέλα που αναλύθηκαν στην Ενότητα 4.2. Παρουσιάζεται η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα global μοντέλα που αξιολογήθηκαν ξεχωριστά για κάθε ασθενή αλλά και για όλους τους ασθενείς μαζί (επισημαίνεται ως Global). Η μεγαλύτερη τιμή για κάθε ασθενή σημειώνεται με έντονη γραφή.

Συμπερασματικά, διαπιστώνουμε ότι το global μοντέλο έχει εμφανή χαμηλότερη απόδοση από το median των εξατομικευμένων μοντέλων. Αυτό μπορεί να οφείλεται στο γεγονός ότι αρκετές από τις υποτροπές έχουν χαμηλή σοβαρότητα με αποτέλεσμα το μοντέλο μας να δυσκολεύεται τις εντοπίσει και έτσι η συνολική απόδοση του μοντέλου να είναι τόσο χαμηλή. Για να επαληθεύσουμε αυτό τον ισχυρισμό θα κάνουμε ένα πείραμα στην Ενότητα 5.2.3, για να δούμε πόσο επηρεάζει η σοβαρότητα της υποτροπής τον εντοπισμό της. Επίσης, καταλήγουμε στο συμπέρασμα ότι έχουμε καλύτερη επίδοση στο global μοντέλο FNN AE το οποίο όμως αξιολογείται μεμονωμένα ο κάθε ασθενής με ποσοστά PR και ROC AUC στο 0.77 και 0.62, αντίστοιχα. Η επίδοση αυτού του μοντέλου σε σχέση με το εξολοκλήρου εξατομικευμένο μοντέλο CNN AE είναι πολύ κοντά (κατά 0.01).

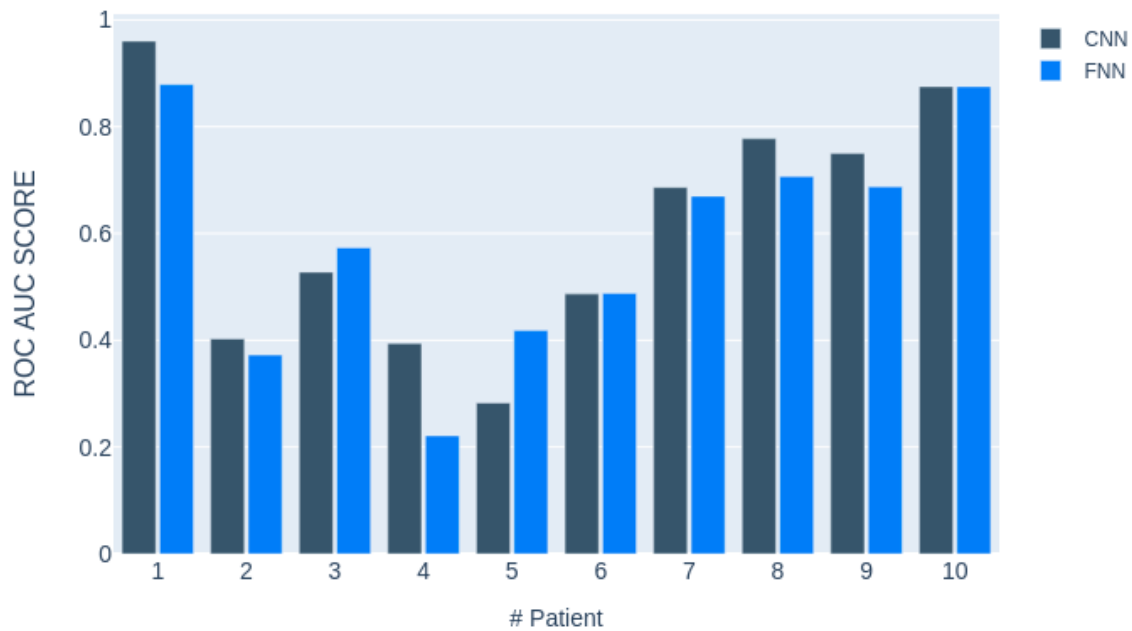
Στα Σχήματα 5.2.7 και 5.2.8 παρουσιάζονται συνοπτικά τα αποτελέσματα για τους ασθενείς στο global μοντέλο FNN AE το οποίο όμως αξιολογείται μεμονωμένα ο κάθε ασθενή (σημειωμένο με γαλάζιο χρώμα) και στο εξολοκλήρου εξατομικευμένο μοντέλο CNN AE (σημειωμένο με μπλε χρώμα).



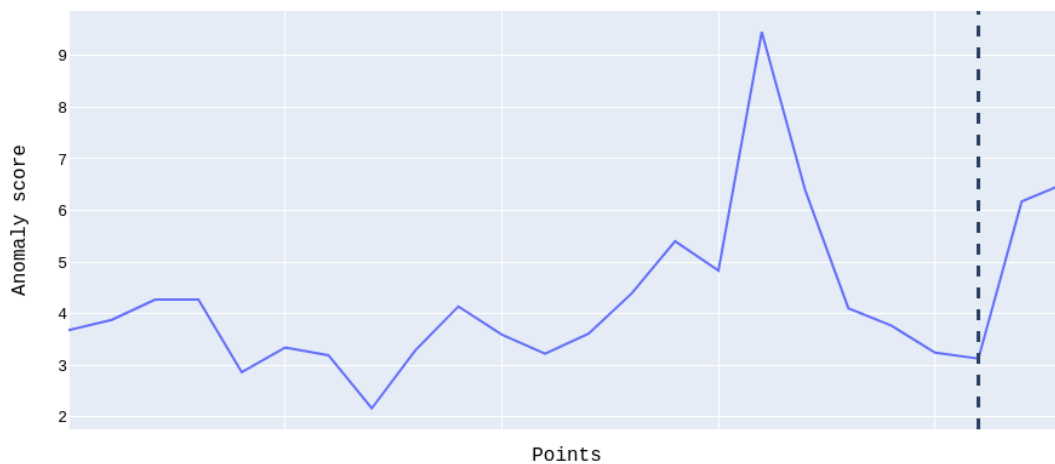
Σχήμα 5.2.7: PR AUC για το global μοντέλο FNN το οποίο όμως αξιολογείται μεμονωμένα ο κάθε ασθενή και για το εξολοκλήρου εξατομικευμένο μοντέλο CNN.

Από τα Σχήματα 5.2.7 και 5.2.8, παρατηρούμε ότι οι ασθενείς #1 και #10 έχουν τις καλύτερες επιδόσεις και για τις δύο μετρικές. Ο ασθενής #1 έχει μία moderate υποτροπή και ο ασθενής #10 μία severe. Αυτό μας οδηγεί στο συμπέρασμα ότι όσο σοβαρότερη είναι μία υποτροπή τόσο πιο εύκολα ανιχνεύεται από το μοντέλο μας. Και στους 2 αυτούς ασθενείς η αναλογία των κανονικών δεδομένων και των δεδομένων υποτροπής είναι σχετικά καλή, δηλαδή δεν μπορούμε να πούμε ότι έχουμε imbalanced dataset, αλλά πρέπει να σημειώσουμε ότι αν είχαμε περισσότερα δεδομένα για όλους τους ασθενείς θα είχαμε σαφώς και καλύτερα αποτελέσματα.

Ο ασθενής #2 έχει imbalanced dataset (Normal δεδομένα 317 ημερών ενώ Relapse δεδομένα 2 ημερών) με το σύνολο δεδομένων να περιέχει μόνο δύο μέρες υποτροπής. Αυτό έχει ως αποτέλεσμα τη χαμηλότερη απόδοση στη μετρική PR AUC η οποία είναι κοντά στο 0, ενώ με τη μετρική ROC AUC η απόδοση του δεν είναι τόσο κακή, λόγω του ότι η μετρική ROC AUC επικεντρώνεται στο θετικό αποτέλεσμα της πλειοψηφίας. Το μοντέλο μας θεωρεί ότι σχεδόν όλα μας τα σημεία δεν είναι σημεία υποτροπής και δεν μπορεί να εντοπίσει τις δύο συγκεκριμένες μέρες υποτροπής. Αυτό πιθανόν να οφείλεται στο γεγονός ότι οι δύο μέρες υποτροπής του συγκεκριμένου ασθενή είναι χαμηλής σοβαρότητας και το μοντέλο μας να μην μπορεί να τις ανιχνεύσει. Όπως φαίνεται και στο Σχήμα 5.2.9, οι μέρες ψυχωτικής υποτροπής για τις οποίες έχουμε δεδομένα είναι μόνο 2, στις οποίες βλέπουμε το αναμενόμενο σε σχέση με τα αποτελέσματά μας, δηλαδή υψηλότερα anomaly scores στις μέρες μη υποτροπής και χαμηλότερα anomaly scores στις μέρες υποτροπής.



Σχήμα 5.2.8: ROC AUC για το global μοντέλο FNN το οποίο όμως αξιολογείται μεμονωμένα ο κάθε ασθενή και για το εξολοκλήρου εξατομικευμένο μοντέλο CNN.



Σχήμα 5.2.9: Γραφική παράσταση του Anomaly Score του ασθενή #2 στο μοντέλο CNN στο διάστημα 5 λεπτών. Το anomaly score δεξιά της διαχωριστικής γραμμής, είναι οι μέρες στις οποίες ο ασθενής είχε διαγνωστεί με ψυχωτική υποτροπή, ενώ το anomaly score αριστερά της διαχωριστικής γραμμής είναι οι μέρες όπου ο ασθενής δεν είχε κάποιο σύμπτωμα υποτροπής ('κανονικές' μέρες). Παρατηρούμε ότι οι μέρες ψυχωτικής υποτροπής για τις οποίες έχουμε δεδομένα είναι μόνο 2, στις οποίες δεν βλέπουμε το αναμενόμενο, δηλαδή υψηλότερα anomaly scores στις μέρες υποτροπής από τις κανονικές μέρες. Να σημειώσουμε ότι τα Points στον άξονα- $x$  είναι οι μέρες, οι οποίες δεν είναι συνεχόμενες, δηλαδή τα απεικονιζόμενα anomaly scores δεν είναι συνεχή.



### 5.2.3 Σοβαρότητα Υποτροπής

Σε αυτή την Ενότητα θα παρουσιάσουμε διάφορα πειράματα για να ελέγξουμε πόσο επηρεάζει η σοβαρότητα υποτροπής, τον εντοπισμό της. Σε αυτό το πείραμα ελέγχουμε πόσο επηρεάζει η σοβαρότητα της υποτροπής, ελέγχοντας για κάθε ασθενή κάθε υποτροπή ξεχωριστά. Μετά από τη προεπεξεργασία των δεδομένων όπως είδαμε στην Ενότητα 3, οι ασθενείς που μας δίνουν τη δυνατότητα να μελετήσουμε τη επιρροή που έχει η σοβαρότητα υποτροπής είναι οι #5, #6 και #7, αφού αντιμετώπισαν περισσότερες από μία περιπτώσεις υποτροπής (Low και Moderate).

Στον Πίνακα 5.12 παρουσιάζεται το σφάλμα ανακατασκευής (reconstruction error), Mean Square Error των εξατομικευμένων μοντέλων μας για τους τρεις ασθενείς.

Patients	FNN AE		CNN AE		Transformer		GRU AE	
	Low	Moderate	Low	Moderate	Low	Moderate	Low	Moderate
#5	0.005946	0.005119	0.007824	0.006617	0.004306	0.003363	0.008364	0.008364
#6	0.004421	0.004348	0.005098	0.005481	0.002341	0.002944	0.006973	0.007549
#7	0.005024	0.005077	0.006257	0.006515	0.004532	0.004618	0.009152	0.010003

Πίνακας 5.12: Σφάλμα Ανακατασκευής - Mean Square Error στα εξατομικευμένα μοντέλα, για Low και Moderate υποτροπές ξεχωριστά.

Από τον Πίνακα 5.12 παρατηρούμε ότι για τον ασθενή #7 σε όλα τα μοντέλα το σφάλμα ανακατασκευής της Moderate υποτροπής είναι μεγαλύτερο, κάτι το οποίο ήταν αναμενόμενο, αφού όσο χειρότερη είναι μία υποτροπή τόσο περισσότερες ανωμαλίες αναμένουμε να έχουμε στα δεδομένα μας. Στον ασθενή #6 σε όλα τα μοντέλα εκτός από το FNN AE παρατηρούμε ότι και πάλι η Moderate υποτροπή έχει μεγαλύτερο σφάλμα ανακατασκευής. Στο ασθενή #5 βλέπουμε ότι για όλα τα μοντέλα το σφάλμα ανακατασκευής της Low υποτροπής είναι μεγαλύτερο από της Moderate. Αυτό πιθανόν να οφείλεται στο γεγονός ότι για τον συγκεκριμένο ασθενή έχουμε πολύ λίγα δεδομένα (2 ωρών) για αυτή την υποτροπή, επομένως τα αποτελέσματά μας πιθανόν να μην είναι ακριβείς.

Στον Πίνακα 5.13 παρουσιάζεται το σφάλμα ανακατασκευής (reconstruction error), Mean Square Error των global μοντέλων μας στα οποία έγινε εξατομικευμένη αξιολόγηση των ασθενών, για τους τρεις ασθενείς οι οποίοι έχουν περισσότερες από μία υποτροπή. Από τα αποτελέσματα διαπιστώνουμε ότι για τον ασθενή #7 το σφάλμα ανακατασκευής δεν είναι σε όλα τα μοντέλα μεγαλύτερο στη Moderate υποτροπή από ότι στη Low, όπως είχαμε παρατηρήσει στα εξατομικευμένα μοντέλα. Στο ασθενή #6, παρατηρούμε ότι για όλα τα μοντέλα το σφάλμα ανακατασκευής της Moderate υποτροπής είναι μεγαλύτερο από της Low. Για τον ασθενή #5 παρατηρούμε ότι μόνο στα μοντέλα Transformer και GRU AE το σφάλμα ανακατασκευής της Moderate υποτροπής είναι μεγαλύτερο από της Low.

Patients	FNN AE		CNN AE		Transformer		GRU AE	
	Low	Moderate	Low	Moderate	Low	Moderate	Low	Moderate
#5	0.004493	0.004342	0.007824	0.006617	0.003788	0.003933	0.006670	0.006690
#6	0.003960	0.003824	0.005098	0.005481	0.002927	0.003607	0.006684	0.006705
#7	0.005915	0.005951	0.008858	0.009142	0.008743	0.008171	0.009788	0.009408

Πίνακας 5.13: Σφάλμα Ανακατασκευής - Mean Square Error στα global μοντέλα με αξιολόγηση στα εξατομικευμένα δεδομένα, για Low και Moderate υποτροπές ξεχωριστά.



Στον Πίνακα 5.14 παρουσιάζεται το σφάλμα ανακατασκευής στα global μοντέλα για Low, Moderate και Severe υποτροπές ξεχωριστά. Δηλαδή, συγκεντρώσαμε τις υποτροπές ανά κατηγορία για όλους τους ασθενείς μαζί. Όπως φαίνεται και στον πιο κάτω πίνακα σε όλα τα μοντέλα εκτός από τους Transformer παρατηρείται μια κλιμακωτή αύξηση του σφάλματος ανακατασκευής σε σχέση με την σοβαρότητα της υποτροπής. Το αποτέλεσμα αυτό είναι αναμενόμενο αφού διαισθητικά περιμένουμε όσο πιο σοβαρή είναι μία υποτροπή, τόσο περισσότερες ανωμαλίες να περιέχουν τα δεδομένα μας άρα να είναι και πιο εύκολα ανιχνεύσιμη. Το γεγονός ότι το σφάλμα ανακατασκευής των low υποτροπών είναι χαμηλό, επαληθεύει την θεωρία που είχαμε ότι τα global μοντέλα μας έχουν τόσο χαμηλή απόδοση, λόγω του ότι τα περισσότερα δεδομένα είναι χαμηλής σοβαρότητας υποτροπής και είναι δύσκολο να εντοπιστούν, ενώ οι σοβαρότητες υποτροπής moderate και severe είναι πιο εύκολο να εντοπιστούν.

Patients	FNN AE	CNN AE	Transformer	GRU AE
Low	0.002598	0.003650	0.003370	0.005126
Moderate	0.002629	0.003759	0.003232	0.005274
Severe	<b>0.002796</b>	<b>0.004076</b>	0.003099	<b>0.005594</b>

Πίνακας 5.14: Σφάλμα Ανακατασκευής - Mean Square Error στα global μοντέλα, για Low, Moderate και Severe υποτροπές ξεχωριστά.

Συνοψίζοντας, συμπεραίνουμε ότι τα μοντέλα με τις καλύτερες αποδόσεις είναι ο FNN και ο CNN Autoencoder. Επίσης, παρατηρήσαμε ότι η σοβαρότητα της υποτροπής επηρεάζει αρκετά την ανίχνευση της. Δηλαδή, όσο πιο σοβαρή είναι μία υποτροπή τόσο πιο εύκολα ανιχνεύσιμη είναι. Τέλος, καταλήγουμε στο συμπέρασμα ότι το σημαντικότερο ρόλο έχουν τα δεδομένα. Όσο περισσότερα δεδομένα έχουμε στη διάθεση μας, τα οποία να μην είναι διακοπτόμενα τόσο καλύτερα αποτελέσματα θα έχει η μέθοδος που εκτελούμε.

#### 5.2.4 Πειράματα με thresholding

Σε όλα τα πειράματα τα οποία εκτελέστηκαν και παρουσιάστηκαν στις παραπάνω Ενότητες, έχει γίνει επιπλέον και η εφαρμογή ενός κατωφλιού (thresholding) για την ανίχνευση της υποτροπής. Για τον εντοπισμό του κατωφλιού (threshold) έγινε μεγιστοποίηση της μέτρησης απόδοσης  $f_{beta}$ , όπου  $beta = 0.05$ . Στον Πίνακα 5.15 παρουσιάζονται τα median αποτελέσματα για τα εξατομικευμένα μοντέλα, δεδομένων διαστήματος 5 λεπτών.

Patients	$f_{0.05}$	Sensitivity	Specificity
FNN AE	0.96(0.92 – 1.00)	0.47(0.14 – 0.74)	1.00(1.00 – 1.00)
CNN AE	0.97(0.88 – 1.00)	0.53(0.14 – 0.78)	1.00(0.98 – 1.00)
Transformer	0.94(0.73 – 0.99)	0.39(0.20 – 0.72)	1.00(0.87 – 1.00)
GRU AE	0.96(0.85 – 0.99)	<b>0.57(0.33 – 0.88)</b>	<b>1.00(0.80 – 1.00)</b>

Πίνακας 5.15: Αποτελέσματα για τα εξατομικευμένα μοντέλα που περιγράφηκαν στην Ενότητα 4.2, σε δεδομένα διαστήματος 5 λεπτών. Παρουσιάζεται συνολική η διάμεση τιμή (median) και το IQR (σε παρένθεση) για τα εξατομικευμένα μοντέλα.

Με τη μέθοδο του thresholding παρατηρούμε καλύτερα αποτελέσματα σε θέματα ακρίβειας. Ο GRU AE έχει τη καλύτερη απόδοση ενώ τη χειρότερη απόδοση έχει ο Transformer.

Τα αποτελέσματα μας δεν είναι άμεσα συγκρίσιμα με το άρθρο του Adler et al., [Adl+20] για πολλούς λόγους. Κατά κύριο λόγο δεν έχουμε ίδια δεδομένα, στο άρθρο [Adl+20] χρησιμοποιούνται δεδομένα από smartphone, ενώ στη δική μας εργασία χρησιμοποιούμε δεδομένα από smartwatch. Επίσης, στο άρθρο για την μέθοδο του thresholding γίνεται μεγιστοποίηση της ποσότητας  $TPR/FPR$ , στο validation set για την εύρεση του κατώφλιού ενώ στη δική μας εργασία μεγιστοποιούμε το  $f_{beta}$  score στο test set μας, όπως προτείνεται και από το άρθρο του Malhotra et al., [Mal+15]. Τέλος, στο άρθρο του Adler et al., [Adl+20] γίνεται προσπάθεια για υλοποίηση global μοντέλων, ενώ εμείς εφαρμόσαμε τη μέθοδο του thresholding στα εξατομικευμένα μοντέλα. Παρ' όλα αυτά, συγκρίνοντας 'έμμεσα' τα καλύτερα μας αποτελέσματα με αυτά του Adler et al. μπορούμε να πούμε ότι έχουμε καλύτερη απόδοση.

Methods	Sensitivity, median (IQR)	Specificity, median (IQR)
FNN AD [Adl+20]	0.25(0.15 – 1.00)	0.88(0.14 – 0.96)
Our GRU AE	0.57(0.33 – 0.88)	1.00(0.80 – 1.00)

Πίνακας 5.16: Σύγκριση απόδοσης με το άρθρο [Adl+20].

Γενικά, τα αποτελέσματα αυτής της μεθόδου είναι συγκεκριμένα για αυτά τα δεδομένα. Δηλαδή, για τα συγκεκριμένα δεδομένα επιλέγεται το καλύτερο κατώφλι που θα μας δώσει τα καλύτερα αποτελέσματα. Πιθανόν, το συγκεκριμένο κατώφλι να μην μπορεί να γενικευτεί σε άλλα δεδομένα ή σε διαφορετικό διαχωρισμό των δεδομένων. Επομένως, για μια πιο γενική αξιολόγηση των μοντέλων μας προτιμήσαμε να εργαστούμε με τις μετρικές AUC, οι οποίες μας δίνουν μια καλύτερη οπτική για την απόδοση των μοντέλων μας. Δηλαδή, τη συμπεριφορά του μοντέλου για όλα τα κατώφλια.

# Κεφάλαιο 6

## Συμπεράσματα και Μελλοντικές Επεκτάσεις

### 6.1 Σύνοψη και Συμπεράσματα

Στη παρούσα διπλωματική εργασία μελετήσαμε σε βάθος το πρόβλημα της ανίχνευσης ψυχωτικών υποτροπών σε ασθενείς με ψυχωτικές διαταραχές του φάσματος της σχιζοφρένειας. Η λύση του προβλήματος αυτού μπορεί να αντιμετωπίσει διάφορα ζητήματα υγείας και ευεξίας σε ασθενείς με ψυχωτικές διαταραχές, οδηγώντας σε ένα καλύτερο αύριο για τους ίδιους τους ασθενείς και τους οίκους τους. Καθώς επίσης και σε μια πιο άμεση αντιμετώπιση των διαταραχών, κάτι το οποίο ίσως θα διευκόλυνε τους ιατρούς και τις ψυχιατρικές δομές.

Η μελέτη μας, ξεκίνησε με μία βιβλιογραφική ανασκόπηση στα σχετικά θέματα. Οι μελέτες κυμαίνονται από την κατανόηση της συμπεριφοράς των ατόμων με ψυχωτικές διαταραχές μέσω στατιστικών μοντέλων και νευρωνικών δικτύων έως στην έγκαιρη πρόβλεψη υποτροπών. Στη δική μας εργασία, εκμεταλλευτήκαμε τα φυσιολογικά σήματα κίνησης και παρακολούθησης του καρδιακού παλμού, που συλλέγονται από έξυπνα ρολόγια, τα οποία επεξεργαστήκαμε κατάλληλα, εξάγοντας διάφορες αναπαραστάσεις και χαρακτηριστικά, που θεωρούνται ως συναφή για την ανίχνευση υποτροπών. Τέλος, σχεδιάσαμε διάφορα μοντέλα (encoder-decoder) νευρωνικών δικτύων, όπως Fully Connected, Convolutional, GRU Autoencoder και Transformer για τη ανίχνευση των ανωμαλιών.

Αρχικά, μια από τις αρχικές μας παρατηρήσεις είναι το πόσο σημαντικός είναι ο όγκος των δεδομένων, αφού πολλά από τα προβλήματα που είχαμε στη δική μας έρευνα οφείλονταν στα ελλιπή δεδομένα. Παρά το γεγονός ότι τα δεδομένα μας ήταν μακράς διάρκειας, πολλά από αυτά τα οποία συλλέχθηκαν από το smartwatch που φορούσαν οι ασθενείς, ήταν διακοπτόμενα με αποτέλεσμα να μην μπορούμε να τα συμπληρώσουμε κατάλληλα και έτσι να αναγκαζόμαστε να τα αφαιρούμε. Επίσης, να σημειώσουμε ότι με τα δεδομένα που είχαμε μέχρι τον Ιούνιο του 2021 μπορούσαμε να ελέγξουμε για υποτροπή μόνο τους 5 από τους 10 ασθενείς, όμως όταν προστέθηκαν δεδομένα επιπλέον τεσσάρων μηνών, δηλαδή από τον Ιούνιο του 2021 μέχρι των Σεπτέμβριο του 2021, καταφέραμε να αυξήσουμε τους ασθενείς στους δέκα, όπως παρουσιάστηκαν και στην εργασία μας.

Τα πειράματα τα οποία εκτελέστηκαν ήταν αρκετά για να βγάλουμε συμπεράσματα, χρήσιμα για

την ανίχνευση υποτροπών σε ασθενής με ψυχωτικές διαταραχές. Εκπαιδεύσαμε τα μοντέλα μας (Fully Connected Autoencoder, Convolutional Autoencoder, GRU Autoencoder και Transformer) ως εξατομικευμένα μοντέλα και ως global. Παρατηρήσαμε ότι, υπάρχουν ενθαρρυντικά αποτελέσματα για τις περιπτώσεις όπου έχουμε περισσότερα δεδομένα υποτροπής ή η υποτροπή έχει υψηλότερη σοβαρότητα. Σύμφωνα με τα αποτελέσματά μας, ένα μικρότερο granularity στο σύνολο δεδομένων (5-min) καταλήγει σε καλύτερη απόδοση στην ανίχνευση των υποτροπών. Έχουμε παρατηρήσει ότι όταν προσθέσαμε χαρακτηριστικά στα raw δεδομένα μας, η απόδοση των μοντέλων μας έχει αυξηθεί, με μεγαλύτερη αύξηση για τη μετρική PR και ROC AUC να είναι στο 15% και στο 9%, αντίστοιχα. Επιπλέον, στην περίπτωση μας τα μοντέλα FNN και CNN AE στα πειράματα με δεδομένα από τα 10 χαρακτηριστικά, έχουν καλύτερες επιδόσεις είτε ως εξατομικευμένα μοντέλα είτε ως global μοντέλα που αξιολογήθηκαν σε εξατομικευμένα δεδομένα. Παρόλο που έχουν παρατηρηθεί μικρότερες επιδόσεις για κάποια από τα μοντέλα μας, όλα μας τα αποτελέσματα έχουν υψηλότερες τιμές από το random classifier. Τέλος, έχουμε παρατηρήσει ότι η σοβαρότητα της υποτροπής επηρεάζει την ανίχνευση της, με άλλα λόγια όσο σοβαρότερη είναι μία υποτροπή τόσο πιο εύκολη είναι και η ανίχνευση της.

Στην επόμενη ενότητα, με βάση τα συμπεράσματα αυτά προτείνουμε κάποιες προτάσεις για βελτίωση των παραπάνω μεθόδων και επίτευξη ακόμη καλύτερων αποτελεσμάτων.

## 6.2 Μελλοντικές Επεκτάσεις

Προτεινόμενες μελλοντικές επεκτάσεις για την παρούσα μελέτη είναι οι ακόλουθες:

- Το έργο e-Prevention πραγματοποιείται σχεδόν δύο χρόνια με συνεχόμενη συλλογή δεδομένων από ασθενείς με ψυχωτικές διαταραχές. Σε μελλοντικό στάδιο, η ύπαρξη περισσότερων δεδομένων και υποτροπών, όπως και η ύπαρξη περισσότερων ασθενών θα μπορούσε να βοηθήσει στο να έχουμε καλύτερα αποτελέσματα αλλά και στο να βγάλουμε καλύτερα συμπεράσματα. Επίσης, επιπλέον χαρακτηριστικά και αναπαραστάσεις των αρχικών δεδομένων τα οποία έχουν προταθεί και σε περαιτέρω έρευνες, μπορούν να προστεθούν έτσι ώστε να δώσουν και αυτά κάποια επιπρόσθετη πληροφορία.
- Χρήσιμες πληροφορίες μπορεί να αποκτηθούν χωρίζοντας τα δεδομένα μας ανάλογα με τη κατάσταση του ασθενή, δηλαδή αν βρίσκεται σε κατάσταση εγρήγορσης (awake) ή ύπνου (sleep). Αυτή η μελέτη μπορεί να οδηγήσει σε χρήσιμα συμπεράσματα για την ανίχνευση της υποτροπής.
- Για επέκταση αυτής της μελέτης, προτείνεται περισσότερη διερεύνηση σε πιο σύνθετες αρχιτεκτονικές νευρωνικών δικτύων. Προτείνεται διερεύνηση των Convolutional αρχιτεκτονικών [Cho+22], αφού ο δικός μας απλός convolutional autoencoder μας έχει δώσει μία αρκετά καλή απόδοση. Επομένως, μια πιο σύνθετη αρχιτεκτονική ίδιας λογικής μπορεί να επιφέρει ακόμη καλύτερα αποτελέσματα. Επίσης, η υλοποίηση αρχιτεκτονικών με δυνατότητα πρόβλεψης θα ήταν χρήσιμη για την πρόβλεψη των υποτροπών.
- Η πρόβλεψη μιας υποτροπής θα οδηγήσει τους ιατρούς να βοηθήσουν τους ασθενείς για την αποτροπή της. Επομένως, ένα σύστημα το οποίο θα κάνει έγκαιρη πρόβλεψη υποτροπής θα είναι πολύ χρήσιμο. Στην πρώτη προσέγγιση, η οποία έχει προταθεί από τον Adler et al., [Adl+20], έχουν αφαιρεθεί τα πραγματικά δεδομένα υποτροπής και θεωρήθηκαν δεδομένα υποτροπής τα δεδομένα ενός μήνα πριν την υποτροπή. Έτσι, γίνεται

ανίχνευση υποτροπής στα δεδομένα αυτά, με αποτέλεσμα να μπορεί κάποιος να προβλέψει την υποτροπή περίπου ένα μήνα πριν αυτή συμβεί. Εναλλακτικά, αν θέλουμε να προβλέψουμε πότε ο ασθενής θα έχει υποτροπή, μπορούμε αντί να βασιστούμε στην ιδέα του Autoencoder, να δημιουργήσουμε ένα προβλέπτη (predictor) χρονοσειρών, ο οποίος θα προβλέπει τα δεδομένα μετά από ένα χρονικό διάστημα, και έπειτα θα γίνεται η ανίχνευση της υποτροπής στα προβλεπόμενα δεδομένα με την ίδια λογική.

## 6.3 Υποβολή Συνεδριακού Άρθρου

Βασισμένη σε αυτή τη διπλωματική εργασία, θα γίνει υποβολή ενός άρθρου στο συνέδριο EUSIPCO-2022, παρουσιάζοντας κάποιους από τους πειραματισμούς και τα αποτελέσματα μας. Για την αντιμετώπιση του προβλήματος της ανίχνευσης υποτροπών, παρουσιάζουμε τέσσερις αυτοencoder αρχιτεκτονικές (FNN AE, CNN AE, GRU AE και Transformer), που υλοποιούνται εξατομικευμένα για κάθε ασθενή και σε global επίπεδο με βάση τα συνολικά δεδομένα των ασθενών. Παρουσιάζουμε τα εκτεταμένα πειράματα, για δέκα ασθενείς που πάσχουν από ψυχωτικές διαταραχές του φάσματος της σχιζοφρένειας, τα οποία μας έδωσαν υποσχόμενα αποτελέσματα. Τέλος, πραγματοποιήσαμε μια post hoc ανάλυση χρησιμοποιώντας τα μοντέλα που απέδωσαν καλύτερα, για να εξετάσουμε τη σημασία του επιπέδου της σοβαρότητας υποτροπής μεταξύ 3 ασθενών που υποτροπίασαν πολλές φορές με διαφορετικό επίπεδο σοβαρότητας υποτροπής.

M. Panagiotou, A. Zlantintsi, P. P. Filntisis, A. J. Roumeliotis, N. Efthymiou and P. Maragos, “Unsupervised learning for relapse detection of patients suffering from psychosis using wearable sensors data”, under submission in the *30th European Signal Processing Conference (EUSIPCO-2022)*.



# Παράρτημα Α

## Βιβλιογραφία

- [Adl+20] Adler, D. A. et al. “Predicting early warning signs of psychotic relapse from passive sensing data: An approach using encoder-decoder neural networks”. In: *JMIR mHealth and uHealth* 8.8 (2020). ISSN: 22915222. DOI: [10.2196/19962](https://doi.org/10.2196/19962).
- [Age+02] Agelink, M. W. et al. “Relationship between major depression and heart rate variability.: Clinical consequences and implications for antidepressive treatment”. In: *Psychiatry Research* 113.1-2 (2002), pp. 139–149.
- [AM17] Albelwi, S. and Mahmood, A. “A framework for designing the architectures of deep convolutional neural networks”. In: *Entropy* 19.6 (2017), p. 242.
- [Alo+19] Alom, M. Z. et al. “A state-of-the-art survey on deep learning theory and architectures”. In: *Electronics* 8.3 (2019), p. 292.
- [AAM21] Alsobhani, A., ALabbodi, H. M., and Mahdi, H. “Speech Recognition using Convolution Deep Neural Networks”. In: *Journal of Physics: Conference Series*. Vol. 1973. 1. IOP Publishing. 2021, p. 012166.
- [AHS12] Anderson, I. M., Haddad, P. M., and Scott, J. “Bipolar disorder”. In: *British Medical Journal* 345 (2012). ISSN: 17561833. DOI: [10.1136/bmj.e8508](https://doi.org/10.1136/bmj.e8508). URL: <https://www.bmj.com/lookup/doi/10.1136/bmj.e8508>.
- [And95] Andreasen, N. C. “Symptoms, signs, and diagnosis of schizophrenia”. In: *The Lancet* 346.8973 (Aug. 1995), pp. 477–481. ISSN: 01406736. DOI: [10.1016/S0140-6736\(95\)91325-4](https://doi.org/10.1016/S0140-6736(95)91325-4).
- [AMC17] Aung, M. H., Matthews, M., and Choudhury, T. “Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies”. In: *Depression and Anxiety* 34.7 (2017), pp. 603–609.
- [Bal12] Baldi, P. “Autoencoders, unsupervised learning, and deep architectures”. In: *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*. JMLR Workshop and Conference Proceedings. 2012, pp. 37–49.
- [Bär+07] Bär, K.-J. et al. “Non-linear complexity measures of heart rate variability in acute schizophrenia”. In: *Clinical Neurophysiology* 118.9 (2007), pp. 2009–2015.
- [Bär+05] Bär, K. et al. “Loss of efferent vagal activity in acute schizophrenia”. In: *Journal of Psychiatric Research* 39.5 (2005), pp. 519–527.
- [Bar+18] Barnett, I. et al. “Relapse prediction in schizophrenia through digital phenotyping: a pilot study”. In: *Neuropsychopharmacology* 43.8 (2018), pp. 1660–1666.



- [Bau+04] Baumert, M. et al. “Forecasting of life threatening arrhythmias using the compression entropy of heart rate”. In: *Methods of Information in Medicine* 43.02 (2004), pp. 202–206.
- [Bau+05] Baumert, M. et al. “Estimating the complexity of heart rate fluctuations—an approach based on compression entropy”. In: *Fluctuation and Noise Letters* 5.04 (2005), pp. L557–L563.
- [Bir+19] Birnbaum, M. L. et al. “Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook”. In: *NPJ Schizophrenia* 5.1 (2019), pp. 1–9.
- [Bis94] Bishop, C. M. “Neural networks and their applications”. In: *Review of Scientific Instruments* 65.6 (1994), pp. 1803–1832.
- [BKH21] Błażkiewicz, M., Kędziorek, J., and Hadamus, A. “The Impact of Visual Input and Support Area Manipulation on Postural Control in Subjects after Osteoporotic Vertebral Fracture”. In: *Entropy* 23.3 (2021), p. 375.
- [Boe+06] Boettger, S. et al. “Altered diurnal autonomic variation and reduced vagal information flow in acute schizophrenia”. In: *Clinical Neurophysiology* 117.12 (2006), pp. 2715–2722.
- [Bon+17] Bone, D. et al. “Signal processing and machine learning for mental health research and clinical applications [perspectives]”. In: *IEEE Signal Processing Magazine* 34.5 (2017), pp. 196–195.
- [BPK01] Brennan, M., Palaniswami, M., and Kamen, P. “Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability?” In: *IEEE Transactions on Biomedical Engineering* 48.11 (2001), pp. 1342–1347.
- [Can+20] Can, Y. S. et al. “Personal stress-level clustering and decision-level smoothing to enhance the performance of ambulatory stress detection with smartwatches”. In: *IEEE Access* 8 (2020), pp. 38146–38163.
- [CFV05] Carney, R. M., Freedland, K. E., and Veith, R. C. “Depression, the autonomic nervous system, and coronary heart disease”. In: *Psychosomatic Medicine* 67 (2005), S29–S33.
- [CC19] Chalapathy, R. and Chawla, S. “Deep learning for anomaly detection: A survey”. In: *arXiv preprint arXiv:1901.03407* (2019).
- [CA02] Chambers, A. S. and Allen, J. J. “Vagal tone as an indicator of treatment response in major depression”. In: *Psychophysiology* 39.6 (2002), pp. 861–864.
- [Cha+17] Chapman, J. J. et al. “Quantification of free-living activity patterns using accelerometry in adults with mental illness”. In: *Scientific Reports* 7.1 (2017), pp. 1–12.
- [Cho+22] Choi, T. et al. “Multivariate Time-series Anomaly Detection using SeqVAE-CNN Hybrid Model”. In: *2022 International Conference on Information Networking (ICOIN)*. IEEE. 2022, pp. 250–253.
- [Chu+14] Chung, J. et al. “Empirical evaluation of gated recurrent neural networks on sequence modeling”. In: *arXiv preprint arXiv:1412.3555* (2014).
- [Cow+18] Cowton, J. et al. “A combined deep learning gru-autoencoder for the early detection of respiratory disease in pigs using multiple environmental sensors”. In: *Sensors* 18.8 (2018), p. 2521.
- [CC08] Cryer, J. D. and Chan, K.-S. *Time series analysis: with applications in R*. Vol. 2. Springer, 2008.

- 
- [DM21] Dalmeida, K. M. and Masala, G. L. “HRV features as viable physiological markers for stress detection using wearable devices”. In: *Sensors* 21.8 (2021), p. 2873.
- [Dek+05] Dekking, F. M. et al. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Vol. 488. Springer, 2005.
- [EBM95] Eaton, W. W., Badawi, M., and Melton, B. “Prodromes and precursors: epidemiologic data for primary prevention of disorders with slow onset.” In: *The American Journal of Psychiatry* (1995).
- [Ebr+18] Ebrahimzadeh, E. et al. “Prediction of paroxysmal Atrial Fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal”. In: *Computer Methods and Programs in Biomedicine* 165 (2018), pp. 53–67.
- [Eke+02] Eke, A. et al. “Fractal characterization of complexity in temporal physiological signals”. In: *Physiological Measurement* 23.1 (2002), R1.
- [Elu+21] Elujide, I. et al. “Application of deep and machine learning techniques for multi-label classification performance on psychotic disorder diseases”. In: *Informatomics in Medicine Unlocked* 23 (2021), p. 100545.
- [Fal04] Falconer, K. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.
- [Faw06] Fawcett, T. “An introduction to ROC analysis”. In: *Pattern Recognition Letters* 27.8 (2006), pp. 861–874.
- [Fer+20] Fernando, T. et al. “Deep Learning for Medical Anomaly Detection—A Survey”. In: *arXiv preprint arXiv:2012.02364* (2020).
- [Fil+20] Filntisis, P. P. et al. “Identifying differences in physical activity and autonomic function patterns between psychotic patients and controls over a long period of continuous monitoring using wearable sensors”. In: *arXiv preprint arXiv:2011.02285* (2020).
- [FRG04] Filzmoser, P., Reimann, C., and Garrett, R. *A multivariate outlier detection method*. Citeseer, 2004.
- [FT98] Friedman, B. H. and Thayer, J. F. “Autonomic balance revisited: panic anxiety and heart rate variability”. In: *Journal of Psychosomatic Research* 44.1 (1998), pp. 133–151.
- [Gae+93] Gaebel, W. et al. “Early neuroleptic intervention in schizophrenia: are prodromal symptoms valid predictors of relapse?” In: *The British Journal of Psychiatry* 163.S21 (1993), pp. 8–12.
- [Gar+21] Garoufis, C. et al. “An Unsupervised Learning Approach for Detecting Relapses from Spontaneous Speech in Patients with Psychosis”. In: *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, 2021, pp. 1–5.
- [Goo78] Good, I. *Fractals: Form, Chance, and Dimension*. 1978.
- [GBC16] Goodfellow, I., Bengio, Y., and Courville, A. *Deep learning*. MIT press, 2016.
- [Gra+15] Gravenhorst, F. et al. “Mobile phones as medical devices in mental disorder treatment: an overview”. In: *Personal and Ubiquitous Computing* 19.2 (2015), pp. 335–353.
- [Gu+17] Gu, J. et al. “An empirical study of language cnn for image captioning”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 1222–1231.
-

- [Had+21] Hadamus, A. et al. “Assessment of the Effectiveness of Rehabilitation after Total Knee Replacement Surgery Using Sample Entropy and Classical Measures of Body Balance”. In: *Entropy* 23.2 (2021), p. 164.
- [Haw80] Hawkins, D. M. *Identification of outliers*. Vol. 11. Springer, 1980.
- [Hay09] Hayes, M. H. *Statistical digital signal processing and modeling*. John Wiley & Sons, 2009.
- [Hen+10] Henry, B. L. et al. “Heart rate variability in bipolar mania and schizophrenia”. In: *Journal of Psychiatric Research* 44.3 (2010), pp. 168–176.
- [Hen+20] Henson, P. et al. “Towards clinically actionable digital phenotyping targets in schizophrenia”. In: *NPJ Schizophrenia* 6.1 (2020), pp. 1–7.
- [Hen+21] Henson, P. et al. “Anomaly detection to predict relapse risk in schizophrenia”. In: *Translational Psychiatry* 11.1 (2021), pp. 1–6.
- [Hig88] Higuchi, T. “Approach to an irregular time series on the basis of the fractal theory”. In: *Physica D: Nonlinear Phenomena* 31.2 (1988), pp. 277–283.
- [HS97] Hochreiter, S. and Schmidhuber, J. “Long short-term memory”. In: *Neural Computation* 9.8 (1997), pp. 1735–1780.
- [Hoy+06] Hoyer, D. et al. “Prognostic impact of autonomic information flow in multiple organ dysfunction syndrome patients”. In: *International Journal of Cardiology* 108.3 (2006), pp. 359–369.
- [Hui+09] Huikuri, H. V. et al. “Clinical impact of evaluation of cardiovascular control by novel methods of heart rate dynamics”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367.1892 (2009), pp. 1223–1238.
- [KY79] Kaplan, J. L. and Yorke, J. A. “Chaotic behavior of multidimensional difference equations”. In: *Functional Differential Equations and Approximation of Fixed Points*. Springer, 1979, pp. 204–227.
- [Kar+14] Karpathy, A. et al. “Large-scale video classification with convolutional neural networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 1725–1732.
- [KP09] Kasper, S. and Papadimitriou, G. N. *Schizophrenia*. Informa Healthcare, 2009.
- [KS16] Kesić, S. and Spasić, S. Z. “Application of Higuchi’s fractal dimension from basic to clinical neurophysiology: a review”. In: *Computer Methods and Programs in Biomedicine* 133 (2016), pp. 55–70.
- [KSB05] Kleiger, R. E., Stein, P. K., and Bigger Jr, J. T. “Heart rate variability: measurement and clinical utility”. In: *Annals of Noninvasive Electrocardiology* 10.1 (2005), pp. 88–101.
- [Kos19] Kostadinov, S. “Understanding encoder-decoder sequence to sequence model”. In: *Towards Data Science*, <https://towardsdatascience.com/understanding-encoder-decoder-sequence-to-sequence-model-679e04af4346> (2019).
- [Kou+12] Koutsouleris, N. et al. “Early recognition and disease prediction in the at-risk mental states for psychosis using neurocognitive pattern classification”. In: *Schizophrenia Bulletin* 38.6 (2012), pp. 1200–1215.
- [Kri+97] Krittayaphong, R. et al. “Heart rate variability in patients with coronary artery disease: differences in patients with higher and lower depression scores”. In: *Psychosomatic Medicine* 59.3 (1997), pp. 231–235.

- 
- [LBH15] LeCun, Y., Bengio, Y., and Hinton, G. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444.
- [Li+19] Li, S. et al. “Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting”. In: *Advances in Neural Information Processing Systems* 32 (2019), pp. 5243–5253.
- [Liu+19] Liu, G. et al. “Assessing the potential of longitudinal smartphone based cognitive assessment in schizophrenia: A naturalistic pilot study”. In: *Schizophrenia Research: Cognition* 17 (2019), p. 100144.
- [Liu+20] Liu, G.-D. et al. “A brief review of artificial intelligence applications and algorithms for psychiatric disorders”. In: *Engineering* 6.4 (2020), pp. 462–467.
- [Lon16] London, I. “Encoding cyclical continuous features-24-hour time”. In: *Ian London’s Blog* (2016).
- [LB09] Lopes, R. and Betrouni, N. “Fractal and multifractal analysis: a review”. In: *Medical Image Analysis* 13.4 (2009), pp. 634–649.
- [Mag+20] Maglogiannis, I. et al. “An Intelligent Cloud-Based Platform for Effective Monitoring of Patients with Psychotic Disorders”. In: *Proc. Int’l Conf. on Artificial Intelligence Applic. and Innovation*. 2020. ISBN: 9783030491857. DOI: [10.1007/978-3-030-49186-4\\_25](https://doi.org/10.1007/978-3-030-49186-4_25).
- [Mah36] Mahalanobis, P. C. “On the generalized distance in statistics”. In: National Institute of Science of India. 1936.
- [Mäk+02] Mäkikallio, T. H. et al. “Clinical applicability of heart rate variability analysis by methods based on nonlinear dynamics”. In: *Cardiac Electrophysiology Review* 6.3 (2002), pp. 250–255.
- [Mal+15] Malhotra, P. et al. “Long short term memory networks for anomaly detection in time series”. In: *Proceedings*. Vol. 89. 2015, pp. 89–94.
- [Mal+16] Malhotra, P. et al. “LSTM-based encoder-decoder for multi-sensor anomaly detection”. In: *arXiv preprint arXiv:1607.00148* (2016).
- [MM82] Mandelbrot, B. B. and Mandelbrot, B. B. *The fractal geometry of nature*. Vol. 1. WH freeman New York, 1982.
- [Mar94] Maragos, P. “Fractal signal analysis using mathematical morphology”. In: *Advances in Electronics and Electron Physics* 88 (1994), pp. 199–246.
- [Max+16] Maxhuni, A. et al. “Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients”. In: *Pervasive and Mobile Computing* 31 (2016), pp. 50–66.
- [McC+86] McCandless-Glimcher, L. et al. “Use of symptoms by schizophrenics to monitor and regulate their illness”. In: *Psychiatric Services* 37.9 (1986), pp. 929–933.
- [McG+14] McGorry, P. et al. “Biomarkers and clinical staging in psychiatry”. In: *World Psychiatry* 13.3 (2014), pp. 211–223.
- [Nwa+18] Nwankpa, C. et al. “Activation functions: Comparison of trends in practice and research for deep learning”. In: *arXiv preprint arXiv:1811.03378* (2018).
- [Oh+18] Oh, J. et al. “Prediction and early detection of delirium in the intensive care unit by using heart rate variability and machine learning”. In: *Physiological Measurement* 39.3 (2018), p. 035004.
- [Ola15] Olah, C. *Understanding lstm networks*. 2015.
- [Opp99] Oppenheim, A. V. *Discrete-time signal processing*. Pearson Education India, 1999.
-

- [Pan+21] Pang, G. et al. “Deep learning for anomaly detection: A review”. In: *ACM Computing Surveys (CSUR)* 54.2 (2021), pp. 1–38.
- [PS21] Park, H. and Son, J.-H. “Machine learning techniques for THz imaging and time-domain spectroscopy”. In: *Sensors* 21.4 (2021), p. 1186.
- [Pen+94] Peng, C.-K. et al. “Mosaic organization of DNA nucleotides”. In: *Physical Review* 49.2 (1994), p. 1685.
- [Pin91] Pincus, S. M. “Approximate entropy as a measure of system complexity.” In: *Proceedings of the National Academy of Sciences* 88.6 (1991), pp. 2297–2301.
- [PG07] Piskorski, J. and Guzik, P. “Geometry of the Poincaré plot of RR intervals and its asymmetry in healthy adults”. In: *Physiological Measurement* 28.3 (2007), p. 287.
- [Rab78] Rabiner, L. R. “Digital processing of speech signal”. In: *Digital Processing of Speech Signal* (1978).
- [RW17] Rawat, W. and Wang, Z. “Deep convolutional neural networks for image classification: A comprehensive review”. In: *Neural Computation* 29.9 (2017), pp. 2352–2449.
- [Ret+] Retsinas, G. et al. “Person Identification Using Deep Convolutional Neural Networks on Short-Term Signals from Wearable Sensors”. In: *Proc. Int’l Conf. on Acoustics, Speech and Signal Processing (ICASSP-2020)*.
- [RLM04] Richman, J. S., Lake, D. E., and Moorman, J. R. “Sample entropy”. In: *Methods in Enzymology* 384 (2004), pp. 172–184.
- [RM00] Richman, J. S. and Moorman, J. R. “Physiological time-series analysis using approximate entropy and sample entropy”. In: *American Journal of Physiology-Heart and Circulatory Physiology* (2000).
- [Sae+15] Saeb, S. et al. “Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study”. In: *Journal of Medical Internet Research* 17.7 (2015), e175.
- [Sat+18] Sato, D. et al. “A primitive study on unsupervised anomaly detection with an autoencoder in emergency head CT volumes”. In: *Medical Imaging 2018: Computer-Aided Diagnosis*. Vol. 10575. International Society for Optics and Photonics. 2018, 105751P.
- [Sca82] Scargle, J. D. “Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data”. In: *The Astrophysical Journal* 263 (1982), pp. 835–853.
- [SP97] Schuster, M. and Paliwal, K. K. “Bidirectional recurrent neural networks”. In: *IEEE Transactions on Signal Processing* 45.11 (1997), pp. 2673–2681.
- [SE98] Shi, Y. and Eberhart, R. “A modified particle swarm optimizer”. In: *1998 IEEE International Conference on Evolutionary Computation proceedings. IEEE World Congress on Computational intelligence (Cat. No. 98TH8360)*. IEEE. 1998, pp. 69–73.
- [Sin95] Singer, D. “Changes in heart rate variability associated with sudden cardiac death”. In: *In Heart Rate Variability* (1995), pp. 429–448.
- [Sri+14] Srivastava, N. et al. “Dropout: a simple way to prevent neural networks from overfitting”. In: *The Journal of Machine Learning Research* 15.1 (2014), pp. 1929–1958.

- 
- [SB18] Sutton, R. S. and Barto, A. G. *Reinforcement learning: An introduction*. MIT press, 2018.
- [TBC05] Todder, D., Bersudsky, Y., and Cohen, H. “Nonlinear analysis of RR interval in euthymic bipolar disorder”. In: *Autonomic Neuroscience* 117.2 (2005), pp. 127–131.
- [Tor+16] Torous, J. et al. “New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research”. In: *JMIR Mental Health* 3.2 (2016), e16.
- [Tsi+16] Tsinalis, O. et al. “Automatic sleep stage scoring with single-channel EEG using convolutional neural networks”. In: *arXiv preprint arXiv:1610.01683* (2016).
- [Uça+18] Uçar, M. K. et al. “Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques”. In: *Neural Computing and Applications* 29.8 (2018), pp. 1–16.
- [Vas+17] Vaswani, A. et al. “Attention is all you need”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 5998–6008.
- [Vos+06] Voss, A. et al. “Linear and nonlinear methods for analyses of cardiovascular variability in bipolar disorders”. In: *Bipolar Disorders* 8.5p1 (2006), pp. 441–452.
- [Wan+16] Wang, R. et al. “CrossCheck: toward passive sensing and detection of mental health changes in people with schizophrenia”. In: *Proc. Int’l Conf. on Pervasive and Ubiquitous Computing*. 2016, pp. 886–897.
- [Wan+20] Wang, W. et al. “Social sensing: assessing social functioning of patients living with schizophrenia using mobile phone sensing”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–15.
- [Wie+98] Wiersma, D. et al. “Natural course of schizophrenic disorders: a 15-year followup of a Dutch incidence cohort”. In: *Schizophrenia Bulletin* 24.1 (1998), pp. 75–85.
- [Wu+20] Wu, N. et al. “Deep transformer models for time series forecasting: The influenza prevalence case”. In: *arXiv preprint arXiv:2001.08317* (2020).
- [Yak+18] Yakura, H. et al. “Malware analysis of imaged binary samples by convolutional neural network with attention mechanism”. In: *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*. 2018, pp. 127–134.
- [YSA18] Yildirim, O., San Tan, R., and Acharya, U. R. “An efficient compression of ECG signals using deep convolutional autoencoders”. In: *Cognitive Systems Research* 52 (2018), pp. 198–211.
- [Yin19] Ying, X. “An overview of overfitting and its solutions”. In: *Journal of Physics: Conference Series*. Vol. 1168. 2. IOP Publishing. 2019, p. 022022.
- [Yil+18] Yildirim, Ö. et al. “Arrhythmia detection using deep convolutional neural network with long duration ECG signals”. In: *Computers in Biology and Medicine* 102 (2018), pp. 411–420.
- [ZYX17] Zhang, S., Yao, L., and Xu, X. “Autosvd++ an efficient hybrid collaborative filtering model via contractive auto-encoders”. In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2017, pp. 957–960.
-