



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Τεχνικές βαθιάς μάθησης για την αναγνώριση των επιπέδων προσήλωσης
σε πραγματικό χρόνο κατά την αλληλεπίδραση με παιχνίδι σοβαρού σκοπού
για την υγεία**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Πάλμος Στέφανος

Επιβλέπουσα:

Κωνσταντίνα Σ. Νικήτα

Καθηγήτρια Ε.Μ.Π.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Τεχνικές βαθιάς μάθησης για την αναγνώριση των επιπέδων προσήλωσης
σε πραγματικό χρόνο κατά την αλληλεπίδραση με παιχνίδι σοβαρού σκοπού
για την υγεία**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Πάλμος Στέφανος

Επιβλέπουσα: Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή κριτική επιτροπή την 14^η Οκτωβρίου 2024.

Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια Ε.Μ.Π.

Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.

Αθανάσιος Βουλόδημος
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2024

Πάλμος Στέφανος

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

Copyright © Πάλμος Στέφανος

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή και διανομή της παρούσας εργασίας εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτική 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 ερμηνευθούν ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η παρούσα Διπλωματική Εργασία πραγματεύεται την αναγνώριση των επιπέδων προσήλωσης παικτών κατά την αλληλεπίδραση τους με παιχνίδια σοβαρού σκοπού για την υγεία (ΠΣΣΥ) από καταγραφές των συνεδριών τους και από υποκειμενικές αναφορές που παρείχαν. Συγκεκριμένα, στόχος της είναι η χρήση μοντέλων βαθιάς μηχανικής μάθησης που αξιοποιούν τα δεδομένα καταγραφής οθόνης για την πρόβλεψη των επιπέδων προσήλωσης των χρηστών και τη δυναμική προσαρμογή του ΠΣΣΥ.

Για τον σκοπό αυτό χρησιμοποιούνται δεδομένα από την αλληλεπίδραση των συμμετεχόντων με ένα ΠΣΣΥ με θέμα τις διατροφικές συνήθειες. Αρχικά πραγματοποιείται ανάλυση και επεξεργασία στα διαθέσιμα δεδομένα. Έπειτα, τα δεδομένα χρησιμοποιούνται για την εκπαίδευση μοντέλων βαθιάς μηχανικής μάθησης, και συλλέγονται μετρικές, ώστε να αξιολογηθεί η ικανότητά τους να προβλέπουν τα επίπεδα της προσήλωσης των παικτών.

Με βάση τα αποτελέσματα που εξάγονται για την εκπαίδευση των μοντέλων στα δεδομένα καταγραφής οθόνης όλων των συμμετεχόντων με τη μεθοδολογία του αλγορίθμου "διασταυρούμενη επικύρωση με εξαγωγή ενός δείγματος" ("leave-one-out-cross-validation") και μίας παραλλαγής του, προκύπτει ότι τα μοντέλα βαθιάς μηχανικής μάθησης που αξιοποιούνται δεν επιτυγχάνουν ικανοποιητικά επίπεδα ακρίβειας στην πρόβλεψη των επιπέδων προσήλωσης. Για να διερευνηθεί περαιτέρω αυτό το ζήτημα, εκπαιδεύονται ξεχωριστά μοντέλα για τον κάθε παίκτη. Σε αυτή την περίπτωση παρατηρείται βελτίωση στην ακρίβεια των αποτελεσμάτων. Παρ' όλα αυτά, η συνολικά καλύτερη προβλεπτική ικανότητα δεν είναι καθολική για όλους τους συμμετέχοντες και υποδεικνύει την ανάγκη περαιτέρω έρευνας για την εξέταση της ποιότητας των δεδομένων καταγραφής της προσήλωσης αλλά και στις μεθόδους ανάλυσης τους, με σκοπό τη βελτίωση των αποτελεσμάτων.

Λέξεις Κλειδιά: παιχνίδια σοβαρού σκοπού, συναισθηματική υπολογιστική, επίπεδα προσήλωσης, καταγραφές ροών/βίντεο, βαθιά μηχανική μάθηση

Abstract

The present Diploma Thesis investigates the recognition of player engagement during interaction with serious games for health (SGH), based on session recordings and self-reports. Specifically, it aims to use deep learning models that leverage screen recording data to predict user engagement levels and dynamically adapt the SGH experience.

The study utilizes data from participant interactions with an SGH focused on dietary habits. Initially, the available data is analysed and processed. These processed data are then used to train deep learning models, from which metrics are gathered to evaluate the models' ability to predict players' engagement levels.

Based on the results of training the models using screen recording data from all participants with the "leave-one-out-cross-validation" method and a variation of it, it is evident that deep learning models do not achieve satisfactory accuracy in predicting engagement levels. To further investigate this issue, individual models are trained for each player. In this case, an improvement in prediction accuracy is observed. However, this improvement is not uniform across all participants, highlighting the need for further research to enhance the quality of engagement data recordings and analysis methods, with the goal of improving future outcomes.

Keywords: serious games, affective computing, engagement levels, video recordings, deep learning

Στη Μαργαρίτα και τους γονείς μου...

Πίνακας περιεχομένων

Περίληψη	5
Abstract.....	7
Ευρετήριο Εικόνων	13
Ευρετήριο Πινάκων	14
Κεφάλαιο 1: Εισαγωγή	15
1.1 Εισαγωγή στο αντικείμενο της Διπλωματικής Εργασίας.....	15
1.2 Στόχος της Διπλωματικής Εργασίας.....	15
1.3 Οργάνωση της Διπλωματικής Εργασίας.....	16
Κεφάλαιο 2: Θεωρητικό Υπόβαθρο.....	17
2.1 Παιχνίδια Σοβαρού Σκοπού για την Υγεία (ΠΣΣΥ).....	17
2.1.1 Παιχνίδια Σοβαρού Σκοπού (ΠΣΣ).....	17
2.1.2 ΠΣΣ στον τομέα της υγείας.....	18
2.1.3 Αξιολόγηση και Επικύρωση των ΠΣΣΥ	20
2.1.4 Μελλοντικές προκλήσεις των ΠΣΣΥ	22
2.2 Συναισθηματική Υπολογιστική Νοημοσύνη (ΣΥΝ).....	23
2.2.1 Ανθρώπινα συναισθήματα και η συσχέτιση τους με την Τεχνητή Νοημοσύνη	23
2.2.2 Πεδία Εφαρμογών ΣΥΝ.....	25
2.2.3 Ενσωμάτωση της Τεχνητής Νοημοσύνης με τα ΠΣΣ	26
2.2.4 Η ΣΥΝ στα ΠΣΣ	27
2.3 Δεδομένα Καταγραφής Βίντεο.....	28
2.3.1 Ανάλυση δεδομένων βίντεο.....	28
2.3.2 Εφαρμογές με χρήση δεδομένων από βίντεο	30
2.3.3 Προκλήσεις και μελλοντικές κατευθύνσεις	31
Κεφάλαιο 3: Μεθοδολογία	32
3.1 Περιγραφή της Διαδικασίας.....	32
3.1.1 Το ΠΣΣΥ που χρησιμοποιήθηκε	32
3.1.2 Καταγραφή επιπέδων προσήλωσης	32
3.1.3 Οι συμμετέχοντες.....	33
3.2 Ανάλυση και Προεπεξεργασία των Δεδομένων	33
3.2.1 Δεδομένα καταγραφών βίντεο	33
3.2.2 Δεδομένα προσήλωσης	35
3.2.3 Συσχέτιση των δεδομένων καταγραφής και προσήλωσης.....	38
3.3 Μέθοδοι Βαθιάς Μηχανικής Μάθησης (DL).....	39

3.3.1 Ανάλυση αλγορίθμων DL που χρησιμοποιήθηκαν	39
3.3.2 Το μοντέλο που χρησιμοποιήθηκε	41
3.3.3 Τεχνική Leave-One-Out Cross Validation (LOOCV)	42
3.3.4 Μετρικές	43
Κεφάλαιο 4: Αποτελέσματα.....	45
4.1 Απλό LOOCV χωρίς τους συμμετέχοντες με τις προβληματικές μετρήσεις	45
4.2 Παραλλαγή του LOOCV χωρίς τους Συμμετέχοντες με τις προβληματικές μετρήσεις	47
4.3 Εκπαίδευση κάθε Συμμετέχοντα ξεχωριστά	50
Κεφάλαιο 5: Συζήτηση.....	55
5.1 Στατιστική Ανάλυση στην εκπαίδευση των μοντέλων κάθε συμμετέχοντα ξεχωριστά	55
5.2 Προβλήματα που παρουσιάστηκαν κατά την πειραματική διαδικασία	56
5.3 Τρόποι επίλυσης των προβλημάτων	56
Κεφάλαιο 6: Επίλογος.....	58
6.1 Σύνοψη.....	58
6.2 Όρια και Περιορισμοί	58
Κεφάλαιο 7: Βιβλιογραφία	60
Παράρτημα Εικόνων.....	63

Ευρετήριο Εικόνων

Εικόνα 1: Αναπαράσταση των ΠΣΣ στο φάσμα των σκοπών των παιχνιδιών	18
Εικόνα 2: Στιγμιότυπα από το ΠΣΣΥ “Express Cooking Train”	19
Εικόνα 3: Στιγμιότυπο από το ΠΣΣΥ “Wake Up for the Future!”	20
Εικόνα 4: Κάποια από τα χρησιμοποιούμενα εργαλεία αξιολόγησης και επικύρωσης σε ΠΣΣ [17] ..	22
Εικόνα 5: Τομείς ΠΣΣΥ	23
Εικόνα 6: Καταγραφή μορφασμών προσώπου (ΣΥΝ).....	24
Εικόνα 7: Επεξεργασία βίντεο δεδομένων.....	29
Εικόνα 8: “Express Cooking Train”	32
Εικόνα 9: Καρέ του ΠΣΣΥ με και χωρίς μαύρες μπάρες αντίστοιχα.....	34
Εικόνα 10: Διάγραμμα κατανομής της μετρικής της προσήλωσης για τους πρώτους 12 συμμετέχοντες	37
Εικόνα 11: Διάγραμμα κατανομής της μετρικής της προσήλωσης για τους επόμενους 12 συμμετέχοντες.....	37
Εικόνα 12: Διάγραμμα κατανομής της μετρικής της προσήλωσης για τον τελευταίο συμμετέχοντα .	37
Εικόνα 13: Παράδειγμα σχήματος δεδομένων βίντεο (σειριακά δεδομένα).....	39
Εικόνα 14: Παράδειγμα αρχιτεκτονικής CNN.....	40
Εικόνα 15: Παράδειγμα αρχιτεκτονικής RNN.....	40
Εικόνα 16: Παράδειγμα αρχιτεκτονικής LSTM	41
Εικόνα 17: Οπτική αναπαράσταση της αρχιτεκτονικής που χρησιμοποιήθηκε (τυχαίες τιμές)	42
Εικόνα 18: Παράδειγμα λειτουργίας βαθμίδας dropout.....	42
Εικόνα 19: Αρχιτεκτονική LOOCV	43
Εικόνα 20: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 18, 6	52
Εικόνα 21: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 10, 15	52
Εικόνα 22: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 1, 7	54
Εικόνα 23: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων:21, 22	54
Εικόνα 24: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 1, 2	63
Εικόνα 25: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 7, 9	63
Εικόνα 26: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 11, 12	63
Εικόνα 27: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 14, 17	64
Εικόνα 28: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 21, 5	64
Εικόνα 29: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 22, 25	64
Εικόνα 30: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 3, 4	65
Εικόνα 31: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 8, 13	65
Εικόνα 32: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 16, 19	65
Εικόνα 33: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 20, 23	66
Εικόνα 34: Confusion Matrix για το ατομικό μοντέλο (9) του συμμετέχοντα: 24	66
Εικόνα 35: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 2, 9	66
Εικόνα 36: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 3, 4	68
Εικόνα 37: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 5, 6	68
Εικόνα 38: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 8, 10.....	68
Εικόνα 39: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 13, 15	69
Εικόνα 40: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 16, 19	69
Εικόνα 41: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 20, 23	69
Εικόνα 42: Confusion Matrix για το ατομικό μοντέλο (10) του συμμετέχοντα: 24	70

Ευρετήριο Πινάκων

Πίνακας 1: Κωδικοποίηση Συμμετεχόντων	33
Πίνακας 2: Μορφή Δεδομένων Προσήλωσης των Συμμετεχόντων	35
Πίνακας 3: Κατανομή του μέσου όρου της προσήλωσης των επικαλυπτόμενων παραθύρων καρέ (μέση τιμή πριν - μέση τιμή μετά) $\geq 0 \rightarrow 0$ (μέση τιμή πριν - μέση τιμή μετά) $< 0 \rightarrow 1$	36
Πίνακας 4: Οι συμμετέχοντες που αφαιρέθηκαν από ένα μερίδιο των πειραμάτων και οι λόγοι για τον καθένα από αυτούς	38
Πίνακας 5: Απλό LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 1: 2 x Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), Optimizer: ADAM (lr=3e-4)	46
Πίνακας 6: Απλό LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 2: 2 x Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(256, relu), Dense(256, relu), Dense(1, sigmoid), Optimizer: ADAM (lr=3e-4)	46
Πίνακας 7: Απλό LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 3: 2 x Conv2D(32, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), Optimizer: ADAM (lr=3e-4)	47
Πίνακας 8: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 4: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 30, patience 5)	48
Πίνακας 9: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 5: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 30, patience 10)	48
Πίνακας 10: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 6: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(64, relu), Dense(64, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 30, patience 5)	49
Πίνακας 11: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 7: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer SGD(learning_rate=3e-4, momentum=0.9, nesterov=True), Early Stopping (from Epoch 60, patience 20)	49
Πίνακας 12: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 8: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=1e-4), Early Stopping (from Epoch 60, patience 20)	50
Πίνακας 13: Ατομικό μοντέλο κάθε συμμετέχοντα: Όλες οι τιμές μετρικών για το Μοντέλο 9: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 100, patience 20)	51
Πίνακας 14: Ατομικό μοντέλο κάθε συμμετέχοντα: Όλες οι τιμές μετρικών για το Μοντέλο 10: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(256, relu), Dense(256, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 100, patience 20)	53

Κεφάλαιο 1: Εισαγωγή

1.1 Εισαγωγή στο αντικείμενο της Διπλωματικής Εργασίας

Η επιστήμη της υγείας υιοθετεί έναν πιο ανθρωποκεντρικό προσανατολισμό, εστιάζοντας στην πρόγνωση, την πρόληψη, καθώς και στην εξατομικευμένη φροντίδα των ασθενών, γνώστη και ως ιατρική P4 [1]. Τέτοιες σύγχρονες τεχνολογίες επικεντρώνονται στην μοναδικότητα του κάθε ατόμου και έτσι η συνεισφορά που παρέχεται είναι διπλού σκοπού. Αρχικά, οι ειδικοί εκπαιδεύονται περεταίρω με βασικό γνώμονα αντιμετώπισης περιστατικών με βάση την ατομικότητα των ασθενών. Επιπλέον, η ποιότητα ζωής των ασθενών βελτιώνεται μέσω της υιοθέτησης υγιεινότερων καθημερινών συνηθειών, καθώς αυτές οι τεχνολογίες δεν εστιάζουν αποκλειστικά σε παραδοσιακούς τρόπους αντιμετώπισης.

Για να μπορέσει να επιτευχθεί η αποκατάσταση χρόνιων ασθενειών και παθήσεων, η χρήση καινοτόμων τεχνολογιών αποτελεί, επίσης, σημαντικό βήμα στην εξέλιξη του τομέα της υγείας. Αυτές οι τεχνολογικές εξελίξεις περιλαμβάνουν τη χρήση έξυπνων συσκευών τηλεϊατρικής, την παρακολούθηση ασθενειών, τους ψηφιακούς φακέλους ασθενών, καθώς και τη συλλογή και διαχείριση πληροφοριών υγείας [2]. Παράλληλα, μπορούν να συνδυαστούν με τα ΠΣΣΥ και άλλες τεχνολογίες αιχμής, όπως τη συναισθηματική υπολογιστική νοημοσύνη (ΣΥΝ), η οποία ενώνει και τις επιστήμες των υπολογιστών με σκοπό τη συνολική βελτίωση της υγειονομικής φροντίδας [3]. Στην παρούσα Διπλωματική Εργασία χρησιμοποιούνται καινοτόμες τεχνολογίες και εργαλεία, όπως η παρακολούθηση και ανάλυση επιπέδων προσήλωσης μέσω ΣΥΝ, η χρήση ΠΣΣΥ, και οι καταγραφές του ΠΣΣΥ των συμμετεχόντων. Η διερεύνηση της σχέσης μεταξύ των καταγραφών και των επιπέδων προσήλωσης πραγματοποιείται μέσω της εκπαίδευσης βαθιών μοντέλων μηχανικής μάθησης.

Η παρούσα Διπλωματική Εργασία, καθώς και η συλλογή των δεδομένων που χρησιμοποιήθηκαν σε αυτή την έρευνα, πραγματοποιήθηκαν στο Εθνικό Μετσόβιο Πολυτεχνείο, στο Εργαστήριο Βιοϊατρικών Προσομοιώσεων και Απεικονιστικής Τεχνολογίας (BioSim Laboratory). Ιδιαίτερες ευχαριστίες οφείλω στον Διδάκτορα Κωνσταντίνο Μήτση και την Καθηγήτρια Κωνσταντίνα Νικήτα για την πολύτιμη καθοδήγηση και υποστήριξή τους.

1.2 Στόχος της Διπλωματικής Εργασίας

Η παρούσα Διπλωματική Εργασία έχει ως στόχο την κατανόηση και ανάλυση των δεδομένων των καταγραφών που προέρχονται από ΠΣΣΥ, καθώς και των δεδομένων προσήλωσης των συμμετεχόντων μέσω της υποκειμενικής αξιολόγησής τους. Ταυτόχρονα, διεξάγεται βιβλιογραφική μελέτη πάνω σε έννοιες που αφορούν στον τομέα της υγείας τα ΠΣΣΥ, τη ΣΥΝ, την Τεχνητή Νοημοσύνη και ακόμα τη γενικότερη χρήση καταγραφών στο χώρο της Μηχανικής Μάθησης. Ένας από τους βασικούς στόχους της συγκεκριμένης έρευνας είναι η εκπαίδευση αλγορίθμων βαθιάς μηχανικής μάθησης και η αναζήτηση τρόπων βελτίωσης των μετρικών τους για την επίτευξη αξιόπιστων αποτελεσμάτων. Ιδιαίτερο ενδιαφέρον παρουσιάζει η ερμηνεία των δεδομένων που προκύπτουν από τα μοντέλα, με σκοπό να προσδιοριστεί μία αρχή που συνδέει τις καταγραφές του ΠΣΣΥ με τα επίπεδα προσήλωσης των ατόμων.

1.3 Οργάνωση της Διπλωματικής Εργασίας

Η τρέχουσα Διπλωματική Εργασία χωρίζεται σε τέσσερα βασικά κεφάλαια. Αναλυτικότερα:

Στο Κεφάλαιο 2 πραγματοποιείται ανάλυση του θεωρητικού υπόβαθρου, το οποίο περιλαμβάνει βασικές έννοιες σχετικές με τα Παιχνίδια Σοβαρού Σκοπού για την Υγεία (ΠΣΣΥ), τα Παιχνίδια Σοβαρού Σκοπού (ΠΣΣ), καθώς και την Τεχνητή Νοημοσύνη. Η έμφαση δίνεται στη συμβολή αυτών των τεχνολογιών σε παρεμβάσεις που αφορούν την υγεία. Επιπλέον, εξετάζεται η χρησιμότητα των δεδομένων καταγραφής βίντεο, τα οποία χρησιμοποιούνται ως εργαλεία εκπαίδευσης και αξιολόγησης της απόδοσης.

Στο Κεφάλαιο 3 εισάγεται η μεθοδολογία που ακολουθήθηκε στο πρακτικό μέρος της διπλωματικής εργασίας. Αναλύεται λεπτομερώς η διαδικασία επεξεργασίας, ανάλυσης και προετοιμασίας των δεδομένων καταγραφής βίντεο, όπως και η επισημείωση των δεδομένων που αφορούν τα επίπεδα προσήλωσης των συμμετεχόντων. Κεντρικό στοιχείο της μεθοδολογίας αποτελεί η χρήση μεθόδων Βαθιάς Μηχανικής Μάθησης, όπου παρουσιάζονται οι αλγόριθμοι που χρησιμοποιήθηκαν και εξάγονται οι αντίστοιχες μετρικές.

Στο Κεφάλαιο 4 παρουσιάζονται τα τελικά αποτελέσματα της έρευνας μέσω πινάκων και γραφημάτων. Τα αποτελέσματα αποτυπώνουν τη συνολική απόδοση των αλγορίθμων, καθώς και την αξιολόγηση των επιπέδων προσήλωσης των συμμετεχόντων κατά τη διάρκεια των καταγραφών βίντεο.

Το Κεφάλαιο 5 επικεντρώνεται στη συζήτηση των ευρημάτων της έρευνας, όπου γίνεται μια σύνοψη των αποτελεσμάτων και ανάλυση των δεδομένων που συλλέχθηκαν κατά την πειραματική διαδικασία. Εξετάζονται επίσης τα προβλήματα που προέκυψαν κατά τη διεξαγωγή των πειραμάτων και προτείνονται πιθανοί τρόποι επίλυσής τους.

Τέλος, στο Κεφάλαιο 6, παρουσιάζονται οι μελλοντικές προκλήσεις που αναδύονται από την παρούσα έρευνα, ενώ αναφέρονται και οι περιορισμοί που εντοπίστηκαν κατά την εκπόνησή της. Το τελευταίο Κεφάλαιο 7 περιλαμβάνει τη βιβλιογραφία που μελετήθηκε και χρησιμοποιήθηκε κατά τη διάρκεια της έρευνας, παρέχοντας μια πλήρη επισκόπηση των πηγών που συνέβαλαν στην ανάπτυξή της.

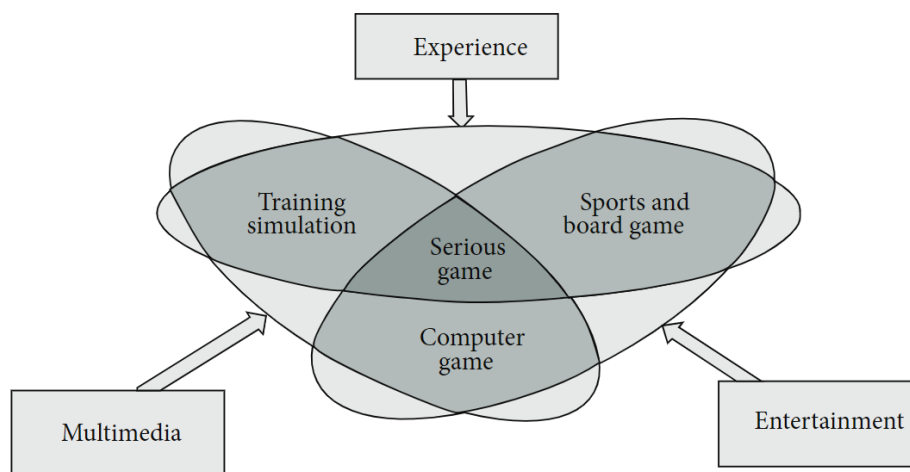
Κεφάλαιο 2: Θεωρητικό Υπόβαθρο

2.1 Παιχνίδια Σοβαρού Σκοπού για την Υγεία (ΠΣΣΥ)

2.1.1 Παιχνίδια Σοβαρού Σκοπού (ΠΣΣ)

Σήμερα, ο όρος ΠΣΣ γίνεται όλο και πιο δημοφιλής, λόγω της διάκρισης τους από τα απλά βιντεοπαιχνίδια στον σχεδιαστικό τους στόχο [4]. Τα παιχνίδια αυτά δημιουργούνται για σκοπούς πέρα από την ψυχαγωγία, όπως είναι η επιστημονική έρευνα, η αντιμετώπιση καταστάσεων έκτακτης ανάγκης, η εκπαίδευση, η ευαισθητοποίηση αλλά και οι βελτιώσεις σε διάφορους τομείς του τρόπου ζωής. Τα ΠΣΣ ανήκουν σε μια ιδιαίτερη κατηγορία παιχνιδιών που έχουν αφήγηση βασισμένη σε μοτίβα υψηλής ποιότητας και ενσωματώνουν στοχαστικά στοιχεία ενώ η ανάπτυξη τους εστιάζει στην προσομοίωση και στην ενσωμάτωση του χρήστη σε καθημερινά σενάρια, προβάλλοντας την εκπαιδευτική αξία της διασκέδασης και του ανταγωνισμού. Η φιλοσοφία πίσω από αυτά τα παιχνίδια είναι ότι μπορούν να προσφέρουν πληροφορίες σε πολλαπλά επίπεδα, προσαρμοσμένα στον κάθε χρήστη και με αυτόν τον τρόπο αποκτούν έναν πιο ανθρώπινο χαρακτήρα, προσαρμόζοντας τις δυσκολίες και τις δυνατότητες σύμφωνα με τις ανάγκες του παίκτη [5],[6]. Επιπλέον, τα παιχνίδια αυτά προσαρμόζουν τη δυσκολία τους με βάση τις αντιδράσεις του παίκτη, επιτρέποντας την κλιμάκωση των επιπέδων με ομαλή μετάβαση από το ένα στο άλλο. Αυτή η προσέγγιση βελτιστοποιεί την εμπειρία του χρήστη, αυξάνοντας την απόδοσή του και την ενασχόληση με το παιχνίδι. Όταν ένα ΠΣΣ έχει σχεδιαστεί με προσοχή στις λεπτομέρειες και στους μηχανισμούς αλληλεπίδρασης, δημιουργεί ένα περιβάλλον όπου ο χρήστης μπορεί εύκολα να εξοικειωθεί, γεγονός που αυξάνει τα επίπεδα της προσήλωσης και της αποδοτικότητάς του.

Τα ΠΣΣ μπορούν να συνδυαστούν και να ομαδοποιηθούν σε μερικές βασικές κατηγορίες, οι οποίες αντανακλούν τους πολλαπλούς τομείς εφαρμογής. Το ευρύ φάσμα των εφαρμογών αυτών εκτείνεται στην εκπαίδευση, την υγεία, τον πολιτισμό καθώς και τη δημιουργικότητα. Χρησιμοποιούνται για την επιμόρφωση ενηλίκων και νέων, από την επαγγελματική κατάρτιση σε τομείς όπως η στρατιωτική εκπαίδευση, μέχρι την ανάπτυξη γνωστικών δεξιοτήτων και την πρόσληψη προσωπικού. Παράλληλα, προσφέρουν σημαντικές λύσεις στην προαγωγή της υγείας και στη θεραπεία, μέσω παιχνιδιών που στοχεύουν στη σωματική άσκηση, τη ψυχική υγεία και την αποκατάσταση ασθενειών. Επιπλέον, συμβάλλουν στην κοινωνική ευαισθητοποίηση και στην κατανόηση πολιτικών ή πολιτιστικών θεμάτων, ενώ παράλληλα προωθούν την καλλιτεχνική έκφραση και τη δημιουργικότητα, βοηθώντας στη σχεδίαση προϊόντων και τη διάδοση μηνυμάτων μέσω καινοτόμων μορφών σχεδίασης των παιχνιδιών [7].



Εικόνα 1: Αναπαράσταση των ΠΣΣ στο φάσμα των σκοπών των παιχνιδιών

2.1.2 ΠΣΣ στον τομέα της υγείας

Η εισαγωγή των ΠΣΣ στον κλάδο της υγείας (ΠΣΣΥ) είναι ιδιαίτερα αξιοσημείωτη επειδή είναι ένα νέο εκπαιδευτικό εργαλείο με μια σειρά από οφέλη. Η διαδραστική μάθηση μέσω παιχνιδιών είναι πιο ευχάριστη και πιο προσιτή, ακόμη και για τους επαγγελματίες υγείας συγκριτικά με τις παραδοσιακές μεθόδους μάθησης. Αυτό προκύπτει διότι μπορούν ταυτόχρονα να μετριάζουν την πιθανότητα να γίνουν λάθη σε περιπτώσεις διάγνωσης, αλλά και να βελτιώνουν την πρακτική εμπειρία τους σε διάφορες πραγματικές καταστάσεις [5].

Τα ΠΣΣΥ μπορούν να κατηγοριοποιηθούν βάσει του κοινού στο οποίο απευθύνονται, είτε αυτό αφορά ασθενείς είτε υγιείς χρήστες. Για τους ασθενείς, τα παιχνίδια αυτά αξιοποιούνται για την παρακολούθηση της υγείας τους, επιτρέποντας τη συλλογή βιοσημάτων και τη διαρκή εποπτεία της κατάστασής τους. Παράλληλα, ορισμένα από αυτά τα παιχνίδια συμβάλλουν στην ανίχνευση και καταγραφή ασυνήθιστων συμπτωμάτων, προσφέροντας μια καινοτόμο προσέγγιση στη διάγνωση. Σε άλλα περιβάλλοντα, λειτουργούν ως θεραπευτικά εργαλεία, βοηθώντας στην αποκατάσταση συγκεκριμένων υγειονομικών προβλημάτων ή στην υποστήριξη των ασθενών σε κοινωνικές και διαγνωστικές δεξιότητες. Επιπλέον, σημαντικός είναι ο ρόλος των ΠΣΣΥ στην αποκατάσταση μετά από ασθένειες ή τραυματισμούς, συμβάλλοντας στην επαναφορά βασικών λειτουργικών ικανοτήτων των ατόμων αυτών. Ακόμη, συνεισφέρουν και ως εκπαιδευτικά εργαλεία για τους ασθενείς σχετικά με την αυτοδιαχείριση των προβλημάτων τους, προάγοντας την κατανόηση και την εξοικείωση με τις καταστάσεις υγείας τους.

Για τους μη ασθενείς, τα ΠΣΣΥ διαδραματίζουν έναν εξίσου σημαντικό ρόλο στην προαγωγή της υγείας και της ευεξίας, ενθαρρύνοντας τη βελτίωση του τρόπου ζωής μέσω της άσκησης, του ύπνου και της ισορροπημένης διατροφής. Επιπλέον, ως παιχνίδια προσομοίωσης μπορούν να αποτελέσουν ένα πολύτιμο εργαλείο για την εκπαίδευση των επαγγελματιών υγείας, προσφέροντάς τους τη δυνατότητα να εξασκηθούν σε κλινικά σενάρια και καταστάσεις έκτακτης ανάγκης. Τα ΠΣΣΥ μπορούν να συμβάλουν και για το ευρύ κοινό παρέχοντας σημαντικές γνώσεις σχετικά με την προσωπική υγιεινή και τη σωστή διατροφή, ενισχύοντας

την προσωπική φροντίδα αλλά και τη διατήρηση της συνολικότερης υγείας του κάθε ατόμου [5].

Στην κατηγορία με τα ΠΣΣΥ για τους μη ασθενείς ανήκουν ακόμη δύο παιχνίδια που έχουν αναπτυχθεί στα πλαίσια του Εργαστηρίου Βιοϊατρικών Προσομοιώσεων και Απεικονιστικής Τεχνολογίας του Εθνικού Μετσόβιου Πολυτεχνείου (BioSim, NTUA), στο οποίο εκπονείται και η παρούσα διπλωματική εργασία. Το πρώτο λέγεται “Express Cooking Train”, το οποίο λαμβάνει χώρα σε ένα μετα-αποκαλυπτικό περιβάλλον, όπου ο χρήστης έχει τον έλεγχο ενός τρένου που ταξιδεύει σε περιοχές γεμάτες τέρατα. Αυτά τα τέρατα τρέφονται κυρίως με ανθυγιεινά και ενεργειακά πλούσια επεξεργασμένα τρόφιμα, γεγονός που καθιστά την πρόκληση για τον χρήστη ιδιαίτερα δύσκολη καθώς ο στόχος του είναι να διασφαλίσει ότι το τρένο θα φτάσει στον προορισμό του, μαγειρεύοντας υγιεινά γεύματα που μοιάζουν οπτικά με ανθυγιεινά. Μέσω αυτής της διαδικασίας, το παιχνίδι προωθεί την εκπαίδευση των παικτών σε υγιεινές διατροφικές επιλογές με δημιουργικό τρόπο. Επιπλέον συνδυάζει στοιχεία από τρία είδη παιχνιδιών, τα οποία είναι μαγειρικής, roguelike και παζλ. Ο χρήστης ξεκινά με βασικές συνταγές και, με την πάροδο του χρόνου, εμπλουτίζει το μενού του μέσω πειραματισμών με νέες τεχνικές και υλικά, προσομοιώνοντας την προετοιμασία φαγητού. Το στοιχείο roguelike αναδεικνύεται με τη δημιουργία τυχαίων σταδίων κάθε φορά που παίζει ο παίκτης, ενώ όταν δεν μπορεί να ανταποκριθεί στις απαιτήσεις συνεπάγεται στην επανεκκίνηση του παιχνιδιού, διατηρώντας όμως τις συνταγές και τις αναβαθμίσεις που έχει κατακτήσει έτσι ώστε να υπάρχει πρόοδος. Το στοιχείο του παζλ βασίζεται στη στρατηγική ρίψη του μαγειρεμένου γεύματος προς τα τέρατα μέσω ενός μηχανισμού που προσομοιώνει την εκτόξευση ενός καταπέλτη, δημιουργώντας έτσι ένα παιχνίδι με στοιχεία φυσικής για τη λύση προβλημάτων [8].



Εικόνα 2: Στιγμιότυπα από το ΠΣΣΥ “Express Cooking Train”

Το δεύτερο ΠΣΣΥ είναι το “Wake Up for the Future!", το οποίο έχει ως στόχο την ευαισθητοποίηση και την προώθηση της αυτοδιαχείρισης της αποφρακτικής άπνοιας ύπνου, μιας από τις πιο κοινές διαταραχές αναπνοής σχετικές με τον ύπνο. Η αποφρακτική άπνοια ύπνου χαρακτηρίζεται από επαναλαμβανόμενα επεισόδια κατάρρευσης των ανώτερων αεραγωγών, που προκαλούν μείωση ή διακοπή της ροής του αέρα για τουλάχιστον 10 δευτερόλεπτα. Ενώ υπάρχουν διάφορες επιλογές θεραπείας, ορισμένες αλλαγές στη συμπεριφορά, όπως η απώλεια βάρους, η υγιεινή διατροφή, η περιορισμένη κατανάλωση αλκοόλ, ο περιορισμός του καπνίσματος, καθώς και η τήρηση μίας κατάλληλης ρουτίνας γύρω από τον ύπνο, μπορούν να ωφελήσουν σημαντικά τους ασθενείς. Περιλαμβάνει μια καινοτόμο και αγωνιώδη πλοκή, όπου ο παίκτης ταξιδεύει από ένα δυστοπικό μέλλον, όπου έχει χαθεί κάθε γνώση σχετικά με την αποφρακτική άπνοια ύπνου, στο παρόν. Ο σκοπός του παιχνιδιού είναι η ευαισθητοποίηση του παίκτη, που βασίζεται σε ένα εννοιολογικό πλαίσιο με τη χρήση των μηχανισμών του [9].



Εικόνα 3: Στιγμιότυπο από το ΠΣΣΥ “Wake Up for the Future!”

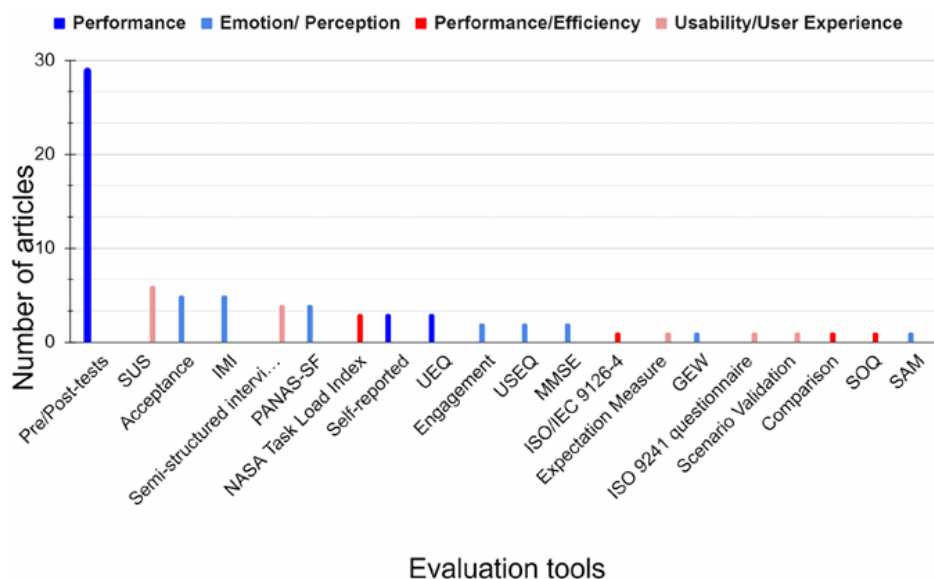
2.1.3 Αξιολόγηση και Επικύρωση των ΠΣΣΥ

Η αξιολόγηση των ΠΣΣΥ αποτελεί ουσιαστικό βήμα για να εξασφαλιστεί ότι αυτά τα εργαλεία είναι αποτελεσματικά και ωφέλιμα για τους παίκτες ασθενείς αλλά και ειδικούς. Η διαδικασία αξιολόγησης περιλαμβάνει την ποσοτική και ποιοτική ανάλυση των αποτελεσμάτων που προκύπτουν από τη χρήση τους σε διάφορες κλινικές εφαρμογές. Επιπλέον, η διασύνδεση μετρήσιμων αποτελεσμάτων υγείας, όπως η μείωση των συμπτωμάτων, και η ενσωμάτωση ανατροφοδότησης από τους χρήστες, είναι βασικοί παράγοντες που καθορίζουν την επιτυχία των ΠΣΣΥ. Παρά το γεγονός ότι η αφήγηση παίζει ρόλο στα περισσότερα παιχνίδια σοβαρού σκοπού (68%), τα είδη των παιχνιδιών διαφέρουν σημαντικά. Για αυτό το λόγο ο σχεδιασμός τους θα πρέπει να δομείται πάνω σε ένα εννοιολογικό πλαίσιο που λαμβάνει υπόψη την εμπειρία και την αλληλεπίδραση του παίκτη

[8]. Είναι προφανές, λοιπόν, ότι η σαφής και αποτελεσματική αξιολόγηση είναι κρίσιμη για την ανάπτυξη ενός παιχνιδιού σοβαρού σκοπού. Πρέπει να βασίζεται κυρίως στον συμμετέχοντα, είτε πρόκειται για παίκτη, είτε για ασθενή, είτε για επαγγελματία υγείας, ο οποίος θα προσφέρει ουσιαστική ανατροφοδότηση σχετικά με τους μηχανισμούς, το περιεχόμενο γνώσεων που μεταδίδεται, καθώς και τον αντίκτυπο που έχει το παιχνίδι μετά από κάποια περίοδο αλληλεπίδρασης. Επιπλέον, είναι σημαντικό τα παιχνίδια σοβαρού σκοπού να είναι ελκυστικά και να προκαλούν το ενδιαφέρον του χρήστη, αποφεύγοντας να γίνονται ανιαρά.

Επιπροσθέτως, η αξιολόγηση ενός ΠΣΣΥ απαιτεί σαφή περιγραφή του παιχνιδιού και των πολιτικών του κατασκευαστή, ειδικά σε ό,τι αφορά έσοδα και διαφημίσεις. Είναι κρίσιμο να αποκαλύπτονται οι πηγές εσόδων και τυχόν σχέσεις με τη φαρμακευτική βιομηχανία, καθώς μπορεί να επηρεάσουν την αξιοπιστία του παιχνιδιού. Ο εξωτερικός σκοπός τους θα πρέπει να αποσαφηνίζεται, ο οποίος μπορεί να διαφέρει από τον εσωτερικό σκοπό του παιχνιδιού, όπως στη χειρουργική εξάσκηση ή την εκπαίδευση. Τα ΠΣΣΥ που χρησιμοποιούνται για διαγνωστικούς ή θεραπευτικούς σκοπούς ενδέχεται να υπόκεινται σε κανονισμούς των οργανισμών FDA ή CE, ενώ η ένταξή τους σε εκπαιδευτικά περιβάλλοντα μπορεί να απαιτεί συγκεκριμένες τεχνικές προδιαγραφές. Οι υπάρχοντες κανονισμοί, όμως, δεν είναι επαρκείς από μόνοι τους, έτσι χρειάζονται επιπλέον βήματα αξιολόγησης και πιστοποίησης για να εξασφαλιστεί ότι ένα παιχνίδι είναι ασφαλές και αποτελεσματικό για τους χρήστες του. Ακόμη, η λειτουργικότητα και η εγκυρότητα τους είναι πιο σύνθετα στοιχεία από τα αντίστοιχα μιας απλής εφαρμογής υγείας, καθώς εμπλέκει τον χρήστη σε μια διαδικασία που αποσκοπεί στην αλλαγή συμπεριφοράς, με στόχο την εκμάθηση [10].

Σε μια συστηματική αναθεώρηση της βιβλιογραφίας πάνω στους τρόπους αξιολογήσεων των ΠΣΣΥ αναδείχτηκε ότι χρησιμοποιούνται ποικίλα εργαλεία αξιολόγησης, τα οποία χωρίζονται σε δύο κατηγορίες: αυτά που αξιολογούν το σύστημα ενός ΠΣΣΥ και εκείνα που αξιολογούν τον παίκτη. Η αξιολόγηση των συστημάτων περιλαμβάνει τόσο την τεχνική απόδοση όσο και τις υποκειμενικές απόψεις των παικτών, όπως η χρηστικότητα και η εμπειρία χρήσης. Για την αξιολόγηση των παιχτών, συχνά χρησιμοποιούνται ερωτηματολόγια πριν και μετά την ενασχόληση τους με το παιχνίδι, τα οποία εξετάζουν αν υπήρξε μάθηση ή βελτίωση μέσω του συστήματος. Επίσης, αξιολογούνται συναισθηματικές αντιδράσεις και αντιλήψεις των χρηστών, καθώς και η εσωτερική τους παρακίνηση μέσω εργαλείων όπως το Intrinsic Motivation Inventory (IMI) [11]. Για την αξιολόγηση της απόδοσης και της αποτελεσματικότητας του συστήματος, χρησιμοποιούνται εργαλεία όπως το NASA Task Load Index [12], ενώ η χρηστικότητα και η εμπειρία χρήσης αξιολογούνται συχνά μέσω της κλίμακας System Usability Scale (SUS) [13]. Επιπλέον, μέθοδοι όπως η προσέγγιση “think-aloud” [14] και οι συνεντεύξεις βοηθούν στη γρήγορη αξιολόγηση των διεπαφών των συστημάτων. Τέλος, στατιστικά τεστ, όπως το t-test [15] και το ANOVA [16], χρησιμοποιούνται για την ανάλυση δεδομένων που προκύπτουν από τις αξιολογήσεις, επιβεβαιώνοντας την εγκυρότητα των αποτελεσμάτων [17]. Το διάγραμμα στην Εικόνα 4 παρουσιάζει κάποια από τα εργαλεία που χρησιμοποιούνται για την αξιολόγηση των ΠΣΣΥ [17].



Εικόνα 4: Κάποια από τα χρησιμοποιούμενα εργαλεία αξιολόγησης και επικύρωσης σε ΠΣΣΥ [17]

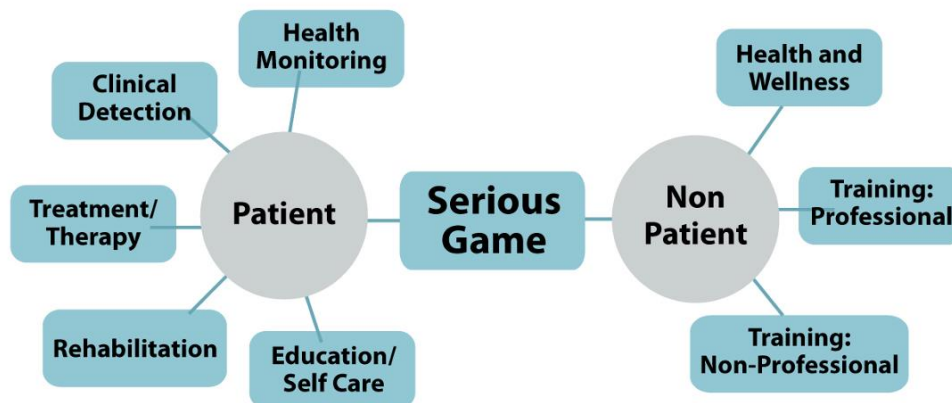
2.1.4 Μελλοντικές προκλήσεις των ΠΣΣΥ

Αν και οι μελλοντικές προοπτικές για τα ΠΣΣΥ είναι ελπιδοφόρες, εξακολουθούν να υπάρχουν πολλές προκλήσεις που πρέπει να ξεπεραστούν πριν αυτά ενσωματωθούν ευρέως σε κλινικές εφαρμογές. Πρώτον, οι επαγγελματίες υγείας είναι, και δικαίως, δύσπιστοι όσον αφορά την αποτελεσματικότητα των νέων τεχνολογιών, ενώ η βιομηχανία των παιχνιδιών πιέζει για γρήγορη προσαρμογή από επιχειρηματική σκοπιά. Παρά το γεγονός ότι οι σχεδιαστές των ΠΣΣΥ και οι πρώτοι που τα υιοθετούν στον ιατρικό χώρο αρχίζουν να κατανοούν τη σημασία της δοκιμής και της επικύρωσης νέων παιχνιδιών σοβαρού σκοπού, τα διαθέσιμα στοιχεία παραμένουν περιορισμένα. Οι συστηματικές ανασκοπήσεις που αναφέρονται σε αυτό το κεφάλαιο καταλήγουν στο συμπέρασμα ότι η ποιότητα των τρεχουσών κλινικών μελετών είναι, στην καλύτερη περίπτωση, μέτρια. Υπάρχει έλλειψη τυχαιοποιημένων κλινικών δοκιμών και ελάχιστες αρνητικές μελέτες, γεγονός που υποδηλώνει κάποια μορφή μεροληψίας στη δημοσίευση. Δεύτερον, δεν υπάρχουν ακόμη αποδείξεις για την οικονομική αποδοτικότητα των θεραπειών και της εκπαίδευσης με τη χρήση παιχνιδιών. Σε μια εποχή όπου πολλές ανεπτυγμένες χώρες εφαρμόζουν περικοπές στο κόστος υγειονομικής περίθαλψης, αυτό το ενδεχόμενο όφελος των βιντεοπαιχνιδιών απαιτεί μεγαλύτερη έμφαση. Τρίτον, η κατανόηση σχετικά με το τι παρακινεί τα άτομα να αλληλοεπιδρούν με ένα παιχνίδι είναι πολύ περιορισμένη. Είναι σημαντικό να γίνει κατανοητό ποια χαρακτηριστικά ενεργοποιούν συγκεκριμένες ομάδες χρηστών, ώστε να προβλέπεται η μακροπρόθεσμη αποτελεσματικότητα των παιχνιδιών [18].

Εκτός από αυτά τα επιστημονικά εμπόδια, πρέπει να ξεπεραστούν και πρακτικά ζητήματα. Για παράδειγμα, πολλοί λιγότερο προοδευτικοί κλινικοί και ασθενείς παραμένουν ανενημέρωτοι σχετικά με την ύπαρξη των σχετικών παιχνιδιών και για τα επιστημονικά στοιχεία που υποστηρίζουν τη χρήση τους. Πληροφορίες σχετικά με ΠΣΣΥ είναι συχνά δύσκολο να βρεθούν και οι ισχυρισμοί για την αποτελεσματικότητά τους είναι δύσκολο να

κριθούν. Επιπλέον, οι περισσότεροι κλινικοί ιατροί δεν είναι σε θέση να αξιολογήσουν την εγκυρότητα των παιχνιδιών αυτών [18].

Παρ' όλες αυτές τις αναφερθέντες δυσκολίες, η καθιέρωση επιστημονικών συνεδρίων και περιοδικών που ασχολούνται με τα ΠΣΣΥ, όπως το *Games for Health Journal*, το *BMJ Innovations* και το *JMIR Serious Games*, έχει αυξήσει σημαντικά την προβολή και την ευαισθητοποίηση του κοινού για τη σημασία τους. Έχουν γίνει προσπάθειες για τη δημιουργία πλαισίων επικύρωσης που καθοδηγούν τους ενδιαφερόμενους στην αναζήτηση των απαραίτητων πληροφοριών για την αξιολόγηση του σκοπού και της αποτελεσματικότητας ενός παιχνιδιού. Για την ενίσχυση της κλινικής προβολής, προτείνεται η δημιουργία μιας δημόσιας διαθέσιμης βιβλιοθήκης για τα ΠΣΣΥ και για αντίστοιχες ψηφιακές εφαρμογές. Η πλήρης διαφάνεια σχετικά με τα οφέλη και τους περιορισμούς τους τόσο προς το κοινό όσο και προς τους επαγγελματίες υγείας θα διευκολύνει τελικά την ένταξή τους σε πρωτόκολλα θεραπείας και προγράμματα εκπαίδευσης [18].



Εικόνα 5: Τομείς ΠΣΣΥ

2.2 Συναισθηματική Υπολογιστική Νοημοσύνη (ΣΥΝ)

2.2.1 Ανθρώπινα συναισθήματα και η συσχέτιση τους με την Τεχνητή Νοημοσύνη

Τα τελευταία τριάντα χρόνια, ένας μεγάλος αριθμός μεθόδων αναπτύσσεται συνεχώς για να διευκολυνθεί η ανάλυση των ανθρώπινων συναισθημάτων, από χειροκίνητες μεθόδους όπως τα ερωτηματολόγια που συντάσσονται από ψυχολόγους έως μεθόδους που αξιοποιούν υπολογιστικά συστήματα. Παραμένει περίπλοκο ζήτημα στην έρευνα η ανάλυση των συναισθημάτων, καθώς ο αντικειμενικός ορισμός και η κατανόηση τους δεν μπορούν φέρουν σαφή απάντηση. Σήμερα, η αναγνώριση συναισθημάτων μέσω των υπολογιστικών συστημάτων έχει αναπτυχθεί σε πολλές εφαρμογές διαφορετικών αναγκών, διότι έχει μελετηθεί με διάφορους τρόπους, συμπεριλαμβανομένης της αναγνώρισης εκφράσεων προσώπου, σημάτων ομιλίας, φυσιολογικών σημάτων, καθώς και συναισθημάτων που μεταδίδονται μέσω γραπτού κειμένου [19].

Η πλειονότητα των ερευνών που επικεντρώνονται στη μελέτη συναισθημάτων και συμπεριφορών στην εκπαίδευση, έχει συγκεντρώσει δεδομένα είτε μέσω ανάλυσης από ειδικούς, είτε με δηλωτικές μεθόδους, όπως από τις αναφορές των χρηστών, είτε και με συνδυασμό αυτών των δύο προσεγγίσεων. Αν και οι συγκεκριμένες μέθοδοι προσφέρουν πλεονεκτήματα, όπως η χαμηλή τους δαπάνη και η προσαρμοστικότητα τους σε εκπαιδευτικά πλαίσια, εντούτοις αντιμετωπίζουν σημαντικούς περιορισμούς. Για να υπερκεραστούν τα προβλήματα που παρουσιάζονται, εφαρμόζονται διάφορες στρατηγικές που χρησιμοποιούν νευροβιολογικές και συμπεριφορικές μετρικές, οι οποίες ενσωματώνονται σε εκπαιδευτικές μελέτες. Σημαντική είναι η παρατήρηση ότι οι έρευνες δείχνουν στενή συσχέτιση ανάμεσα στον τρόπο με τον οποίο τα συναισθήματα βιώνονται και εκφράζονται μέσω του εγκεφάλου. Αυτό σημαίνει πως τόσο οι νευροβιολογικές όσο και οι συμπεριφορικές προσεγγίσεις παρέχουν αξιόπιστα και συνεπή συναισθηματικά δεδομένα. Παρ' όλα αυτά, σε πραγματικά περιβάλλοντα, αυτές οι μέθοδοι συχνά αποτυγχάνουν να αποδώσουν τα επιθυμητά αποτελέσματα, λόγω της πολυπλοκότητας των παραμέτρων που επηρεάζουν την ανθρώπινη έκφραση στην καθημερινή ζωή [20].

Η τεχνητή νοημοσύνη ως μία από τις νεότερες τεχνολογίες με την υπόσχεση επίτευξης αξιόπιστων και χρήσιμων αποτελεσμάτων σε πολυάριθμες ερευνητικές περιοχές έχει, προφανώς, εφαρμογή και στον τομέα των ανθρώπινων συναισθημάτων. Η διαδικασία κατασκευής ενός μοντέλου αναγνώρισης συναισθημάτων περιλαμβάνει τη συλλογή δεδομένων, την εξαγωγή χαρακτηριστικών που σχετίζονται με το συναίσθημα, τη μείωση των χαρακτηριστικών με κατάλληλη επεξεργασία στα δεδομένα και, τέλος, τη δημιουργία ενός μοντέλου. Μετά την ολοκλήρωση των τριών πρώτων βημάτων, η τελική εργασία είναι να σχεδιαστεί ένα αποτελεσματικό μοντέλο συναισθημάτων με βάση ορισμένα κριτήρια απόδοσης ταξινόμησης, το οποίο είναι και το βασικό κριτήριο για την ακρίβεια της πρόβλεψης που θα επιχειρηθεί. Επιπλέον, δεν χρησιμοποιούνται μόνο μέθοδοι απλής μηχανικής μάθησης για την επίτευξη του παραπάνω στόχου, αλλά και μέθοδοι βαθιάς μηχανικής μάθησης είναι διαδεδομένες, διότι φαίνεται να είναι αποτελεσματικές στην εξαγωγή και στην αναπαράσταση χαρακτηριστικών εις βάθος [21].



Εικόνα 6: Καταγραφή μορφασμών προσώπου (ΣΥΝ)

2.2.2 Πεδία Εφαρμογών ΣΥΝ

Μία από τις κυριότερες μορφές της ανάλυσης που γίνεται στην βιβλιογραφία για την αναγνώριση και την καταγραφή των ανθρωπίνων συναισθημάτων είναι μέσω των εκφράσεων του προσώπου. Αυτή η μέθοδος είναι καθιερωμένη ευρέως λόγω της αλληλεπίδρασης των ανθρώπων και των υπολογιστών σε διάφορες περιπτώσεις, όπως για παράδειγμα με τα έξυπνα τηλέφωνα, τα συστήματα έξυπνου ελέγχου, τα γενικότερα κοινωνικά δίκτυα, καθώς και σε μελέτες για τη συμπεριφορά και την ψυχολογία των ανθρώπων. Αυτά τα συστήματα αξιολογούνται με βάση την ικανότητά τους να κωδικοποιούν και να μεταδίδουν πληροφορίες από τις εκφράσεις του προσώπου [22]. Ο σχεδιασμός και η ανάπτυξη αλγοριθμικών λύσεων ικανών να ερμηνεύουν τα συναισθήματα του προσώπου από ανθρώπινα πρόσωπα ανοίγει ένα νέο παράθυρο δυνατοτήτων για το πλαίσιο αλληλεπίδρασης ανθρώπου-υπολογιστή, όπως στη ρομποτική και τα παιχνίδια, μεταξύ πολλών άλλων. Ωστόσο, το πώς η συμπεριφορά της ανθρώπινης έκφρασης είναι αμετάβλητη μεταξύ διαφορετικών ατόμων και πώς μπορεί να παρεμβαίνει στην ανθρώπινη επικοινωνία με την πάροδο του χρόνου, είναι ενδιαφέρον ερώτημα που πρέπει να μελετηθεί και να μοντελοποιηθεί υπολογιστικά. Προκειμένου να παραχθούν αναλυτικά μοντέλα για την αναγνώριση ανθρώπινης έκφρασης από εικόνες προσώπου, οι επιστήμονες υπολογιστών διερευνούν διαφορετικούς τρόπους προσεγγίσεων, οι οποίοι είναι ικανοί να εφαρμοστούν αποτελεσματικά με αποδοτικούς αλγόριθμους. Υπάρχουν, επίσης, ισχυροί συσχετισμοί μεταξύ των πεδίων επεξεργασίας εικόνας και τεχνητής νοημοσύνης που διερευνούν αυτό το θέμα. Η εύρεση υπολογιστικών τεχνικών που είναι σε θέση να αναγνωρίζουν αποτελεσματικά τα συναισθήματα από τις εικόνες του προσώπου, είναι επιθυμητή σε ποικίλα πεδία με σημαντικές επιρροές [23].

Η αναγνώριση συναισθημάτων σε κείμενα είναι ένας σημαντικός κλάδος επεξεργασίας φυσικής γλώσσας (NLP) που μπορεί να ωφελήσει πολλές εφαρμογές σε διαφορετικούς τομείς, όπως η εξόρυξη δεδομένων, η ηλεκτρονική μάθηση, τα συστήματα φιλτραρίσματος πληροφοριών, η αλληλεπίδραση ανθρώπου-υπολογιστή αλλά και η ψυχολογία. Η ρητή αναγνώριση συναισθημάτων σε ένα κείμενο είναι το πρόβλημα που αντιμετωπίζεται περισσότερο στη βιβλιογραφία, διότι αντιμετωπίζεται σαν πρόβλημα κυρίως με τον προσδιορισμό λέξεων-κλειδιών. Η άρρητη αναγνώριση συναισθημάτων είναι το πιο δύσκολο πρόβλημα που απαιτεί βαθύτερη ανάλυση, καθώς επειδή ένα τέτοιο συναίσθημα συνήθως κρύβεται μέσα στο κείμενο, η επίλυσή του απαιτεί κατανόηση του γενικότερου νοήματος που πηγάζει από τη συλλογική οπτική του κειμένου. Υπάρχουν τέσσερις κύριες προσεγγίσεις για την άρρητη αναγνώριση συναισθημάτων σε ένα κείμενο, από τις οποίες η μια βασίζεται, όπως στη ρητή αναγνώριση, στον εντοπισμό λέξεων-κλειδιών και στην αντιστοίχιση μιας ετικέτας συναισθήματος με βάση τη λέξη αυτή κάθε φορά. Μια άλλη προσέγγιση βασίζεται σε κανόνες για να ερμηνευθούν συναισθηματικά τα δεδομένα μέσω επεξεργασίας κειμένου και εξαγωγής των κανόνων αυτών με γλωσσικές, στατιστικές και υπολογιστικές τεχνικές. Οι επιλεγμένοι κανόνες εφαρμόζονται στο σύνολο των δεδομένων για να καθορίσουν τις ετικέτες των συναισθημάτων. Τέλος, οι δύο τελευταίες μέθοδοι, η επιβλεπόμενη και μη επιβλεπόμενη μάθηση, καθώς και η βαθιά μάθηση, επιτρέπουν στα συστήματα να μαθαίνουν αυτόματα, με τη βαθιά μάθηση να ξεχωρίζει μέσω μιας ιεραρχίας εννοιών που τη χαρακτηρίζουν, όπου οι πιο πολύπλοκες έννοιες βασίζονται σε απλούστερες [24].

Ένας ακόμη ενδιαφέρων τομέας στην αναγνώριση συναισθημάτων είναι η χρήση ηλεκτροεγκεφαλογραφημάτων (EEG), στην οποία οι πρόσφατες εξελίξεις δείχνουν έντονο

ενδιαφέρον από την ερευνητική κοινότητα, καθώς σε φορετές λύσεις καταναλωτικής ποιότητας μπορούν να προσφέρουν μια φθηνή και απλή λύση για την αναγνώριση συναισθημάτων. Το ηλεκτροεγκεφαλογράφημα θεωρείται μια φυσιολογική ένδειξη στην οποία οι ηλεκτρικές δραστηριότητες των νευρικών κυττάρων συγκεντρώνονται στον ανθρώπινο εγκεφαλικό φλοιό και χρησιμοποιείται για την καταγραφή δραστηριοτήτων του εγκεφάλου. Είναι αξιόπιστο στην αναγνώριση συναισθημάτων λόγω της σχετικά αντικειμενικής του αξιολόγησης σε σύγκριση με μη φυσιολογικές ενδείξεις, όπως οι εκφράσεις στο πρόσωπο και οι χειρονομίες. Γενικότερα, η νευρολογία στοχεύει να διασαφηνίσει τα νευρικά δίκτυα που βρίσκονται σε συναισθηματικές καταστάσεις και τις συνέπειές τους στη φυσιολογία, τη γνωστική λειτουργία και τη συμπεριφορά ενός ατόμου. Το πεδίο έχει ιστορικά επικεντρωθεί γύρω από τον καθορισμό των ανθρώπινων συναισθημάτων και τον προσδιορισμό συνδυαστικών ρόλων του σώματος, των αισθήσεων και των συναισθημάτων [25].

2.2.3 Ενσωμάτωση της Τεχνητής Νοημοσύνης με τα ΠΣΣ

Η ενσωμάτωση της Τεχνητής Νοημοσύνης με τα ΠΣΣ αποτελεί μια καινοτόμο προσέγγιση που αναδεικνύει τις δυνατότητες της τεχνολογίας στη βελτίωση της εκπαίδευσης, της υγείας και άλλων κρίσιμων τομέων. Όπως αναλύθηκε και προηγουμένως, τα ΠΣΣ σχεδιάζονται για να επιτυγχάνουν μαθησιακούς ή θεραπευτικούς στόχους πέρα από την απλή ψυχαγωγία. Η ενσωμάτωση αλγορίθμων τεχνητής νοημοσύνης μπορεί να ενισχύσει την εξατομίκευση, την προσαρμοστικότητα και την αλληλεπίδραση του χρήστη. Επιπλέον, μπορεί να παρακολουθείται η πρόοδος των χρηστών, ακόμη και να αναλύονται τα δεδομένα σε πραγματικό χρόνο και να προσαρμόζεται το περιεχόμενο ανάλογα με τις ανάγκες και το επίπεδο εμπλοκής των παικτών, καθιστώντας την εμπειρία πιο αποτελεσματική και καθηλωτική. Στη βιβλιογραφία αναλύονται σε δύο βασικές κατηγορίες οι τομείς της τεχνητής νοημοσύνης συνδυαστικά με τα ΠΣΣ, οι οποίες είναι η λήψη αποφάσεων και η μηχανική μάθηση [26].

Οι αλγόριθμοι των λήψεων αποφάσεων χρησιμοποιούνται για λογικές και ορθολογικές διαδικασίες λήψης αποφάσεων με βάση τις πληροφορίες που συγκεντρώνει ένα σύστημα. Η είσοδος των συστημάτων αυτών είναι η γνώση που έχει συλλεχθεί προηγουμένως και η έξοδος είναι ένα αίτημα κάποιας απαιτούμενης δράσης. Μια από τις κύριες τεχνικές είναι τα δέντρα αποφάσεων (Decision Trees), τα οποία χρησιμοποιούν σύνολα κανόνων για την πρόβλεψη, βασισμένα σε δεδομένα που συλλέγονται κατά τη διάρκεια της αξιολόγησης του συστήματος. Η ασαφής λογική (Fuzzy Logic) προσφέρει έναν πιο ευέλικτο τρόπο διαχείρισης της αβεβαιότητας, επιτρέποντας την εξέταση πολλαπλών ενδεχομένων, και όχι απλώς δυαδικών επιλογών. Επίσης, τα Μαρκοβιανά συστήματα αποτελούν μια ακόμη τεχνική, η οποία στηρίζεται σε πιθανές καταστάσεις και μεταβάσεις, χωρίς να απαιτείται μνήμη από το σύστημα, προσφέροντας έτσι μια πιο αποδοτική μορφή λήψης αποφάσεων. Παράλληλα, τα συστήματα βασισμένα σε κανόνες (Rule-Based Systems) επιτρέπουν τη χρήση προκαθορισμένων οδηγιών για τη λήψη αποφάσεων και την εκτέλεση τακτικών κινήσεων, ενώ οι μηχανές πεπερασμένων καταστάσεων (Finite-State Machines) λειτουργούν με μεταβάσεις μεταξύ περιορισμένων καταστάσεων βάσει συγκεκριμένων συνθηκών. Όλες αυτές οι τεχνικές συνεργάζονται για να βελτιώσουν τη λήψη αποφάσεων στα ΠΣΣ, προσφέροντας εξατομικευμένες και αποτελεσματικές εμπειρίες στους παίκτες [26].

Ένα σύστημα μηχανικής μάθησης, από την άλλη, μαθαίνει από τα δεδομένα που του δίνονται όταν βελτιώνεται στην εκτέλεση μιας καθορισμένης εργασίας, βελτιστοποιώντας την απόδοσή του με τη χρήση διάφορων μετρικών που το αποδεικνύουν. Στον τομέα των ΠΣΣ, αυτή η διαδικασία εφαρμόζεται στα πρότυπα των ίδιων των παικτών και επίσης χρησιμοποιείται στη βελτιστοποίηση της συμπεριφοράς των χαρακτήρων χωρίς παίκτη (NPC) και των παραγόντων του παιχνιδιού, παρέχοντάς τους καλύτερο ρεαλισμό. Το εγγενές πρόβλημα των αλγορίθμων μάθησης στον συγκεκριμένο τομέα είναι ο χρόνος χρήσης που απαιτείται για τη δημιουργία ενός αποτελεσματικού συστήματος μάθησης. Μια από τις βασικές τεχνικές είναι ο ταξινομητής Naive Bayes, ο οποίος χρησιμοποιεί το θεώρημα του Bayes για να ταξινομεί και να προβλέπει αποτελέσματα, βασιζόμενος στην πιθανότητα που έχει ένα χαρακτηριστικό να ανήκει σε μία ή περισσότερες κλάσεις, με όλα τα χαρακτηριστικά να θεωρούνται ανεξάρτητα. Επιπλέον, τα τεχνητά νευρωνικά δίκτυα (ANNs) προσφέρουν μια άλλη σημαντική τεχνική μάθησης, προσομοιώνοντας τη λειτουργία του ανθρώπινου εγκεφάλου μέσω ενός συστήματος "νευρώνων" που συνδέονται με βάρη, προσδίδοντας έτσι γνώση στο σύστημα. Ακόμη μία τεχνική είναι αυτή της λογικής βασισμένης σε περιπτώσεις (Case-Based Reasoning), η οποία επιτρέπει στα συστήματα να μαθαίνουν από προηγούμενες λύσεις και να προσαρμόζουν το περιεχόμενο του παιχνιδιού σύμφωνα με την αλληλεπίδραση του χρήστη, δημιουργώντας εξατομικευμένες εμπειρίες. Οι μηχανές υποστήριξης διανυσμάτων (SVMs) προσφέρουν, επίσης, έναν αποτελεσματικό τρόπο ανάλυσης δεδομένων και αναγνώρισης προτύπων, χρησιμοποιώντας υπερεπίπεδα για τη διάκριση δεδομένων σε διαφορετικές κλάσεις με τρόπο που μεγιστοποιεί τη διαχωριστική απόσταση μεταξύ τους. Αντίστοιχα με τις διαφορετικές τεχνικές λήψης αποφάσεων, στο κομμάτι της μηχανικής μάθησης βελτιώνουν την προσαρμοστικότητα και τη λειτουργία των ΠΣΣ, προσφέροντας εξατομικευμένες και δυναμικές εμπειρίες στους χρήστες [26].

2.2.4 Η ΣΥΝ στα ΠΣΣ

Μεταξύ των διαφορετικών πλεονεκτημάτων που παρέχουν τα ΠΣΣ για τη μάθηση, η ικανότητά τους να αξιοποιούν τη συνεργασία μεταξύ συναισθηματικών και μαθησιακών διαδικασιών οικειοποίησης είναι ένας από τους μεγαλύτερους στόχους τους. Τα συναισθήματα των παικτών, στην πράξη, διαδραματίζουν πρωταρχικό ρόλο στη δομή και την εμπειρία των ΠΣΣ, καθώς συνδέονται στενά με τη δράση και προετοιμάζουν τον παίκτη για ενέργειες όπως είναι η προσέγγιση ενός προβλήματος στο παιχνίδι με βάση τους διαθέσιμους πόρους. Επιπλέον, περιλαμβάνουν συγκεκριμένα συναισθηματικά χαρακτηριστικά που μπορούν να ενισχύσουν τα μαθησιακά αποτελέσματα των παικτών, αξιοποιώντας την αλληλεξάρτηση μεταξύ των συναισθημάτων και της προσήλωσης τους [26].

Τα ΠΣΣ διακρίνονται με βάση τα χαρακτηριστικά τους που ενισχύουν τα θετικά συναισθήματα στους παίκτες, συμβάλλοντας έτσι στην ενίσχυση του μαθησιακού τους χαρακτήρα. Διαθέτουν, συνεπώς, ευχάριστα και ικανοποιητικά στοιχεία, τα οποία προσελκύουν τους παίκτες και τους βοηθούν να αναπτύξουν αίσθηση επάρκειας, αυτοπεποίθησης και διεύρυνσης των ικανοτήτων τους. Το αισθητικό επίπεδο ευχαρίστησης σχετίζεται με τα οπτικά ερεθίσματα, όπως τα χρώματα, οι μορφές και οι ρυθμικές εναλλαγές, τα οποία προσελκύουν αρχικά τους παίκτες. Επιπλέον, η ευχαρίστηση από την επίτευξη

στόχων, όπως η νίκη ή η καλή απόδοση, προκαλεί την αίσθηση της ικανοποίησης και του ελέγχου της κατάστασης, ενισχύοντας την αυτοεκτίμηση των παικτών [26].

Πέρα από την αισθητική ευχαρίστηση και την αίσθηση επιτυχίας, όμως, τα ΠΣΣ ενισχύουν επίσης την ευχαρίστηση από την ανακάλυψη, προωθώντας τη φυσική περιέργεια και την ανάγκη για εξερεύνηση νέων μονοπατιών μέσα στα παιχνίδια. Αυτή η επιθυμία ανακάλυψης συνδέεται άμεσα με τη συναισθηματική απαίτηση για την ολοκλήρωση μιας αποστολής, παρέχοντας έτσι έναν επιπλέον κινητήριο παράγοντα για τους παίκτες. Ταυτόχρονα, τα ΠΣΣ επιτρέπουν στους παίκτες να διαχειριστούν την απογοήτευση και την αποθάρρυνση που σχετίζονται με την αποτυχία σε κάποια πρόκληση του παιχνιδιού. Σε αντίθεση με τις παραδοσιακές μαθησιακές διαδικασίες, τα ΠΣΣ ενσωματώνουν τα λάθη ως ένα φυσικό και εκπαιδευτικό στοιχείο, το οποίο όχι μόνο μειώνει τα αρνητικά συναισθήματα, αλλά και μετατρέπει τα λάθη σε ευκαιρίες για εξερεύνηση καινοτόμων λύσεων [26].

Τέλος, ένα από τα σημαντικά χαρακτηριστικά των ΠΣΣ είναι η ικανότητά τους να ενισχύουν τα γνωστικά και συναισθηματικά οφέλη μέσω της διαχείρισης των λαθών του παίκτη. Τα ΠΣΣ παρέχουν στους παίκτες τη δυνατότητα να πειραματίζονται και να κάνουν λάθη χωρίς να φοβούνται τις συνέπειες, μέσα σε ένα ασφαλές περιβάλλον. Η παρουσία ειδικής παιδαγωγικής ανατροφοδότησης, που εστιάζει στην απόδοση του παίκτη και όχι στην προσωπική του ταυτότητα, προάγει μια θετική αξιολόγηση των λαθών και υποστηρίζει την ενεργό συμμετοχή του στη μαθησιακή διαδικασία. Με αυτόν τον τρόπο τα θετικά συναισθήματα που ενισχύονται κατά τη διάρκεια αυτής της διαδικασίας βελτιώνουν την προσήλωση, τη μνήμη αλλά και τη συνολική εμπειρία του παίκτη, δημιουργώντας μια πιο ευχάριστη και αποτελεσματική εμπειρία που έχει και ως αντίκτυπο στο τέλος το μαθησιακό στοιχείο [26].

2.3 Δεδομένα Καταγραφής Βίντεο

2.3.1 Ανάλυση δεδομένων βίντεο

Η χρήση δεδομένων βίντεο σε ερευνητικές εργασίες αποτελεί σημαντική εξέλιξη, καθώς οι σύγχρονες τεχνολογικές δυνατότητες, όπως οι ψηφιακές κάμερες και τα λογισμικά επεξεργασίας, παρέχουν νέες προοπτικές ανάλυσης και παρουσίασης. Τέτοιου τύπου δεδομένα ξεχωρίζουν λόγω της δυνατότητάς τους να καταγράφουν πλήρως μια διαδικασία, συνδυάζοντας εικόνα και ήχο, κάτι που δεν είναι εφικτό με άλλες μεθόδους, όπως η ηχογράφηση ή η καταγραφή γραπτών. Η δυνατότητα καταγραφής μη λεκτικών στοιχείων, όπως οι εκφράσεις του προσώπου και οι κινήσεις, ενισχύει την πληρότητα του περιεχομένου που καταγράφεται και μπορεί να συμβάλει στη βαθύτερη κατανόηση δυναμικών καταστάσεων. Επιπλέον, η αναδρομική επεξεργασία των βίντεο επιτρέπει τόσο στους ερευνητές όσο και στους συμμετέχοντες να επανεξετάσουν το υλικό, προσφέροντας επιπλέον πληροφορίες στη συλλογή των δεδομένων αλλά και στη μετέπειτα ανάλυση. Παρόλο, όμως, που η χρήση βίντεο δεδομένων ενδέχεται να συνοδεύεται από προκλήσεις, όπως το υψηλό κόστος και η αυξημένη δυσκολία ανάλυσης, τα οφέλη τους όσον αφορά την πληρότητα των δεδομένων και τη δυνατότητα καταγραφής πλουσιότερου περιεχομένου καθιστούν τη χρήση τους ιδιαίτερα πολύτιμη για την έρευνα [27].

Οι αλγόριθμοι τεχνητής νοημοσύνης έχουν εξαιρετική επιτυχία στην έρευνα επεξεργασίας βίντεο, όμως, η ποικιλομορφία σε χωρικά και χρονικά στοιχεία καθιστά τη χρήση των δεδομένων από βίντεο ένα δύσκολο πρόβλημα σε πολλές περιπτώσεις. Συγκεκριμένα χρησιμοποιούνται κυρίως τεχνικές βαθιάς μάθησης στην επεξεργασία βίντεο στη βιβλιογραφία, οι οποίες περιλαμβάνουν μια ποικιλία προσεγγίσεων, όπως τα Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), καθώς και υβριδικές μεθόδους. Τα CNN έχουν εφαρμοστεί ευρέως στην αναγνώριση αντικειμένων, στην ανίχνευση ασυνήθιστων γεγονότων και στην παρακολούθηση αντικειμένων. Ένα σύστημα CNN χρησιμοποιεί ροές για την επεξεργασία πληροφοριών τόσο χωρικά όσο και χρονικά, εξασφαλίζοντας έτσι την ανάλυση της κίνησης. Υπάρχουν διάφορες παραλλαγές CNN που συνδυάζουν αυτές τις πληροφορίες για την αναγνώριση κάποιας δραστηριότητας σε πραγματικό χρόνο. Επιπλέον, τα μοντέλα που χρησιμοποιούν CNN για την επεξεργασία βίντεο έχουν δείξει υψηλή απόδοση στην αναγνώριση προσώπων και την ανίχνευση κινούμενων αντικειμένων, ακόμα και σε συνθήκες μειωμένης ποιότητας βίντεο [28].

Οι αλγόριθμοι RNN, και ειδικότερα οι παραλλαγές τους, δηλαδή τα LSTM και GRU, χρησιμοποιούνται για την επεξεργασία χρονικών δεδομένων σε βίντεο, δεδομένου ότι οι ακολουθίες καρέ απαιτούν ανάλυση χωροχρονικών σχέσεων. Αυτές οι μέθοδοι είναι κατάλληλες για την αναγνώριση σύνθετων δραστηριοτήτων ή και για την παρακολούθηση πολλών αντικειμένων σε διαδοχικές σκηνές βίντεο. Μια σημαντική προσέγγιση για την αναγνώριση ανθρώπινης δραστηριότητας περιλαμβάνει τη χρήση LSTM στην προσπάθεια κατανόησης της δυναμικής των ατόμων μέσα σε ένα βίντεο, καθώς και στην αναγνώριση εκφράσεων προσώπου και επιπλέον στην ανίχνευση ανωμαλιών. Ταυτόχρονα, υβριδικές μέθοδοι που συνδυάζουν CNN και RNN αξιοποιούν τα πλεονεκτήματα των CNN για την εξαγωγή οπτικών χαρακτηριστικών και των RNN για την ανάλυση της χρονικής αλληλουχίας, επιτρέποντας πιο ακριβείς προβλέψεις σε εφαρμογές σε ιδιαίτερα σύνθετες εφαρμογές [28].



Εικόνα 7: Επεξεργασία βίντεο δεδομένων

2.3.2 Εφαρμογές με χρήση δεδομένων από βίντεο

Στη σύγχρονη εποχή προσφέρεται ένα ευρύ φάσμα εφαρμογών επεξεργασίας βίντεο σε διάφορους τομείς το οποίο με την όλο και μεγαλύτερη αύξηση της υπολογιστικής ικανότητας θα συνεχίζει να μεγαλώνει. Στην ψυχαγωγία, η επεξεργασία βίντεο χρησιμοποιείται για τη δημιουργία αυτόματων περιγραφών των δράσεων που εμφανίζονται σε βίντεο, ενισχύοντας έτσι την αλληλεπίδραση με το περιεχόμενο. Στην υγειονομική περίθαλψη, οι τεχνολογίες αυτές συμβάλλουν στην παρακολούθηση ασθενών και στη διάγνωση. Παράλληλα, η διαχείριση της κυκλοφορίας και των μεταφορών αξιοποιεί την επεξεργασία βίντεο για την παρακολούθηση της κίνησης οχημάτων, βελτιώνοντας την ασφάλεια στους δρόμους και την αποτελεσματικότητα των συστημάτων μεταφοράς. Επιπλέον, οι εφαρμογές βίντεο επεξεργασίας βρίσκουν χρήση στην αυτοματοποίηση κατοικιών, την ανίχνευση φλόγας και καπνού, καθώς και στην ενίσχυση της ασφάλειας και της προστασίας σε καθημερινές δραστηριότητες [28].

Οι εφαρμογές αυτές δεν περιορίζονται μόνο στους παραπάνω τομείς, αλλά επεκτείνονται και στην εκπαίδευση, όπου χρησιμοποιούνται για τη δημιουργία εκπαιδευτικών βίντεο, καθώς και στην παρακολούθηση έξυπνων πόλεων μέσω συστημάτων επιτήρησης. Ένα από τα πιο δημοφιλή παραδείγματα είναι η διαχείριση πλήθους, όπου η επεξεργασία βίντεο επιτρέπει τον υπολογισμό των ατόμων σε σημεία εισόδου και εξόδου ενός χώρου σε πραγματικό χρόνο ή σε περιοδικά διαστήματα. Επιπλέον, η ανάλυση βίντεο βοηθά στην εξαγωγή πληροφοριών για τη λήψη αποφάσεων, παρέχοντας στους αρμόδιους φορείς τη δυνατότητα να κατανοήσουν και να ανταποκριθούν σε τρέχουσες καταστάσεις με μεγαλύτερη ακρίβεια αλλά και ταχύτητα [28].

Υπάρχουν διάφορες λειτουργίες στους παραπάνω τομείς στα πλαίσια χρήσης βίντεο δεδομένων καθώς η ανάλυση και κατανόηση των δεδομένων βίντεο αποτελεί ένα πολύπλευρο μέσο. Μια από τις κύριες είναι η αναγνώριση ανθρώπινων δράσεων, η οποία επιτρέπει την ταυτοποίηση και κατηγοριοποίηση συγκεκριμένων δραστηριοτήτων, όπως το περπάτημα, το τρέξιμο, ή ακόμα και πτώσεις, που είναι κρίσιμες για την ανάλυση συμπεριφοράς. Αυτή η λειτουργία βρίσκει εφαρμογή σε συστήματα επιτήρησης, αθλητικές αναλύσεις και στην παρακολούθηση ηλικιωμένων ατόμων για την ανίχνευση επικίνδυνων καταστάσεων. Παράλληλα, η ανίχνευση κίνησης είναι μια άλλη σημαντική λειτουργία που στοχεύει στον εντοπισμό της παρουσίας κίνησης μέσα σε μια σκηνή. Μέσω της ανάλυσης της κίνησης των αντικειμένων, γίνεται δυνατός ο υπολογισμός της κατεύθυνσης και της ταχύτητας της κίνησης, με εφαρμογές όπως η παρακολούθηση οχημάτων σε οδική κυκλοφορία ή η επιτήρηση περιοχών για ύποπτες δραστηριότητες [28].

Επιπλέον, η ανίχνευση και η παρακολούθηση αντικειμένων αποτελούν ένα ακόμη σημαντικό κομμάτι των περιοχών εφαρμογής στην ανάλυση βίντεο δεδομένων. Η ανίχνευση επιτρέπει την ταυτοποίηση και εντοπισμό ορατών αντικειμένων σε βίντεο, όπως οχήματα, ανθρώπους ή ζώα, και χρησιμοποιείται σε εφαρμογές όπως η αυτόνομη οδήγηση, η ασφάλεια αλλά και η αθλητική ανάλυση. Η παρακολούθηση των αντικειμένων, από την άλλη, αφορά την παρακολούθηση της κίνησης καθ' όλη τη διάρκεια του βίντεο, επιτρέποντας τη μελέτη της πορείας του και την ανάλυση της συμπεριφοράς του σε συνεχόμενα καρέ. Η κατανόηση της σκηνής και η ανάλυση της συμπεριφοράς των αντικειμένων είναι επίσης σημαντικές, καθώς επιτρέπουν την αναγνώριση σύνθετων γεγονότων και τη διάκριση μεταξύ φυσιολογικών και

ασυνήθιστων δραστηριοτήτων, λειτουργώντας έτσι και ως καλή τεχνική για ανίχνευση ανωμαλιών σε προβλήματα τεχνητής νοημοσύνης. Συνολικά, οι λειτουργίες αυτές καθιστούν την επεξεργασία βίντεο απαραίτητο εργαλείο σε τομείς όπως η επιτήρηση, η αυτόνομη τεχνολογία και η βελτίωση της ασφάλειας [28].

2.3.3 Προκλήσεις και μελλοντικές κατευθύνσεις

Αν και έχει σημειωθεί σημαντική πρόοδος τα τελευταία χρόνια, εξακολουθούν να υπάρχουν πολλές προκλήσεις στην εφαρμογή τεχνικών βαθιάς μάθησης στην επεξεργασία βίντεο και στην ανάπτυξη μοντέλων σε εφαρμογές της πραγματικής ζωής. Μια από τις κύριες δυσκολίες αφορά την κακή ποιότητα των βίντεο, για παράδειγμα όταν προέρχονται από κάμερες επιτήρησης τοποθετημένες σε μεγάλες αποστάσεις αδυνατούν να προσφέρουν τα απαιτούμενα δεδομένα. Σε τέτοιες περιπτώσεις, οι σκηνές μπορεί να περιλαμβάνουν πολυκοσμία ή απόκρυψη αντικειμένων, γεγονός που καθιστά δυσκολότερη την επεξεργασία και την ανάλυση των δεδομένων. Επιπλέον ο υποκείμενος στόχος σε αυτές τις συνθήκες είναι μικρός, με αποτέλεσμα να κάνει την αναγνώρισή του ακόμη πιο δύσκολη, ενώ η χαμηλή ποιότητα των βίντεο λόγω της απόστασης των καμερών επιβαρύνει ακόμη περισσότερο το ήδη απαιτητικό σενάριο. Ένα άλλο πρόβλημα είναι η παρακολούθηση πολλαπλών κινούμενων αντικειμένων, η οποία αποτελεί πρόκληση λόγω της αβεβαιότητας στην κίνηση των στόχων και της αδυναμίας των συστημάτων να μοντελοποιήσουν με ακρίβεια τη δυναμική της κίνησης σε τόσα αντικείμενα ή άτομα [28].

Επιπροσθέτως, η επεξεργασία βίντεο σε πραγματικά περιβάλλοντα συχνά περιλαμβάνει δυναμικά υπόβαθρα, τα οποία μεταβάλλονται συνεχώς. Οι αλλαγές στον φωτισμό, οι αποκρύψεις και οι αλλαγές στις οπτικές γωνίες προσθέτουν σημαντικές δυσκολίες στην επεξεργασία βίντεο σε σύνθετα σκηνικά. Ακόμη ένα κρίσιμο ζήτημα είναι η έλλειψη κατάλληλων συνόλων δεδομένων, τόσο ως προς την ποσότητα όσο και την ποιότητα τους. Από την άλλη, ο ανεπαρκής σχολιασμός των δεδομένων οδηγεί σε ανακρίβειες, ενώ η περιορισμένη διαθεσιμότητα των δημοσίως προσβάσιμων δεδομένων περιορίζει τη δυνατότητα διεξαγωγής διαφορετικών ερευνών σε κάθε τομέα συγκεκριμένα, έτσι ώστε να υπάρχουν πιο αντιπροσωπευτικά αποτελέσματα. Τέλος, η ανάγκη για υπολογιστική ισχύ αποτελεί ακόμη ένα σημαντικό εμπόδιο, καθώς οι σύγχρονες τεχνικές βαθιάς μάθησης απαιτούν εξειδικευμένες μονάδες επεξεργασίας γραφικών (GPUs), προκειμένου να επεξεργαστούν τα μεγάλα σύνολα δεδομένων βίντεο, λόγω διαθεσιμότητας αλλά και κόστους [28].

Είναι σημαντικό να αναφερθεί ότι η επεξεργασία και η αξιοποίηση δεδομένων βίντεο κυρίως μέσω τεχνολογιών υπολογιστικής όρασης εγείρει σημαντικά ηθικά και νομικά ζητήματα. Η ραγδαία ανάπτυξη και εφαρμογή αυτών των τεχνολογιών σε διάφορες συσκευές και πλατφόρμες, όπως κινητά τηλέφωνα, διαδικτυακές εφαρμογές και κάμερες παρακολούθησης, εντείνει τις ανησυχίες σχετικά με την προστασία της ιδιωτικότητας και της κατάλληλης χρήσης των δεδομένων. Οι τεχνολογίες αυτές μπορούν να αναγνωρίζουν πρόσωπα ή να ανιχνεύουν προσωπικές πληροφορίες από βίντεο, γεγονός που καθιστά αναγκαία την εφαρμογή αυστηρών ρυθμίσεων και πρωτοκόλλων για την αποτροπή της καταχρηστικής ή αθέμιτης εκμετάλλευσης των δεδομένων. Επιπλέον, είναι επιτακτική η ανάγκη να γίνεται ολοκληρωμένη εξέταση των ηθικών ζητημάτων που προκύπτουν, ώστε να διασφαλιστεί η υπεύθυνη χρήση αυτών των τεχνολογιών στα πλαίσια της έρευνας και της κοινωνίας γενικότερα [29].

Κεφάλαιο 3: Μεθοδολογία

3.1 Περιγραφή της Διαδικασίας

3.1.1 Το ΠΣΣΥ που χρησιμοποιήθηκε

Το παιχνίδι που χρησιμοποιήθηκε για την εκπόνηση της διπλωματικής αυτής εργασίας ονομάζεται “Express Cooking Train”, το οποίο στοχεύει στη βελτίωση των δεξιοτήτων διατροφικής και μαγειρικής παιδείας του παίκτη. Προσφέρει ένα διαδραστικό περιβάλλον προσομοίωσης μαγειρικής, βασισμένο στη μέθοδο "δοκιμής και σφάλματος", ενώ ενσωματώνει μια οντολογία συνταγών. Το παιχνίδι έχει σχεδιαστεί με βάση επιστημονικές θεωρίες αλλαγής συμπεριφοράς, ακολουθώντας μια προσέγγιση που βασίζεται σε θεωρίες. Επιπλέον, προτείνεται ένα καινοτόμο εννοιολογικό πλαίσιο που έχει στόχο να αυξήσει τη συμμετοχή του χρήστη, ενώ παράλληλα ενδυναμώνει τις δεξιότητές του στη διατροφή και τη μαγειρική [8]. Από το αναφερθέν παιχνίδι συλλέχτηκαν οι καταγραφές βίντεο των συμμετεχόντων, που εκμεταλλεύτηκε η παρούσα διπλωματική εργασία.



Εικόνα 8: “Express Cooking Train”

3.1.2 Καταγραφή επιπέδων προσήλωσης

Η καταγραφή της προσήλωσης σαν διαδικασία δεν έχει κάποια state-of-the-art τεχνική που χρησιμοποιείται ευρέως σε πολλές εφαρμογές στη βιβλιογραφία. Στο συγκεκριμένο πρόβλημα χρησιμοποιήθηκε ένα εργαλείο που βασίστηκε στο “RankTrace” [30] με σκοπό να αποδοθούν οι δυναμικές αλλαγές στην προσήλωση των παικτών σε πραγματικό χρόνο κατά τη διάρκεια του παιχνιδιού. Το εργαλείο αυτό, το οποίο δημιουργήθηκε με το GameMaker Studio, επιτρέπει την συνεχή σημείωση της προσήλωσης, καθώς ενώ οι παίκτες παρακολουθούν καταγραφές των παιχνιδιών τους χρησιμοποιούν τη ροδέλα του ποντικιού για να συλλέγεται ένα δείγμα ανά δευτερόλεπτο. Μέσω αυτής της προσέγγισης, οι παίκτες παράγουν ένα συνεχές

ίχνος, το οποίο δίνει μια συνολική εικόνα για την αντιληπτή εμπλοκή τους με το παιχνίδι με τη επιφύλαξη της υποκειμενικότητας των συμμετεχόντων. Η διαδικασία αυτή πραγματοποιείται αμέσως μετά την αλληλεπίδραση του παίκτη με το παιχνίδι, με σκοπό να ληφθεί το καλύτερο δείγμα που είναι δυνατόν να γίνει με αυτή τη μέθοδο [6].

3.1.3 Οι συμμετέχοντες

Δεκαεννέα άνδρες και επτά γυναίκες, κυρίως προπτυχιακοί και μεταπτυχιακοί φοιτητές του Εθνικού Μετσόβιου Πολυτεχνείου (ΕΜΠ), αποτέλεσαν το σύνολο των 26 συμμετεχόντων που επιλέχθηκαν. Με εξαίρεση δύο που ήταν κάπως παχύσαρκοι (30-35 ετών) και δύο που ήταν ελαφρώς υπέρβαροι (25-27 ετών), κανένα από τα άτομα δεν εμφάνισε εμφανή κινητική ή οπτική αναπηρία και η πλειοψηφία των συμμετεχόντων είχε φυσιολογικό ΔΜΣ (18,5-25 ετών). Αφού οι συμμετέχοντες κάθισαν άνετα έπαιξαν το συγκεκριμένο παιχνίδι και αμέσως μετά χρησιμοποίησαν τη μέθοδο σημείωσης της προσήλωσης με τη ροδέλα του ποντικιού για την καταγραφή και αυτών των δεδομένων. Η κωδικοποίηση των συμμετεχόντων είχε την παρακάτω μορφή (ο ένας συμμετέχοντας έχει αφαιρεθεί καθώς δεν μπόρεσε να ανακτηθεί η καταγραφή του):

Participant	Participant Code	Participant	Participant Code
Participant 1	1	Participant 14	14
Participant 2	2	Participant 15	15
Participant 3	3	Participant 16	16
Participant 4	4	Participant 17	17
Participant 5	5	Participant 18	18
Participant 6	6	Participant 19	19
Participant 7	7	Participant 20	20
Participant 8	8	Participant 21	21
Participant 9	9	Participant 22	22
Participant 10	10	Participant 23	23
Participant 11	11	Participant 24	24
Participant 12	12	Participant 25	25
Participant 13	13		

Πίνακας 1: Κωδικοποίηση Συμμετεχόντων

3.2 Ανάλυση και Προεπεξεργασία των Δεδομένων

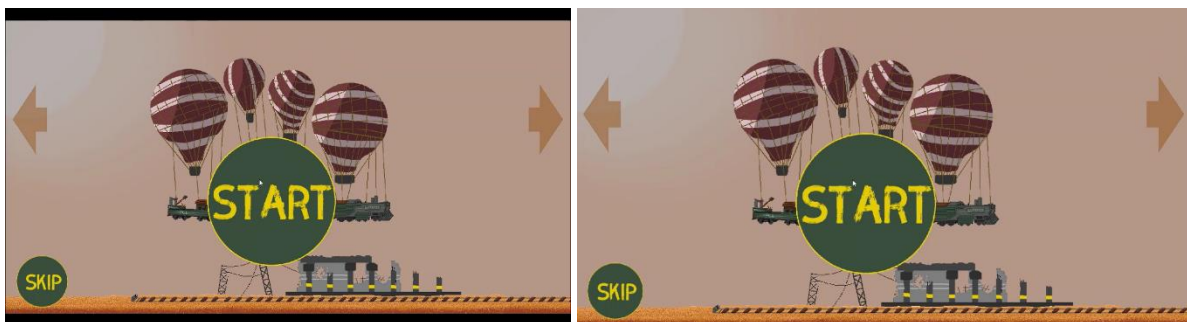
3.2.1 Δεδομένα καταγραφών βίντεο

Τα βίντεο που χρησιμοποιήθηκαν για την ανάλυση της παρούσας διπλωματικής καταγράφηκαν κατά τη διάρκεια της αλληλεπίδρασης των συμμετεχόντων με το παιχνίδι. Οι βιντεοσκοπήσεις αυτές πραγματοποιήθηκαν με ρυθμό 30 καρέ ανά δευτερόλεπτο και ανάλυση καρέ 1680 x 872, επιτρέποντας την λήψη λεπτομερών δεδομένων του κάθε παίκτη και την περαιτέρω επεξεργασία των βίντεο. Αρχικό βήμα στην ανάλυση των βίντεο ήταν η μείωση της κλίμακας των βίντεο καθώς το μέγεθος τους ήταν μεγάλο και επιβάρυνε τη μνήμη και έτσι

έγινε η μείωση της κλίμακας των καρτέ των βίντεο από 1680 x 872 σε 852 x 480 με τη χρήση του εργαλείου Microsoft Clipchamp. Επιπλέον τα δεδομένα προσήλωσης δεν ήταν ακριβώς συγχρονισμένα ανά δευτερόλεπτο με τα αρχικά βίντεο δεδομένα, καθώς σε κάθε συμμετέχοντα η διάρκεια των καταγραφών είναι συγκριτικά ελαφρώς μεγαλύτερη σε διάρκεια. Για το λόγο αυτό, έγινε περικοπή στη διάρκεια των βίντεο έτσι ώστε να γίνει η σωστή αντιστοίχιση με τη διάρκεια των δεδομένων προσήλωσης, ξανά με τη χρήση του Microsoft Clipchamp.

Τα δεδομένα των καταγραφών βίντεο διαβάστηκαν με τη χρήση της βιβλιοθήκης cv2 της python. Το επόμενο βήμα ήταν να αναλυθούν τα δεδομένα από αυτή τη μορφή σε σειριακά καρτέ κατάλληλης χρονικής απόστασης. Λόγω των τιμών της προσήλωσης που είναι μία ανά δευτερόλεπτο, αποθηκεύτηκε ένα από τα τριάντα καρτέ ανά δευτερόλεπτο, έτσι ώστε να γίνει η αντιστοίχιση των τιμών της προσήλωσης και των καρτέ σε ένα προς ένα. Αυτή η διαδικασία επιτεύχθηκε με μια σειρά από συναρτήσεις και για τον κάθε συμμετέχοντα ξεχωριστά, με σκοπό την ευκολότερη ανάλυση στη συνέχεια.

Στη συνέχεια ακολουθήθηκε επεξεργασία των καρτέ, έτσι ώστε να μπορέσουν να χρησιμοποιηθούν στις απαιτήσεις των μοντέλων. Παρατηρήθηκε ότι εμφανίζονται μαύρες μπάρες στα καρτέ από τα καταγεγραμμένα βίντεο, οι οποίες μπορεί να ζημιώσουν την ποιότητα εκπαίδευσης. Για αυτό το λόγο γίνεται μία αρχική περικοπή στα καρτέ, έτσι ώστε να αφαιρεθούν οι μαύρες μπάρες και να μην επηρεαστούν αρνητικά τα μοντέλα. Επιπλέον, συρρικνώθηκαν τα καρτέ σε τελικό μέγεθος 170 x 88, με σκοπό την μείωση του μεγέθους των δεδομένων και συνεπώς την επίτευξη της ομαλής και απροβλημάτιστης εκπαίδευσης. Στη βιβλιογραφία αναφέρεται συχνά πως σε μοντέλα που χρησιμοποιούν δεδομένα εικόνων, συνίσταται η χρήση τετραγωνικών διαστάσεων για την επίτευξη ομοιομορφίας των δεδομένων. Στη συγκεκριμένη περίπτωση όμως καθώς υπάρχει ήδη ομοιομορφία, διότι όλα τα καρτέ έχουν ακριβώς την ίδια δομή, κρίθηκε πως μια τέτοια προσέγγιση δεν είναι απαραίτητη και προτιμήθηκε η ανάλυση να συντηρηθεί σε μια αναλογία διαστάσεων παρόμοια με την αρχική μορφή των βίντεο.



Εικόνα 9: Καρτέ του ΠΣΣΥ με και χωρίς μαύρες μπάρες αντίστοιχα

3.2.2 Δεδομένα προσήλωσης

Τα δεδομένα προσήλωσης είναι καταγεγραμμένα σε αρχεία .txt για τον κάθε συμμετέχοντα ξεχωριστά, τα οποία έχουν τη δομή της μίας τιμής ανά συγκεκριμένη χρονική στιγμή. Η τιμή της προσήλωσης παίρνει ακέραιες τιμές, οι οποίες καθορίζονται αναλόγως με το πώς κυλούσε τη ροδέλα του ποντικιού καταγραφής ο κάθε παίκτης. Παρακάτω φαίνεται μία μορφή των τιμών αυτών:

Engagement Value	Time Stamp
9	15:38.47
6	15:38.48
4	15:38.49
4	15:38.50
9	15:38.51

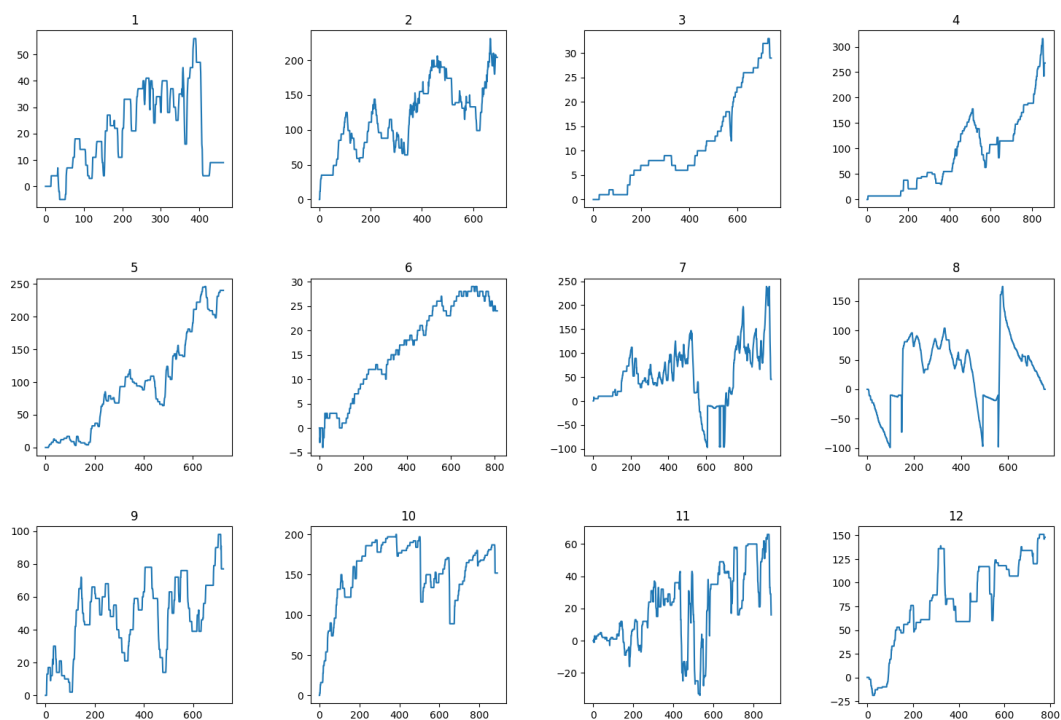
Πίνακας 2: Μορφή Δεδομένων Προσήλωσης των Συμμετεχόντων

Όπως αναλύεται και στο επόμενο κεφάλαιο της ενότητας αυτής, επιλέγονται επικαλυπτόμενα χρονικά παράθυρα ανάλυσης 20 δευτερολέπτων ως δεδομένα για την πειραματική διαδικασία. Σε αυτά τα 20 δευτερόλεπτα υπάρχουν 20 τιμές της προσήλωσης του κάθε συμμετέχοντα που κάθε μία αντιστοιχεί και σε ένα δευτερόλεπτο. Υπολογίζεται ο μέσος όρος των πρώτων 10 τιμών και των υπόλοιπων 10 ξεχωριστά και στη συνέχεια συγκρίνονται μεταξύ τους. Όταν η τιμή των πρώτων 10 είναι μεγαλύτερη από των υπολοίπων ορίζεται η τιμή της ετικέτας ως 0 ενώ στην ανάποδη περίπτωση ορίζεται ως 1. Ουσιαστικά, υπολογίζεται η αύξηση ή η μείωση της τιμής της προσήλωσης σε ένα συγκεκριμένο παράθυρο από 20 καρέ κάθε φορά. Με αυτόν τον τρόπο φτιάχνεται μία κατανομή στις ετικέτες των δεδομένων του κάθε συμμετέχοντα που φαίνεται στον παρακάτω πίνακα:

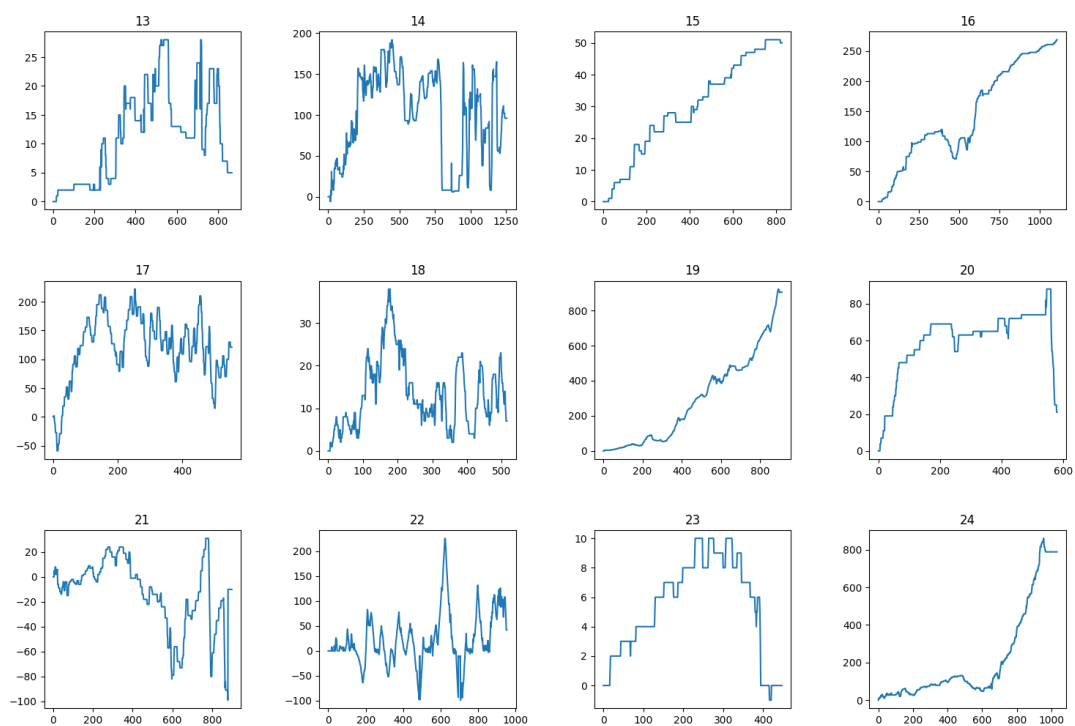
Participant	Amount of 0	Amount of 1
1	210	235
2	329	347
3	322	401
4	423	423
5	303	404
6	288	507
7	456	474
8	489	250
9	374	331
10	259	612
11	392	479
12	389	370
13	536	313
14	588	649
15	420	389
16	279	810
17	229	306
18	266	231
19	207	687
20	294	268
21	442	439
22	496	439
23	235	194
24	361	653
25	368	427

Πίνακας 3: Κατανομή του μέσου όρου της προσήλωσης των επικαλυπτόμενων παραθύρων καρτέ (μέση τιμή πριν - μέση τιμή μετά) $\geq 0 \rightarrow 0$ | (μέση τιμή πριν - μέση τιμή μετά) $< 0 \rightarrow 1$

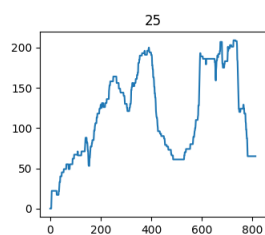
Στη συνέχεια έγινε ανάλυση της κατανομής των τιμών της προσήλωσης των συμμετεχόντων με σκοπό την καλύτερη κατανόηση των δεδομένων, όπως φαίνεται στα παρακάτω σχήματα:



Εικόνα 10: Διάγραμμα κατανομής της μετρικής της προσήλωσης για τους πρώτους 12 συμμετέχοντες



Εικόνα 11: Διάγραμμα κατανομής της μετρικής της προσήλωσης για τους επόμενους 12 συμμετέχοντες



Εικόνα 12: Διάγραμμα κατανομής της μετρικής της προσήλωσης για τον τελευταίο συμμετέχοντα

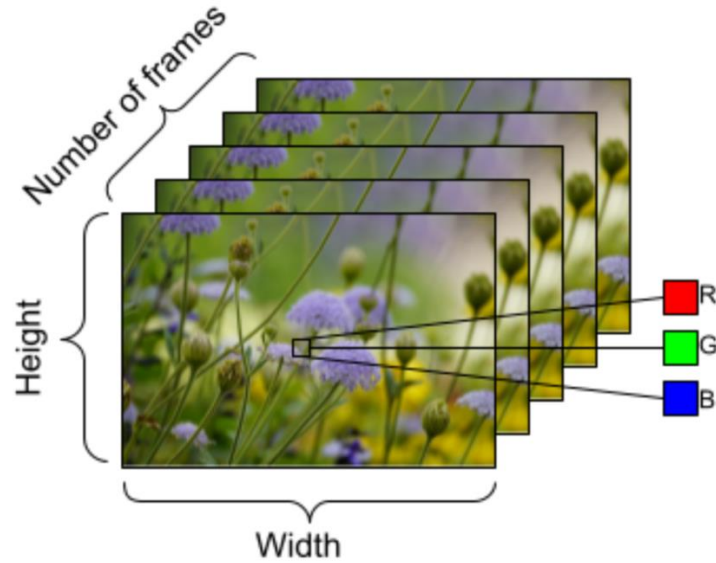
Από τα παραπάνω μπορεί να βγει το συμπέρασμα ότι κάποιοι συμμετέχοντες μπορεί να είναι προβληματικοί είτε λόγω της έντονης ανισορροπίας των δεδομένων τους στις τιμές ετικετών 0 και 1, είτε γιατί οι τιμές της προσήλωσης εμφανίζουν μια συνεχόμενη αύξηση. Ένα ακόμη πρόβλημα που εμφανίστηκε στις τιμές της προσήλωσης 2 συμμετεχόντων είναι η μικρή διακύμανση των τιμών προσήλωσης, δηλαδή της μικρής ευαισθησίας που παρουσιάζουν τα δεδομένα, κάτι που μπορεί να φανεί και από τα διαγράμματα των τιμών της προσήλωσης. Τα παραπάνω λαμβάνονται υπόψη στο κομμάτι της εκπαίδευσης και αφαιρούνται από το σύνολο των δεδομένων σε κάποιες από τις πειραματικές διαδικασίες. Στον παρακάτω πίνακα αναλύονται οι πιθανά προβληματικοί συμμετέχοντες:

Participant	Reason of Exclusion
3	Low Fluctuation on Small Values
4	Continuous Rise of Values
5	Continuous Rise of Values
6	Bad Distribution of 0/1
8	Bad Distribution of 0/1
10	Bad Distribution of 0/1
13	Bad Distribution of 0/1
15	Continuous Rise of Values
16	Bad Distribution of 0/1
19	Bad Distribution of 0/1
20	Continuous Rise of Values
23	Low Fluctuation on Small Values
24	Bad Distribution of 0/1

Πίνακας 4: Οι συμμετέχοντες που αφαιρέθηκαν από ένα μερίδιο των πειραμάτων και οι λόγοι για τον καθένα από αυτούς

3.2.3 Συσχέτιση των δεδομένων καταγραφής και προσήλωσης

Επιλέχθηκαν επικαλυπτόμενα παράθυρα στα καρτέ με σκοπό τη βέλτιστη χρήση των δεδομένων βίντεο και την καλύτερη απεικόνιση της σειριακής εξάρτησης τους. Το παράθυρο που επιλέχτηκε είναι τα 20 καρτέ ανά μία τιμή ετικέτας και όπως αναλύθηκε στο προηγούμενο κεφάλαιο οι ετικέτες έχουν δυαδική μορφή. Για τον συνδυασμό των καρτέ πλέον δεδομένων και των ετικετών κατασκευάστηκε μία κλάση, η οποία συσχετίζει μια εικοσάδα καρτέ με την υπολογισμένη τιμή ετικέτας αυτών των καρτέ από το αντίστοιχο αρχείο με τις τιμές της προσήλωσης. Η συγκεκριμένη διαδικασία επαναλαμβάνεται για όλα τα καρτέ όλων των παικτών που συμμετέχουν κάθε φορά στο πείραμα που διεξάγεται, με αποτέλεσμα να κατασκευάζεται το τελικό dataset το οποίο και διαχωρίζεται στη συνέχεια στα αντίστοιχα training, validation (όπου υπάρχει) και test set.



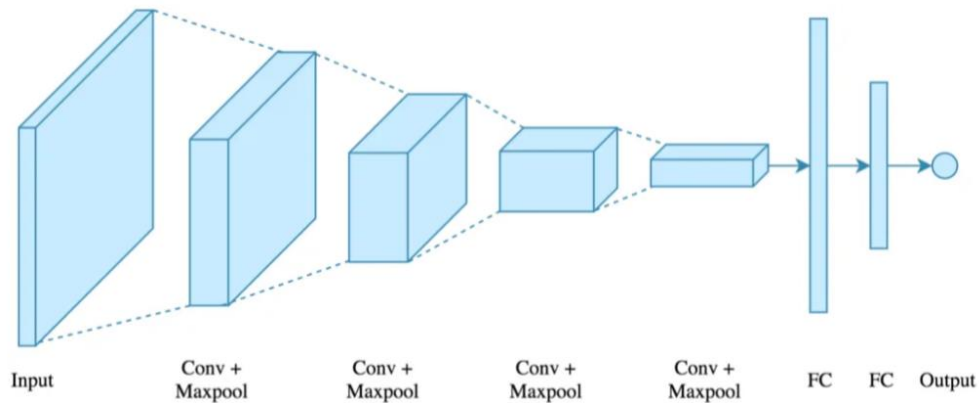
Εικόνα 13: Παράδειγμα σχήματος δεδομένων βίντεο (σειριακά δεδομένα)

3.3 Μέθοδοι Βαθιάς Μηχανικής Μάθησης (DL)

3.3.1 Ανάλυση αλγορίθμων DL που χρησιμοποιήθηκαν

Στην παρούσα διπλωματική εργασία, εφαρμόστηκαν αλγόριθμοι βαθιάς μηχανικής μάθησης μέσω του tensorflow/keras με σκοπό τη βέλτιστη λύση του προβλήματος της ταξινόμησης (classification). Συγκεκριμένα, χρησιμοποιήθηκαν συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Networks, CNN) συνεργατικά με ένα δίκτυο μακροπρόθεσμης-βραχυπρόθεσμης μνήμης (Long Short-Term Memory, LSTM), το οποίο αποτελεί εξελιγμένη μορφή των επαναλαμβανόμενων νευρωνικών δικτύων (Recurrent Neural Networks, RNN). Τα CNN εξειδικεύονται στην επεξεργασία χωρικών δεδομένων, όπως εικόνες, ενώ τα LSTM είναι ιδανικά για την ανάλυση χρονικών ακολουθιών, επιτρέποντας την κατανόηση της δυναμικής και των μοτίβων της συμπεριφοράς σε βάθος χρόνου. Η συνδυαστική χρήση αυτών των τεχνικών εξασφάλισε την αποδοτική ανάλυση των χαρακτηριστικών των βίντεο και την ακριβή εκτίμηση της προσήλωσης των συμμετεχόντων σε πραγματικό χρόνο.

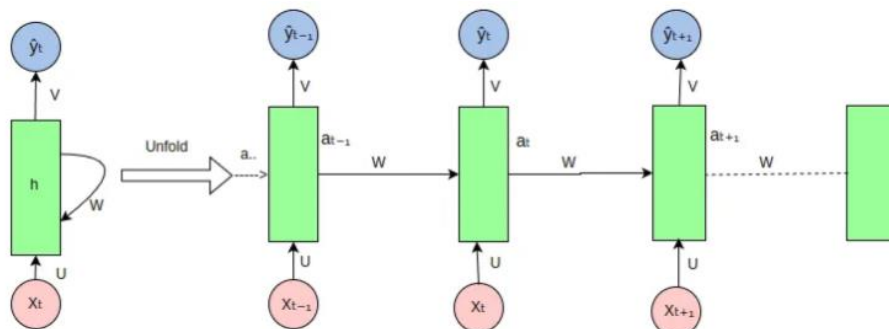
Η βασική αρχή λειτουργίας των CNN είναι η χρήση συνελκτικών φίλτρων (convolutional filters), τα οποία ανιχνεύουν σημαντικά χωρικά χαρακτηριστικά σε μικρότερες περιοχές των δεδομένων εισόδου. Ουσιαστικά, τα φίλτρα αυτά σαρώνουν την εικόνα, εντοπίζοντας μοτίβα όπως άκρα, γωνίες ή πιο σύνθετες δομές, σε διαφορετικά επίπεδα του δικτύου. Καθώς προχωρούν τα επίπεδα του CNN, τα χαρακτηριστικά γίνονται όλο και πιο αφηρημένα και αντιπροσωπευτικά των αρχικών δεδομένων. Μετά τη συνελκτική διαδικασία, τα χαρακτηριστικά μεταφέρονται σε πλήρως συνδεδεμένα επίπεδα (fully connected layers), τα οποία χρησιμοποιούνται για την τελική κατηγοριοποίηση ή πρόβλεψη. Η προσέγγιση αυτή καθιστά τα CNN ιδιαίτερα αποδοτικά στη μοντελοποίηση πολύπλοκων δεδομένων εικόνας, με τη δυνατότητα να μαθαίνουν αυτόματα σημαντικά χαρακτηριστικά χωρίς ανθρώπινη παρέμβαση [31].



A CNN works very similar to how our human eye works. The core operations

Εικόνα 14: Παράδειγμα αρχιτεκτονικής CNN

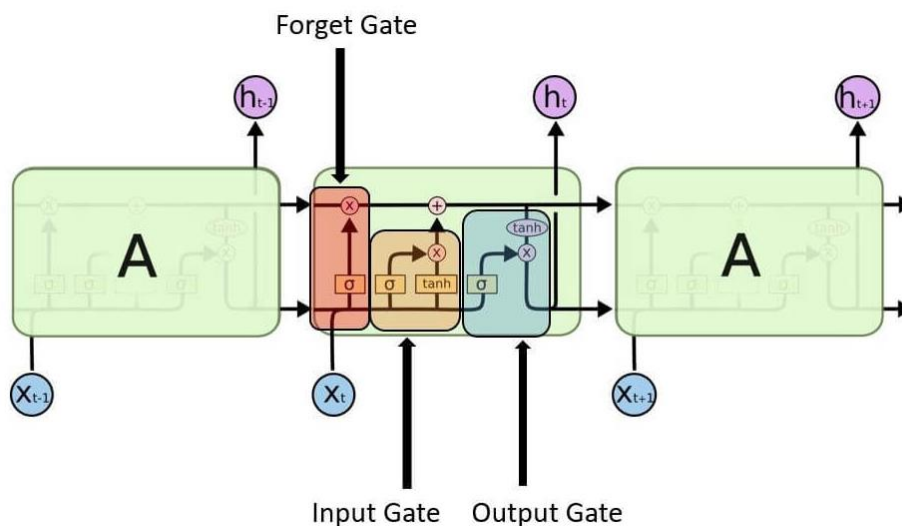
Από την άλλη τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) είναι ένα άλλο εξειδικευμένο σχήμα αρχιτεκτονικών νευρωνικών δικτύων. Τα RNN αναπτύσσονται για την επίλυση μαθησιακών προβλημάτων όπου οι παρελθοντικές πληροφορίες συνδέονται άμεσα με την πραγματοποίηση μελλοντικών προβλέψεων. Τέτοια διαδοχικά παραδείγματα παίζονται συχνά σε πολλές εργασίες του πραγματικού κόσμου, όπως η μοντελοποίηση γλώσσας, όπου οι προηγούμενες λέξεις στην πρόταση χρησιμοποιούνται για να προσδιοριστεί ποια θα είναι η επόμενη λέξη ή και στην πρόβλεψη της χρηματιστηριακής αγοράς όταν οι τιμές των μετοχών της τελευταίας περιόδου καθορίζουν τη μελλοντική κίνηση τους. Συνεπώς τα RNN είναι κατάλληλα να χρησιμοποιούνται σε δεδομένα από χρονοσειρές ή διαδοχικές εργασίες. Επιπλέον τα RNN είναι ιδανικά για την επεξεργασία 1-D διαδοχικών δεδομένων, σε αντίθεση με τα δεδομένα εικόνας 2-D που μοιάζουν με πλέγμα στα CNN [32].



parameters, represented by u , v , w is used consistently throughout the

Εικόνα 15: Παράδειγμα αρχιτεκτονικής RNN

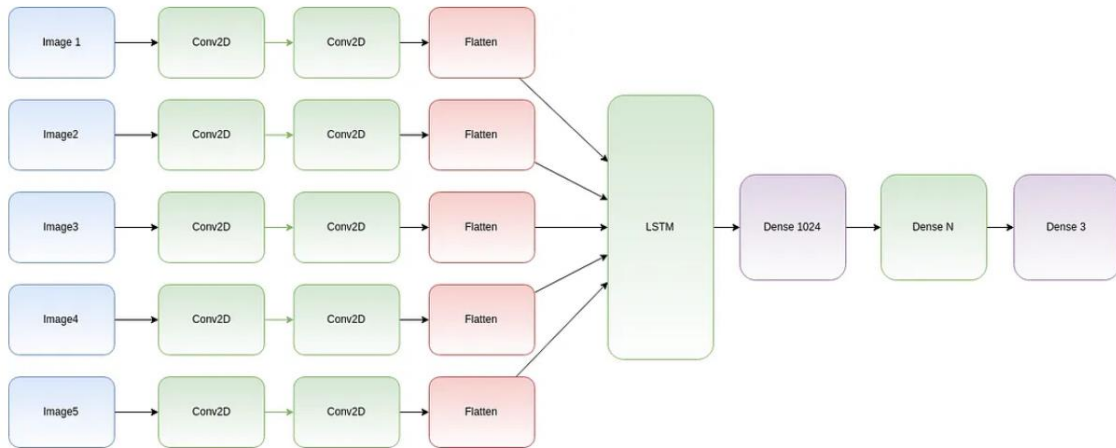
Τα δίκτυα μακροπρόθεσμης-βραχυπρόθεσμης μνήμης (LSTM) δίκτυα αποτελούν μια εξελιγμένη μορφή των βασικών RNN, σχεδιασμένη ειδικά να αντιμετωπίζει τις αδυναμίες τους, ιδίως όσον αφορά την αποθήκευση και αξιοποίηση πληροφοριών μακροπρόθεσμης εξάρτησης. Σε αντίθεση με τα απλά RNN, τα οποία δυσκολεύονται να διατηρήσουν σημαντικές πληροφορίες καθώς αυξάνεται το χρονικό εύρος των δεδομένων, τα LSTM ενσωματώνουν μηχανισμούς πυλών που ελέγχουν τη ροή πληροφορίας στο εσωτερικό τους. Οι πύλες αυτές επιτρέπουν στα LSTM να αποφασίζουν ποια πληροφορία θα αποθηκευτεί, ποια θα ξεχαστεί, και ποια θα χρησιμοποιηθεί για την παραγωγή της εξόδου, αντιμετωπίζοντας το πρόβλημα της εξαφάνισης ή έκρηξης λόγω των βαθμίδων που παρατηρείται στα παραδοσιακά RNN. Ως εκ τούτου, τα LSTM υπερέχουν στην ανάλυση δεδομένων με μακροχρόνιες εξαρτήσεις, καθιστώντας τα ιδανικά για εφαρμογές όπως η επεξεργασία φυσικής γλώσσας, η ανάλυση σειρών χρόνου και η πρόβλεψη ακολουθιακών γεγονότων, όπου οι βασικές μονάδες των RNN αποτυγχάνουν να αποδώσουν επαρκώς [32].



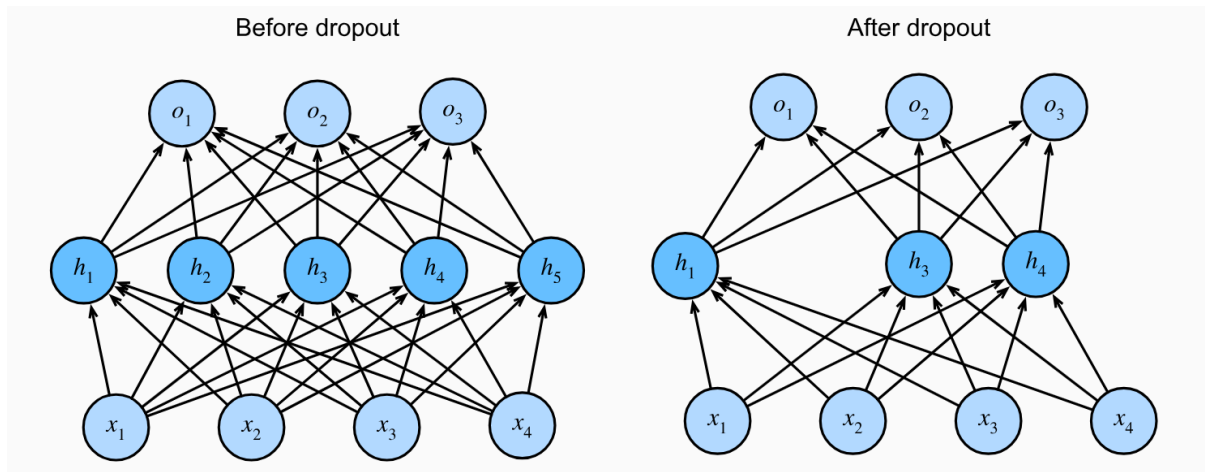
Εικόνα 16: Παράδειγμα αρχιτεκτονικής LSTM

3.3.2 Το μοντέλο που χρησιμοποιήθηκε

Το μοντέλο που χρησιμοποιήθηκε τελικά αξιοποιεί τα δίκτυα αλγορίθμων που αναλύθηκαν παραπάνω. Αρχικά, το επίπεδο εισόδου διαμορφώνεται χρησιμοποιώντας ένα περιτύλιγμα TimeDistributed σε κάθε επίπεδο αρχιτεκτονικής CNN, λόγω της μετάδοσης 20 εικόνων ως μία μόνο είσοδο. Με αυτό τον τρόπο προστίθενται 2 επίπεδα CNN και στη συνέχεια η έξοδος από το σύμπλεγμα αυτό εισάγεται σε ένα επίπεδο LSTM, του οποίου η έξοδος ενώνεται με ένα fully connected επίπεδο και τέλος με ένα μονό κόμβο, από τον οποίο προκύπτει το αποτέλεσμα της πρόβλεψης (0 ή 1). Τέλος, προστίθεται και ένα επίπεδο dropout, το οποίο έχει ως σκοπό την μείωση της υπερπροσαρμογής (overfit) του μοντέλου. Η λειτουργία του επιπέδου αυτού βασίζεται στην τυχαία επιλογή ενός κλάσματος νευρώνων που θα μηδενιστεί κατά τη διάρκεια κάθε εκπαιδευτικής συνεδρίας.



Εικόνα 17: Οπτική αναπαράσταση της αρχιτεκτονικής που χρησιμοποιήθηκε (τυχαίες τιμές)

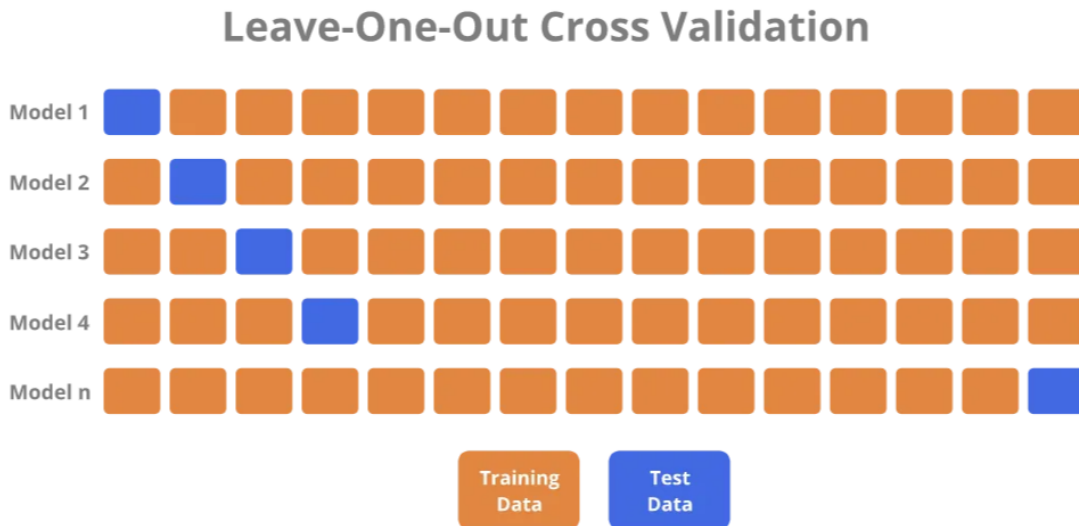


Εικόνα 18: Παράδειγμα λειτουργίας βαθμίδας dropout

3.3.3 Τεχνική Leave-One-Out Cross Validation (LOOCV)

Ο αλγόριθμος "διασταυρούμενη επικύρωση με εξαγωγή ενός δείγματος" ("leave-one-out-cross-validation (LOOCV)") είναι μια παραλλαγή της διασταυρούμενης επικύρωσης που χρησιμοποιείται για την αξιολόγηση της απόδοσης μοντέλων μηχανικής μάθησης. Στον LOOCV, το σύνολο δεδομένων χωρίζεται σε n υποσύνολα, όπου n είναι το πλήθος των δειγμάτων, όπου στο πρόβλημα της παρούσας διπλωματικής κάθε δείγμα είναι ένας συμμετέχοντας. Με αυτόν τον τρόπο σε κάθε επανάληψη, ένας συμμετέχοντας χρησιμοποιείται ως σύνολο ελέγχου, ενώ οι υπόλοιποι χρησιμοποιούνται για την εκπαίδευση του μοντέλου. Η διαδικασία αυτή επαναλαμβάνεται n φορές, ώστε κάθε δείγμα να χρησιμοποιηθεί μία φορά ως σύνολο ελέγχου. Η τελική απόδοση του μοντέλου υπολογίζεται ως ο μέσος όρος των σφαλμάτων από όλες τις επαναλήψεις. Αυτός ο αλγόριθμος προσφέρει μια αμερόληπτη εκτίμηση της απόδοσης του μοντέλου, καθώς κάθε δείγμα αξιολογείται με βάση ένα μοντέλο που δεν το έχει δει κατά την εκπαίδευση. Ωστόσο, μπορεί να είναι υπολογιστικά δαπανηρό σε μεγάλα σύνολα δεδομένων, καθώς απαιτείται η εκπαίδευση του μοντέλου πολλές φορές, όμως, είναι ιδιαίτερα χρήσιμο όταν τα δεδομένα είναι περιορισμένα.

Στην παρούσα διπλωματική χρησιμοποιήθηκε μία επιπλέον παραλλαγή του παραπάνω αλγορίθμου, κατά την οποία τα n υποσύνολα των δεδομένων χωρίζονται σε τρεις κατηγορίες. Τα $n-2$ χρησιμοποιούνται ως το training set, το ένα χρησιμοποιείται ως το validation set και το τελευταίο ως το test set, το οποίο είναι ο συμμετέχων που θέλουμε να τεστάρουμε σε κάθε επανάληψη. Η επιλογή του συμμετέχοντα για το validation set γίνεται με τυχαία επιλογή σε κάθε επανάληψη του παραλλαγμένου αλγορίθμου.



Εικόνα 19: Αρχιτεκτονική LOOCV

3.3.4 Μετρικές

Για την αξιολόγηση των πειραμάτων που εφαρμόστηκαν στην παρούσα διπλωματική εργασία χρησιμοποιήθηκαν διάφορες μετρικές, όπως αναλύονται παρακάτω:

Η πρώτη μετρική είναι ο Confusion Matrix, ο οποίος αποτελεί έναν πίνακα που παρουσιάζει τις σωστές και λανθασμένες προβλέψεις που έκανε το μοντέλο ταξινόμησης. Ο πίνακας περιλαμβάνει τέσσερα βασικά στοιχεία: τα True Positives (TP), δηλαδή τα δείγματα που προβλέφθηκαν σωστά ως θετικά, τα True Negatives (TN), δηλαδή τα δείγματα που προβλέφθηκαν σωστά ως αρνητικά, τα False Positives (FP), που είναι τα δείγματα που προβλέφθηκαν λανθασμένα ως θετικά ενώ ήταν αρνητικά, και τα False Negatives (FN), που είναι τα δείγματα που προβλέφθηκαν λανθασμένα ως αρνητικά ενώ ήταν θετικά.

Η δεύτερη μετρική που χρησιμοποιήθηκε είναι η ακρίβεια (Accuracy), η οποία μετρά το ποσοστό των σωστών προβλέψεων σε σχέση με το σύνολο των δειγμάτων. Η ακρίβεια αποτελεί μια από τις πιο απλές μετρικές, αλλά μπορεί να είναι παραπλανητική όταν τα δεδομένα δεν είναι ισορροπημένα. Η ακρίβεια υπολογίζεται ως το σύνολο των σωστών προβλέψεων διαιρούμενο με το σύνολο των δειγμάτων. Υπολογίζεται ως:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Η τρίτη μετρική είναι η προσέγγιση ακρίβειας (Precision), η οποία αξιολογεί την ακρίβεια των θετικών προβλέψεων. Πρόκειται για το ποσοστό των σωστών θετικών προβλέψεων σε σχέση με το σύνολο των θετικών προβλέψεων που έγιναν από το μοντέλο. Αυτή η μετρική είναι ιδιαίτερα σημαντική σε περιπτώσεις όπου το κόστος των ψευδώς θετικών (FP) είναι υψηλό, όπως για παράδειγμα στα συστήματα ανίχνευσης απάτης. Υπολογίζεται ως:

$$Precision = \frac{TP}{TP+FP}$$

Η τέταρτη μετρική είναι η ευαισθησία (Sensitivity ή Recall), η οποία μετρά την ικανότητα του μοντέλου να εντοπίζει σωστά τα θετικά δείγματα. Πρόκειται για το ποσοστό των σωστών θετικών προβλέψεων σε σχέση με το σύνολο των πραγματικών θετικών δειγμάτων. Η ευαισθησία είναι κρίσιμη σε περιπτώσεις όπου το κόστος των ψευδώς αρνητικών (FN) είναι υψηλό, όπως για παράδειγμα σε ιατρικές διαγνώσεις. Υπολογίζεται ως:

$$Recall = \frac{TP}{TP+FN}$$

Η πέμπτη μετρική είναι η ειδικότητα (Specificity), που μετρά την ικανότητα του μοντέλου να αναγνωρίζει σωστά τα αρνητικά δείγματα. Η ειδικότητα είναι το ποσοστό των σωστών αρνητικών προβλέψεων σε σχέση με το σύνολο των πραγματικών αρνητικών δειγμάτων. Αυτή η μετρική είναι σημαντική σε περιπτώσεις όπου απαιτείται να ελαχιστοποιηθούν τα ψευδώς θετικά αποτελέσματα. Υπολογίζεται ως:

$$Specificity = \frac{TN}{TN+FP}$$

Τέλος, η έκτη μετρική είναι το F1 Score, το οποίο αποτελεί τον αρμονικό μέσο του Precision και του Recall. Το F1 Score παρέχει μια ισορροπημένη μέτρηση που συνδυάζει τόσο την ακρίβεια των προβλέψεων όσο και την ικανότητα του μοντέλου να αναγνωρίζει θετικά δείγματα. Αυτή η μετρική είναι ιδιαίτερα χρήσιμη όταν υπάρχει ανισορροπία στις κατηγορίες δεδομένων. Υπολογίζεται ως:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Κεφάλαιο 4: Αποτελέσματα

Στην παρούσα ενότητα, παρουσιάζονται τα αποτελέσματα που προέκυψαν από την εφαρμογή των αλγορίθμων βαθιάς μάθησης στην ανάλυση των δεδομένων βίντεο, με στόχο την αναγνώριση των επιπέδων προσήλωσης των συμμετεχόντων κατά τη διάρκεια της αλληλεπίδρασής τους το ΠΣΣΥ “Express Cooking Train”. Τα αποτελέσματα αξιολογούνται χρησιμοποιώντας ένα σύνολο μετρικών απόδοσης, όπως το Accuracy, το Precision, το Recall και το F1 Score, όπως αναλύθηκαν και παραπάνω, ώστε να αναλυθεί διεξοδικά η αποτελεσματικότητα του προτεινόμενου συστήματος. Επιπλέον, παρουσιάζονται συγκριτικά γραφήματα και πίνακες για την καλύτερη απεικόνιση των επιδόσεων του συνδυασμού των αλγορίθμων που χρησιμοποιήθηκαν, δίνοντας έμφαση στην ακριβή πρόβλεψη της εμπλοκής των χρηστών σε πραγματικό χρόνο. Η ανάλυση των αποτελεσμάτων επιτρέπει τη συζήτηση για τα πλεονεκτήματα, τις αδυναμίες και τις προκλήσεις που αντιμετωπίζει το σύστημα, προσφέροντας μια ολοκληρωμένη αποτίμηση της απόδοσης του μοντέλου.

Δεδομένου ότι το σύνολο των δεδομένων παρουσιάζει καλή ισορροπία μεταξύ των κατηγοριών, η κύρια μετρική που χρησιμοποιήθηκε για την αξιολόγηση της απόδοσης των αλγορίθμων είναι η ακρίβεια (Accuracy). Η ακρίβεια αποτελεί έναν κατάλληλο δείκτη αξιολόγησης σε περιπτώσεις όπου οι κατηγορίες είναι ισορροπημένες, καθώς υπολογίζει το ποσοστό των σωστών προβλέψεων στο σύνολο των δειγμάτων. Σε αντίθεση με άλλες μετρικές, όπως η Precision ή το Recall, οι οποίες είναι πιο κατάλληλες όταν οι κατηγορίες είναι άνισες, η ακρίβεια παρέχει μια συνολική εκτίμηση της απόδοσης του μοντέλου σε ισορροπημένα σύνολα δεδομένων, αποτυπώνοντας με ακρίβεια την ικανότητα του αλγορίθμου να προβλέπει σωστά τόσο τις θετικές όσο και τις αρνητικές περιπτώσεις. (κεφάλαιο 5)

4.1 Απλό LOOCV χωρίς τους συμμετέχοντες με τις προβληματικές μετρήσεις

Αρχικά, χρησιμοποιούνται τρεις διαφορετικές εκδοχές του μοντέλου CNN-LSTM με στόχο την αξιολόγηση της επίδρασης των υπερπαραμέτρων στην απόδοση του, που πραγματοποιείται με τη χρήση του απλού αλγορίθμου LOOCV, ο οποίος επιτρέπει τη λεπτομερή ανάλυση της απόδοσης σε σύνολα δεδομένων μικρού μεγέθους. Η βασική διαφορά μεταξύ των εκδοχών έγκειται στις υπερπαραμέτρους που χρησιμοποιούνται κατά την εκπαίδευση, όπως η επιλογή του αριθμού φίλτρων CNN, των μονάδων του LSTM, και του αριθμού νευρώνων στο πλήρως συνδεδεμένα επίπεδο. Ο σκοπός αυτής της προσέγγισης είναι να εξεταστεί πώς οι παραλλαγές στις παραμέτρους αυτές επηρεάζουν την ακρίβεια και τη γενικότερη απόδοση του μοντέλου, με έμφαση στη βελτίωση της εκμάθησης είτε με την αύξηση είτε με τη μείωση των παραμέτρων του. Σημειώνεται ότι αυτή η διαδικασία πειραματισμού αποτελεί το πρώτο βήμα των πειραμάτων, παρέχοντας σημαντικές πληροφορίες για τη βέλτιστη διαμόρφωση της διαδικασίας στα επόμενα. Όλες οι δοκιμές έγιναν με την εκπαίδευση να διατρέχει για 100 εποχές.

Τα τρία μοντέλα που χρησιμοποιήθηκαν σε αυτό το βήμα διαφέρουν κυρίως ως προς την αρχιτεκτονική των επιπέδων LSTM και Dense, καθώς και ως προς τον αριθμό των φίλτρων στα Conv2D επίπεδα. Το πρώτο μοντέλο περιλαμβάνει δύο Conv2D επίπεδα με 64 φίλτρα και LSTM με 128 νευρώνες, ενώ στο δεύτερο μοντέλο η κύρια διαφορά είναι η αύξηση των

νευρώνων του LSTM και του Dense σε 256, κάτι που πιθανώς βελτιώνει την ικανότητα του μοντέλου να μαθαίνει πιο σύνθετες σχέσεις από τα δεδομένα. Το τρίτο μοντέλο έχει μικρότερη χωρητικότητα, καθώς τα Conv2D επίπεδα έχουν 32 φίλτρα και το LSTM 128 νευρώνες, γεγονός που ενδεχομένως το καθιστά πιο απλό και λιγότερο υπολογιστικά απαιτητικό. Παρά την ίδια χρήση του optimizer ADAM με την ίδια ρυθμιστική παράμετρο μάθησης ($lr=3e-4$) και στα τρία μοντέλα, οι διαφορές στα αποτελέσματα είναι μικρές, όπως φαίνεται και παρακάτω:

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.6667	0.8533	0.3249	0.9532	0.4706
2	0.4762	0.4568	0.3872	0.5610	0.4191
7	0.4558	0.4406	0.4163	0.4937	0.4281
9	0.4773	0.5455	0.0804	0.9245	0.1402
11	0.5174	0.4500	0.3740	0.6326	0.4085
12	0.3923	0.4066	0.4204	0.3631	0.4134
14	0.5576	0.5450	0.3945	0.7042	0.4577
17	0.5966	0.5435	0.4367	0.7191	0.4843
18	0.5282	0.6115	0.3208	0.7662	0.4208
21	0.5989	0.5894	0.6576	0.5399	0.6217
22	0.4407	0.3247	0.0508	0.8807	0.0879
25	0.4145	0.3344	0.2885	0.5199	0.3098
Mean	0.5102	0.5084	0.3460	0.6715	0.3885

Πίνακας 5: Απλό LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 1**: $2 \times \text{Conv2D}(64, (3,3), \text{padding}='same', \text{strides}=(2,2), \text{relu}), \text{MaxPooling2D}((2,2), \text{strides}=(2,2)), \text{GlobalAveragePooling2D}(), \text{LSTM}(128, \text{relu}), \text{Dense}(128, \text{relu}), \text{Dense}(1, \text{sigmoid}), \text{Optimizer: ADAM} (lr=3e-4)$

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.6620	0.7383	0.4010	0.8809	0.5197
2	0.4107	0.3509	0.2439	0.5698	0.2878
7	0.4914	0.4804	0.4868	0.4958	0.4836
9	0.4929	0.5606	0.1984	0.8248	0.2931
11	0.5914	0.5734	0.3247	0.8058	0.4146
12	0.3949	0.4186	0.4830	0.3035	0.4485
14	0.5073	0.4731	0.3619	0.6379	0.4101
17	0.5663	0	0	1	0
18	0.4940	0.5427	0.3358	0.6753	0.4149
21	0.6318	0.6043	0.7687	0.4943	0.6766
22	0.4731	0.6000	0.0183	0.9862	0.0355
25	0.4145	0.4096	0.6471	0.2201	0.5016
Mean	0.5110	0.4793	0.3558	0.6579	0.3738

Πίνακας 6: Απλό LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 2**: $2 \times \text{Conv2D}(64, (3,3), \text{padding}='same', \text{strides}=(2,2), \text{relu}), \text{MaxPooling2D}((2,2), \text{strides}=(2,2)), \text{GlobalAveragePooling2D}(), \text{LSTM}(256, \text{relu}), \text{Dense}(256, \text{relu}), \text{Dense}(1, \text{sigmoid}), \text{Optimizer: ADAM} (lr=3e-4)$

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.6042	1.0000	0.1320	1.0000	0.2332
2	0.4509	0.4232	0.3445	0.5523	0.3798
7	0.5000	0.4900	0.5374	0.4641	0.5126
9	0.4943	0.5307	0.3941	0.6073	0.4523
11	0.5012	0.4596	0.6805	0.3570	0.5487
12	0.4920	0.5007	0.9138	0.0542	0.6470
14	0.5114	0.4782	0.3568	0.6502	0.4086
17	0.6307	0.5885	0.4934	0.7358	0.5368
18	0.5544	0.5866	0.5623	0.5455	0.5742
21	0.5602	0.5419	0.7914	0.3280	0.6433
22	0.4881	0.5914	0.1118	0.9128	0.1880
25	0.4656	0.2961	0.1261	0.7494	0.1768
Mean	0.5211	0.5406	0.4537	0.5797	0.4418

Πίνακας 7: Απλό LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 3:** `2 x Conv2D(32, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), Optimizer: ADAM (lr=3e-4)`

4.2 Παραλλαγή του LOOCV χωρίς τους Συμμετέχοντες με τις προβληματικές μετρήσεις

Στη συνέχεια, εφαρμόζεται μια παραλλαγή του αλγορίθμου Leave-One-Out Cross Validation (LOOCV), όπως αναλύεται παραπάνω, η οποία στοχεύει στη μείωση της μεροληψίας (bias) των μοντέλων και στην παροχή πιο αντιπροσωπευτικών αποτελεσμάτων. Αυτή η παραλλαγή χρησιμοποιεί έναν συμμετέχοντα ως σύνολο επικύρωσης (validation), έναν ως σύνολο δοκιμών (testing) και τους υπόλοιπους συμμετέχοντες για την εκπαίδευση (training), επιτρέποντας την καλύτερη αξιολόγηση της γενίκευσης των μοντέλων. Ο στόχος αυτής της μεθόδου είναι η χρήση του μηχανισμού Early Stopping για να αποφεύγεται η υπερπροσαρμογή (overfitting), διακόπτοντας την εκπαίδευση όταν η απόδοση στο σύνολο επικύρωσης σταματά να βελτιώνεται. Παράλληλα, γίνεται πειραματισμός με διάφορες παραλλαγές υπερπαραμέτρων, τόσο στη δομή του μοντέλου όσο και στις ρυθμίσεις του early stopping, με σκοπό τη βελτιστοποίηση της απόδοσης και την επίτευξη των βέλτιστων δυνατών αποτελεσμάτων.

Τα μοντέλα που χρησιμοποιήθηκαν σε αυτή τη σειρά πειραμάτων έχουν ξανά παρόμοια αρχιτεκτονική, με δύο επίπεδα Conv2D (32 και 64 φίλτρα) και ένα επίπεδο LSTM. Ωστόσο, οι διαφορές τους εντοπίζονται στις ρυθμίσεις των υπερπαραμέτρων, όπως η στρατηγική Early Stopping και το optimizer. Τα πρώτα δύο μοντέλα χρησιμοποιούν την ίδια αρχιτεκτονική και optimizer (ADAM με $lr=3e-4$), αλλά διαφέρουν στο patience του Early Stopping, με το πρώτο να σταματά την εκπαίδευση μετά από 5 περιόδους χωρίς βελτίωση και το δεύτερο μετά από 10. Το τρίτο μοντέλο, αντίστοιχα, μειώνει το μέγεθος του LSTM και του Dense από 128 σε 64 νευρώνες, ενώ διατηρεί την ίδια στρατηγική Early Stopping με το πρώτο μοντέλο.

Το τέταρτο μοντέλο χρησιμοποιεί διαφορετικό optimizer, τον SGD με Nesterov momentum, γεγονός που το καθιστά πιο συντηρητικό στη μάθηση σε σύγκριση με τον ADAM, και εφαρμόζει μεγαλύτερη διάρκεια εκπαίδευσης (60 epochs) με patience 20. Το τελευταίο μοντέλο είναι παρόμοιο με το τέταρτο, αλλά χρησιμοποιεί τον ADAM με μειωμένο learning rate ($1e-4$), κάτι που ενδέχεται να κάνει το μοντέλο πιο αργό στη σύγκλιση, αλλά σταθερότερο, διατηρώντας τις ίδιες ρυθμίσεις Early Stopping. Παρακάτω παρατίθενται τα πειράματα αυτής της μεθόδου, όπου ο συμμετέχοντας που αναγράφεται είναι το testing σύνολο, ενώ ο συμμετέχοντας για validation επιλέγεται τυχαία σε κάθε επανάληψη.

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.6505	0.6438	0.5228	0.7574	0.5770
2	0.5476	0.5375	0.5244	0.5698	0.5309
7	0.4677	0.4617	0.5308	0.4072	0.4939
9	0.4588	0.4897	0.5094	0.4018	0.4993
11	0.5150	0.4676	0.6364	0.4175	0.5391
12	0.4907	0	0	1	0
14	0.5162	0.4918	0.6672	0.3806	0.5662
17	0.5871	0.5224	0.5590	0.6087	0.5401
18	0.5444	0.5675	0.6189	0.4589	0.5921
21	0.5614	0.5495	0.6916	0.4305	0.6124
22	0.5065	0.5506	0.3760	0.6537	0.4469
25	0.5867	0.5559	0.4594	0.6932	0.5031
Mean	0.5361	0.4865	0.5080	0.5649	0.4917

Πίνακας 8: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 4**: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 30, patience 5)

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.5995	0.6429	0.2741	0.8723	0.3843
2	0.5625	0.5464	0.6098	0.5174	0.5764
7	0.4968	0.4891	0.6454	0.3544	0.5565
9	0.4830	0.5139	0.4450	0.5257	0.4770
11	0.5544	0	0	1	0
12	0.3936	0.4052	0.4073	0.3794	0.4063
14	0.5284	0.5022	0.3928	0.6502	0.4408
17	0.6004	0.5300	0.6943	0.5284	0.6011
18	0.5827	0.6208	0.5623	0.6061	0.5901
21	0.5898	0.6852	0.3356	0.8451	0.4505
22	0.5485	0.5782	0.5488	0.5482	0.5631
25	0.5587	0.5165	0.4818	0.6230	0.4986
Mean	0.5415	0.5025	0.4498	0.6209	0.4621

Πίνακας 9: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 5**: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 30, patience 10)

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.5463	0.5022	0.5787	0.5191	0.5377
2	0.5119	0	0	1	0
7	0.4720	0.4650	0.5264	0.4198	0.4938
9	0.4730	0.5027	0.4987	0.4441	0.5007
11	0.5544	0	0	1	0
12	0.3949	0.4053	0.4021	0.3875	0.4037
14	0.5089	0.4633	0.2384	0.7519	0.3148
17	0.5663	0	0	1	0
18	0.5927	0.6019	0.7019	0.4675	0.6481
21	0.4989	0	0	1	0
22	0.4698	0	0	1	0
25	0.5791	0.5458	0.4510	0.6862	0.4939
Mean	0.5140	0.2905	0.2831	0.7230	0.2827

Πίνακας 10: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 6**: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(64, relu), Dense(64, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 30, patience 5)

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.4468	0.3688	0.2995	0.5702	0.3305
2	0.5119	0	0	1	0
7	0.5108	0	0	1	0
9	0.3864	0.3870	0.2708	0.5166	0.3186
11	0.5810	0.5320	0.4961	0.6493	0.5134
12	0.4907	0	0	1	0
14	0.5268	0	0	1	0
17	0.4697	0.3289	0.2140	0.6656	0.2593
18	0.4657	0	0	1	0
21	0.4989	0	0	1	0
22	0.4698	0	0	1	0
25	0.5510	0.5145	0.2493	0.8033	0.3358
Mean	0.4925	0.1776	0.1275	0.8504	0.1465

Πίνακας 11: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το **Μοντέλο 7**: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer SGD(learning_rate=3e-4, momentum=0.9, nesterov=True), Early Stopping (from Epoch 60, patience 20)

Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.5579	0.5136	0.5736	0.5447	0.5420
2	0.4911	0.4794	0.4970	0.4855	0.4880
7	0.4838	0.4762	0.5507	0.4198	0.5107
9	0.4616	0.4919	0.4906	0.4290	0.4913
11	0.5	0.4601	0.7039	0.3361	0.5565
12	0.5519	0.5390	0.8303	0.2629	0.6536
14	0.5300	0.5074	0.2367	0.7935	0.3228
17	0.5739	0.5105	0.4236	0.6890	0.4630
18	0.5363	0.5732	0.5170	0.5584	0.5437
21	0.4989	0.5	0.2744	0.7244	0.3543
22	0.5409	0.5714	0.5366	0.5459	0.5535
25	0.5191	0.4569	0.2969	0.7049	0.3599
Mean	0.5205	0.5066	0.4943	0.5412	0.4866

Πίνακας 12: Παραλλαγή LOOCV για όλες τις τιμές μετρικών για το Μοντέλο 8: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=1e-4), Early Stopping (from Epoch 60, patience 20)

4.3 Εκπαίδευση κάθε Συμμετέχοντα ξεχωριστά

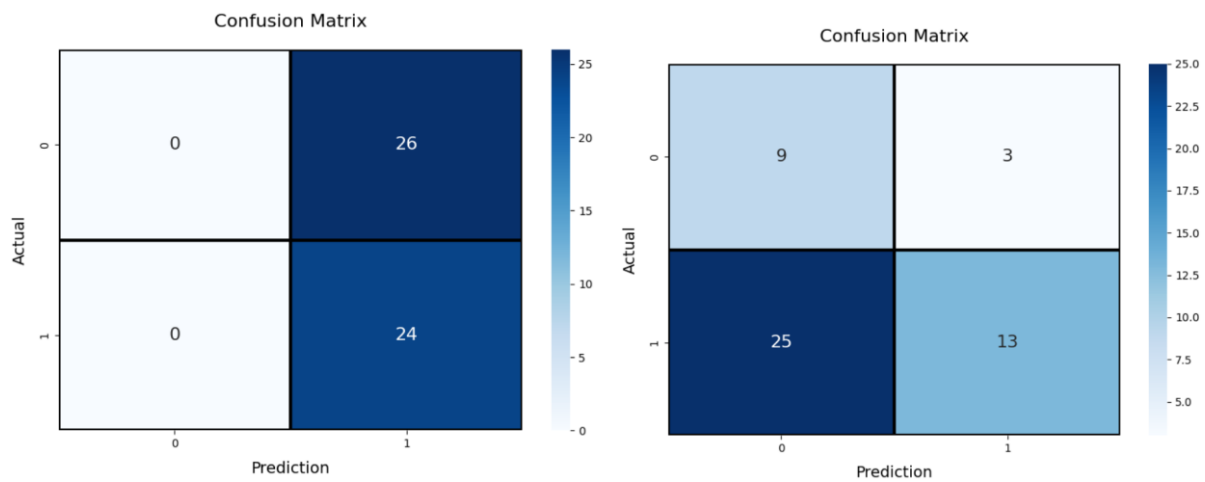
Επιπλέον, πραγματοποιείται ανάλυση κάθε συμμετέχοντα ξεχωριστά, με την εκπαίδευση ενός μοντέλου για το κάθε άτομο. Ο διαχωρισμός των δεδομένων σε σύνολα εκπαίδευσης (train) και δοκιμής (test) γίνεται επιλέγοντας τις τελευταίες 50 εικοσάδες καρτέ για το testing, ενώ οι εικοσάδες από τις θέσεις 51 έως 70 από το τέλος εξαιρούνται για να αποφευχθεί το bias λόγω επικάλυψης των καρτέ. Οι υπόλοιπες εικοσάδες καρτέ, μέχρι και την 71^η από το τέλος, χρησιμοποιούνται για την εκπαίδευση του κάθε μοντέλου. Αυτή η μεθοδολογία εφαρμόζεται για όλους τους συμμετέχοντες (25), όπου αξιολογούνται ξεχωριστά και παράγονται τα δικά του αποτελέσματα, επιτρέποντας μια πιο στοχευμένη και εξατομικευμένη ανάλυση της απόδοσης των μοντέλων.

Τα δύο μοντέλα που παρουσιάζονται σε αυτή την περίπτωση έχουν την ίδια βασική αρχιτεκτονική, ξεκινώντας με δύο επίπεδα Conv2D (32 και 64 φίλτρα), ακολουθούμενα από MaxPooling2D και GlobalAveragePooling2D. Η κύρια διαφορά μεταξύ των δύο έγκειται στο μέγεθος του LSTM και του Dense επιπέδου. Στο πρώτο μοντέλο, το LSTM έχει 128 μονάδες, ενώ στο δεύτερο μοντέλο, το LSTM αυξάνεται σε 256 μονάδες. Αντίστοιχα, το Dense επίπεδο έχει 128 και 256 μονάδες, αντίστοιχα, γεγονός που δίνει στο δεύτερο μοντέλο μεγαλύτερη χωρητικότητα μάθησης, αλλά και μεγαλύτερη υπολογιστική πολυπλοκότητα. Και τα δύο μοντέλα χρησιμοποιούν τον ίδιο optimizer (ADAM με learning rate 3e-4) και εφαρμόζουν Early Stopping από την 100^η περίοδο εκπαίδευσης, με patience 20, που τους επιτρέπει να σταματούν την εκπαίδευση εάν δεν υπάρχει βελτίωση μετά από 20 περιόδους.

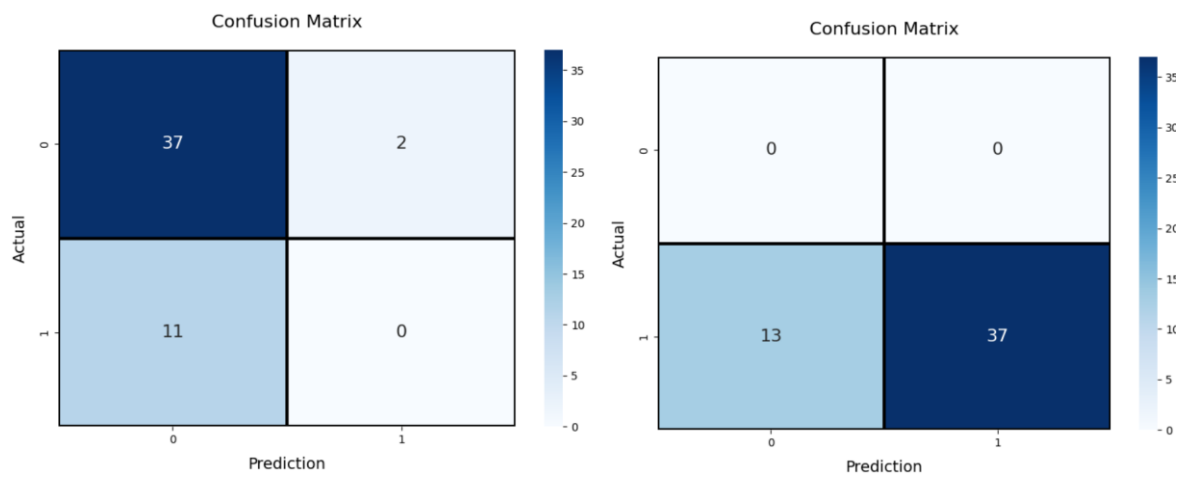
Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.38	0.4839	0.5	0	0.4918
2	0.3	0.3	1	0	0.4615
7	0.68	0	0	1	0
9	0.44	0.3226	0.5882	0.3636	0.4167
11	0.48	0.1429	0.2727	0.5385	0.1875
12	0.44	0.4400	1	0	0.6111
14	0.6	0	0	1	0
17	0.64	0	0	1	0
18	0.48	0.48	1	0	0.6486
21	0.42	0.42	1	0	0.5915
22	0.42	0.5	0.1724	0.7619	0.2564
25	0.12	0	0	1	0
3	0.92	0	0	1	0
4	0.88	0	0	1	0
5	0.52	0	0	1	0
6	0.44	0.8125	0.3421	0.75	0.4815
8	1	1	1	0	1
10	0.74	0	0	0.9487	0
13	1	1	1	0	1
15	0.74	1	0.74	0	0.8506
16	0.58	0	0	1	0
19	0.88	0	0	1	0
20	0.42	0.5455	0.5625	0.1667	0.5538
23	0.32	0.6667	0.2105	0.667	0.32
24	0.46	0	0	1	0
Mean	0.5720	0.3247	0.3755	0.5679	0.3148

Πίνακας 13: Ατομικό μοντέλο κάθε συμμετέχοντα: Όλες οι τιμές μετρικών για το **Μοντέλο 9:** Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(128, relu), Dense(128, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 100, patience 20)

Ακολουθούν και τα Confusion Matrices του κάθε συμμετέχοντα στο παραπάνω Μοντέλο 9:



Εικόνα 20: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 18, 6

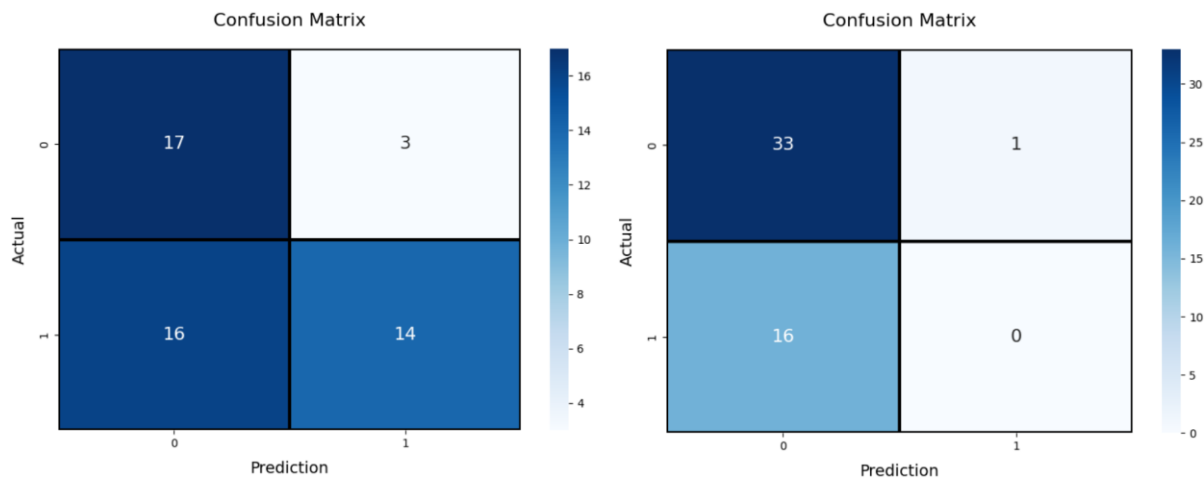


Εικόνα 21: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 10, 15

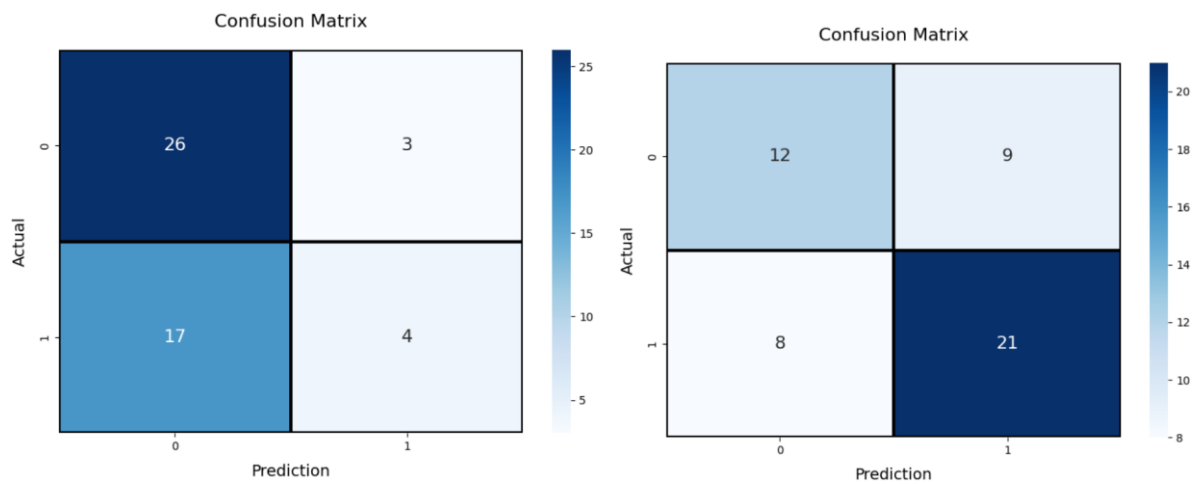
Participant	Accuracy	Precision	Recall	Specificity	F1 Score
1	0.62	0.8235	0.4667	0.85	0.5957
2	0.3	0.3	1	0	0.4615
7	0.66	0	0	0.9706	0
9	0.44	0.3590	0.8235	0.2424	0.5
11	0.78	0	0	1	0
12	0.44	0.44	1	0	0.6111
14	0.6	0	0	1	0
17	0.64	0	0	1	0
18	0.48	0.48	1	0	0.6486
21	0.6	0.5714	0.1905	0.8966	0.2857
22	0.66	0.7	0.7241	0.5714	0.7119
25	0.12	0	0	1	0
3	0.92	0	0	1	0
4	0.88	0	0	1	0
5	0.52	0	0	1	0
6	0.24	0	0	1	0
8	1	1	1	0	1
10	0.78	0	0	1	0
13	1	1	1	0	1
15	0.12	1	0.12	0	0.2143
16	0.58	0	0	1	0
19	0.88	0	0	1	0
20	0.36	0	0	1	0
23	0.34	0.6316	0.3158	0.4167	0.4211
24	0.46	0	0	1	0
Mean	0.5768	0.2922	0.3056	0.6779	0.2580

Πίνακας 14: Ατομικό μοντέλο κάθε συμμετέχοντα: Όλες οι τιμές μετρικών για το **Μοντέλο 10**: Conv2D(32, (3,3), padding='same', strides=(2,2), relu), Conv2D(64, (3,3), padding='same', strides=(2,2), relu), MaxPooling2D((2,2), strides=(2,2)), GlobalAveragePooling2D(), LSTM(256, relu), Dense(256, relu), Dense(1, sigmoid), optimizer ADAM(lr=3e-4), Early Stopping (from Epoch 100, patience 20)

Ακολουθούν και τα Confusion Matrices του κάθε συμμετέχοντα στο παραπάνω Μοντέλο 10:



Εικόνα 22: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 1, 7



Εικόνα 23: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 21, 22

Κεφάλαιο 5: Συζήτηση

Στο παρόν κεφάλαιο παρουσιάζεται η ανακεφαλαίωση της παρούσας Διπλωματικής Εργασίας, η οποία επικεντρώνεται στη χρήση βίντεο δεδομένων καταγραφής και την εφαρμογή μοντέλων βαθιάς μάθησης για την ανάλυση της συμπεριφοράς των συμμετεχόντων. Μετά τη θεωρητική ανάλυση μέσω βιβλιογραφικής έρευνας, παρουσιάστηκαν οι σύγχρονες τεχνικές ανάλυσης δεδομένων βίντεο και οι αρχές της βαθιάς μάθησης, με έμφαση στους αλγόριθμους CNN και LSTM, οι οποίοι χρησιμοποιήθηκαν στην παρούσα εργασία. Στη συνέχεια, εξετάστηκε η διαδικασία συλλογής και επεξεργασίας των δεδομένων καταγραφής βίντεο, καθώς και η δυνατότητα αξιοποίησης αυτών μέσω μοντέλων μηχανικής μάθησης για την ανίχνευση συμπεριφορικών προτύπων. Στο πρακτικό μέρος της διπλωματικής, πραγματοποιήθηκε η εκπαίδευση και αξιολόγηση πολλαπλών εκδοχών του μοντέλου CNN-LSTM, χρησιμοποιώντας ως κύρια μετρική την ακρίβεια (Accuracy), δεδομένου ότι τα δεδομένα ήταν καλά ισορροπημένα. Τα αποτελέσματα αναλύθηκαν με τη χρήση τεχνικών όπως το Leave-One-Out Cross Validation και της παραλλαγής του, καθώς και με μεμονωμένα μοντέλα ανά συμμετέχοντα, προκειμένου να επιτευχθούν βέλτιστες επιδόσεις και να αναδειχθούν οι περιορισμοί και τα πλεονεκτήματα των προτεινόμενων μοντέλων.

5.1 Στατιστική Ανάλυση στην εκπαίδευση των μοντέλων κάθε συμμετέχοντα ξεχωριστά

Με σκοπό την περεταίρω διερεύνηση των αποτελεσμάτων επιλέχθηκε να γίνει μια στατιστική ανάλυση στην περίπτωση της εκπαίδευσης ενός μοντέλου για κάθε συμμετέχοντα ξεχωριστά, στο Μοντέλο 9 του υποκεφαλαίου 4.3. Παρατηρείται πως η μέση ακρίβεια που προκύπτει είναι 57,2%, γεγονός που υποδεικνύει ότι το μοντέλο επιτυγχάνει μια συνολική απόδοση ανώτερη της τυχαίας πρόβλεψης (50%). Ωστόσο, ο βαθμός απόκλισης των επιδόσεων μεταξύ των συμμετεχόντων, όπως αντικατοπτρίζεται στη τυπική απόκλιση (23,12%) και τη διακύμανση (5,35%), καταδεικνύει ότι η απόδοση του μοντέλου παρουσιάζει σημαντικές διαφοροποιήσεις ανάλογα με τα χαρακτηριστικά του εκάστοτε ατόμου. Ορισμένοι συμμετέχοντες σημείωσαν ιδιαίτερα υψηλά επίπεδα ακρίβειας, ενώ άλλοι επέδειξαν αρκετά χαμηλή απόδοση.

Η κατασκευή ενός διαστήματος εμπιστοσύνης 95% για την ακρίβεια προσφέρει περαιτέρω πληροφορίες. Το εύρος του διαστήματος [47,7%, 66,7%] υποδηλώνει ότι με 95% βεβαιότητα η πραγματική μέση ακρίβεια του μοντέλου στον πληθυσμό βρίσκεται μεταξύ αυτών των τιμών. Αυτό το εύρος, ωστόσο, φανερώνει μια σχετική αβεβαιότητα ως προς την ακριβή απόδοση του μοντέλου, καθώς περιλαμβάνει τόσο τιμές κοντά στην τυχαία πρόβλεψη όσο και σημαντικά βελτιωμένες αποδόσεις. Παρόλο που η μέση απόδοση είναι ενθαρρυντική, ο στατιστικός έλεγχος υπόθεσης δεν απέδειξε στατιστικά σημαντική διαφορά σε σύγκριση με την τυχαία πρόβλεψη. Το p-value (0,133) είναι μεγαλύτερο από το καθιερωμένο όριο σημαντικότητας (0,05), πράγμα που σημαίνει ότι δεν μπορεί να απορριφτεί η υπόθεση ότι η μέση ακρίβεια του μοντέλου δεν διαφέρει ουσιαστικά από το 50%, όπως μπορεί να παρατηρηθεί και από τον πίνακα των αποτελεσμάτων παραπάνω. Αυτό προσθέτει επιπλέον

αβεβαιότητα στην αξιοπιστία των αποτελεσμάτων και υποδεικνύει ότι δεν υπερέχει σε σχέση με μια βασική γραμμή απόδοσης.

Συμπερασματικά, η ανάλυση αυτή υποδεικνύει ότι το μοντέλο παρουσιάζει μια τάση προς καλύτερες επιδόσεις από την τυχαία πρόβλεψη, αλλά η ύπαρξη σημαντικής διακύμανσης και η αβεβαιότητα των αποτελεσμάτων θέτουν ερωτήματα για την ανάγκη περαιτέρω βελτιώσεων. Απαιτούνται κατάλληλες προσαρμογές που θα ενισχύσουν την απόδοση του μοντέλου

5.2 Προβλήματα που παρουσιάστηκαν κατά την πειραματική διαδικασία

Κατά την πειραματική διαδικασία προκύπτουν αρκετές προκλήσεις, οι οποίες επηρεάζουν σημαντικά την ακρίβεια και την αξιοπιστία των αποτελεσμάτων. Καταρχάς, παρατηρούνται αντιφατικά αποτελέσματα ανεξαρτήτως του μοντέλου που χρησιμοποιείται, γεγονός που δυσχεραίνει την πρόβλεψη ως προς το ποιο μοντέλο αποδίδει καλύτερα. Αν και η πιο ορθή μέθοδος για την αντιμετώπιση αυτών των προβλημάτων θεωρείται η χρήση του αλγορίθμου Leave-One-Out Cross Validation (είτε η απλή είτε η παραλλαγμένη έκδοση), εντούτοις η μέθοδος αυτή δεν κατορθώνει να παράξει ικανοποιητικά αποτελέσματα όσον αφορά τη μετρική της ακρίβειας (Accuracy). Παρά την ισορροπία στα δεδομένα, η ακρίβεια κυμαίνεται σταθερά γύρω από το 0.5, υποδηλώνοντας αδυναμία του αλγορίθμου να προσφέρει αξιόπιστη εκτίμηση της ποιότητας των αποτελεσμάτων.

Για την αντιμετώπιση των μη ικανοποιητικών αποτελεσμάτων γίνεται η επιλογή χρήσης ατομικών μοντέλων για κάθε συμμετέχοντα, με στόχο την επίτευξη καλύτερων αποτελεσμάτων, καθώς το bias κάθε συμμετέχοντα μπορεί να διαδραματίζει καθοριστικό ρόλο στην ακρίβεια των προβλέψεων. Πράγματι, αυτή η μεθοδολογία αποδίδει καλύτερα αποτελέσματα, ωστόσο παρουσιάζει σημαντικά προβλήματα επεκτασιμότητας, καθώς απαιτεί δεδομένα από ένα μεμονωμένο άτομο κάθε φορά, γεγονός που περιορίζει τη γενίκευση της τεχνικής σε μεγαλύτερη κλίμακα. Τέλος, παρόλο που τα μοντέλα εκπαιδεύονται με κατάλληλο αριθμό εποχών, ανάλογα με την πολυπλοκότητα του εκάστοτε μοντέλου και τον αριθμό παραμέτρων, η σταθερότητα της ακρίβειας στα δεδομένα εκπαίδευσης δεν συνοδεύεται πάντα από βελτίωση στα σύνολα ελέγχου (test) και επικύρωσης (validation), ενισχύοντας την αβεβαιότητα σχετικά με την αποτελεσματικότητα της διαδικασίας εκπαίδευσης.

5.3 Τρόποι επίλυσης των προβλημάτων

Για την επίλυση των προαναφερθέντων προβλημάτων, μπορούν να υιοθετηθούν διάφορες στρατηγικές τόσο στον τομέα της βελτιστοποίησης των μοντέλων (model exploration) όσο και στη διαχείριση και επεξεργασία των δεδομένων (data enhancement). Αρχικά, η εφαρμογή τεχνικών Grid Search στο υπάρχον μοντέλο θα μπορούσε να επιτρέψει τη διερεύνηση ενός ευρέως φάσματος υπερπαραμέτρων, όπως είναι το learning rate, ο αριθμός των επιπέδων LSTM ή το μέγεθος των νευρώνων, με σκοπό τον προσδιορισμό του βέλτιστου συνδυασμού παραμέτρων που αποδίδει καλύτερα. Παράλληλα, η αντικατάσταση του TimeDistributed 2D-CNN με ένα 3D-CNN μοντέλο θα μπορούσε να ενισχύσει την ικανότητα

του μοντέλου να αναγνωρίζει χωροχρονικές σχέσεις στα δεδομένα βίντεο, βελτιώνοντας έτσι την απόδοση στην κατηγοριοποίηση των συμπεριφορικών προτύπων.

Όσον αφορά τη βελτίωση των δεδομένων (data enhancement), προτείνεται η εφαρμογή data augmentation τεχνικών, όπως περιστροφή, αλλαγή φωτεινότητας, και διαφοροποίηση του μεγέθους των καρτέ, με στόχο τη δημιουργία πιο ανθεκτικών μοντέλων. Επιπλέον, η μεταβολή του μεγέθους των time windows των καρτέ που χρησιμοποιούνται για κάθε πακέτο, όπως η αύξηση ή μείωση του αριθμού των καρτέ (π.χ. από 20 σε 40, 60 ή ακόμα και 10), θα μπορούσε να δώσει σημαντικές πληροφορίες για το βέλτιστο μέγεθος δεδομένων που οδηγεί σε καλύτερη ακρίβεια. Ακόμη, πρόσθετες τεχνικές επεξεργασίας των καρτέ, όπως η χρήση διαφορετικών φίλτρων επεξεργασίας εικόνας (π.χ. η μετατροπή των καρτέ σε μονόχρωμη μορφή), θα μπορούσαν να βελτιώσουν περαιτέρω την ποιότητα των δεδομένων εισόδου.

Επιπροσθέτως, η χρήση εναλλακτικών μεθόδων αξιολόγησης της προσήλωσης, πέρα από τον μέσο όρο, θα μπορούσε να βελτιώσει τη συνολική ποιότητα της ανάλυσης. Αυτές οι μέθοδοι περιλαμβάνουν την περιοχή του ίχνους των δεδομένων, το πλάτος και την μέση κλίση του ίχνους, ή ακόμη και ο συνδυασμός κάποιων από τους παραπάνω, θα μπορούσαν να προσφέρουν μία πιο ολοκληρωμένη και ακριβή απεικόνιση των μοτίβων προσήλωσης των συμμετεχόντων. Μία πιο ενδελεχής διερεύνηση στα γραφήματα των απεικονίσεων της προσήλωσης και η ανεύρεση μίας κοινής αρχής που να διέπει τα δεδομένα προσήλωσης θα ενισχύσει τη διαδικασία επιλογής των συμμετεχόντων που επιλέγονται για εκπαίδευση, σε σύγκριση με αυτά των συμμετεχόντων που δεν επιλέγονται.

Κεφάλαιο 6: Επίλογος

6.1 Σύνοψη

Συνοψίζοντας, η παρούσα Διπλωματική Εργασία επικεντρώνεται σε δύο κύριους άξονες. Ο πρώτος άξονας αφορά τη θεωρητική πλαισίωση της εργασίας, η οποία περιλαμβάνει την ανάλυση και παρουσίαση των ΠΣΣΥ, τη ΣΥΝ και τις σύγχρονες μεθόδους καταγραφής δεδομένων μέσω βίντεο. Δόθηκε έμφαση στη βιβλιογραφική ανασκόπηση με σκοπό τη διερεύνηση των τεχνικών και αλγορίθμων βαθιάς μηχανικής μάθησης, όπως τα CNN και LSTM, που χρησιμοποιούνται για την επεξεργασία χωροχρονικών δεδομένων. Ο δεύτερος άξονας αφορά το πρακτικό μέρος της έρευνας, όπου πραγματοποιήθηκε η συλλογή και επεξεργασία δεδομένων βίντεο από την αλληλεπίδραση συμμετεχόντων με ένα συγκεκριμένο ΠΣΣΥ, καθώς και των δεδομένων προσήλωσης τους στη διάρκεια της αλληλεπίδρασης τους με το ΠΣΣΥ. Ακολούθησε η κατασκευή και εκπαίδευση σύνθετων μοντέλων βαθιάς μάθησης, τα οποία στόχευαν στην πρόβλεψη των επιπέδων προσήλωσης με βάση τα καταγεγραμμένα δεδομένα, όπου για την αξιολόγηση τους χρησιμοποιήθηκαν τεχνικές όπως το Leave-One-Out Cross-Validation και η εκπαίδευση εξατομικευμένων μοντέλων για κάθε συμμετέχοντα, με στόχο την ελαχιστοποίηση του bias και την παραγωγή πιο αξιόπιστων αποτελεσμάτων. Τέλος, διερευνήθηκαν τα προβλήματα που ανέκυψαν κατά τη διαδικασία των πειραμάτων, όπως τα αντιφατικά αποτελέσματα μεταξύ των μοντέλων και η δυσκολία στην εξαγωγή ακριβών προβλέψεων. Προτάθηκαν λύσεις για τη βελτίωση της απόδοσης των μοντέλων μέσω τεχνικών model exploration και data augmentation, καθώς και εναλλακτικών τρόπων αξιολόγησης της προσήλωσης. Με αυτόν τον τρόπο, η παρούσα εργασία συνέβαλε στην κατανόηση της σχέσης μεταξύ δεδομένων καταγραφής βίντεο που έχουν ληφθεί από ένα ΠΣΣΥ και των μετρικών της προσήλωσης των συμμετεχόντων, εξερευνώντας τη δυνατότητα πρόβλεψης της προσήλωσης, παρέχοντας παράλληλα προτάσεις και για πιθανές μελλοντικές βελτιώσεις.

6.2 Όρια και Περιορισμοί

Στην παρούσα διπλωματική εργασία παρουσιάζονται κάποιοι σημαντικοί περιορισμοί, οι οποίοι επηρεάζουν τη δυνατότητα διεξαγωγής των απαιτούμενων πειραμάτων. Αρχικά, για την πραγματοποίηση των πειραμάτων χρησιμοποιείται μια κάρτα γραφικών NVIDIA GeForce RTX 3060, με διαθέσιμη μνήμη 10340 MB, η οποία, αν και ικανοποιητική, παρουσιάζει περιορισμούς όσον αφορά την επεξεργαστική της ισχύ. Αυτός ο περιορισμός αυξάνει την ανάγκη να μειωθεί ο αριθμός των συμμετεχόντων που χρησιμοποιούνται για την εκπαίδευση κάθε μοντέλου που συνδυάζει τα δεδομένα πολλών συμμετεχόντων ταυτόχρονα, καθώς ο χρόνος που απαιτείται για την εκπαίδευση συχνά αγγίζει τις δεκάδες ώρες αλλά και λόγω της ανεπάρκειας της μνήμης της κάρτας. Επιπλέον ένας περιορισμός είναι το μικρό πλήθος των συμμετεχόντων που λαμβάνουν μέρος στην έρευνα, γεγονός που περιορίζει την ικανότητα γενίκευσης των αποτελεσμάτων. Η περιορισμένη ποσότητα δεδομένων δεν επιτρέπει την εξαγωγή συμπερασμάτων που μπορούν να γενικευτούν με ασφαλή τρόπο σε μεγαλύτερο πληθυσμό.

Ένας ακόμη σημαντικός περιορισμός αφορά τον τρόπο υπολογισμού της προσήλωσης, ο οποίος βασίζεται στην υποκειμενική εκτίμηση του κάθε συμμετέχοντα. Δεν υπάρχει ομοιομορφία στον τρόπο που οι συμμετέχοντες ορίζουν την αύξηση ή τη μείωση της προσήλωσης τους, γεγονός που εισάγει ασάφεια και πιθανές αποκλίσεις στα αποτελέσματα. Πιο συγκεκριμένα, η εφαρμογή κάποιων μεθόδων αναγνώρισης συναισθημάτων θα μπορούσε να βελτιώσει σημαντικά τη συλλογή των δεδομένων προσήλωσης των συμμετεχόντων χωρίς να εγείρει προβληματισμούς σχετικά με την ακρίβεια της καταγραφής της, η οποία τώρα περιορίζεται στη μνήμη του κάθε συμμετέχοντα. Τέλος, η ανάλυση που έχει διεξαχθεί αποτελεί μία πλήρως καινοτόμα προσέγγιση, καθώς δεν έχει εντοπιστεί αντίστοιχη μελέτη στη βιβλιογραφία. Αυτό σημαίνει ότι τα αποτελέσματα ενδέχεται να είναι ασαφή και επιπλέον να παρουσιάζουν δυσκολία στην ερμηνεία τους, κάτι που παρατηρείται στα πειράματα της παρούσας έρευνας. Η καινοτομία της συγκεκριμένης μελέτης καθιστά απαραίτητη την περαιτέρω διερεύνηση για να εξαχθούν πιο σαφή και αξιόπιστα συμπεράσματα.

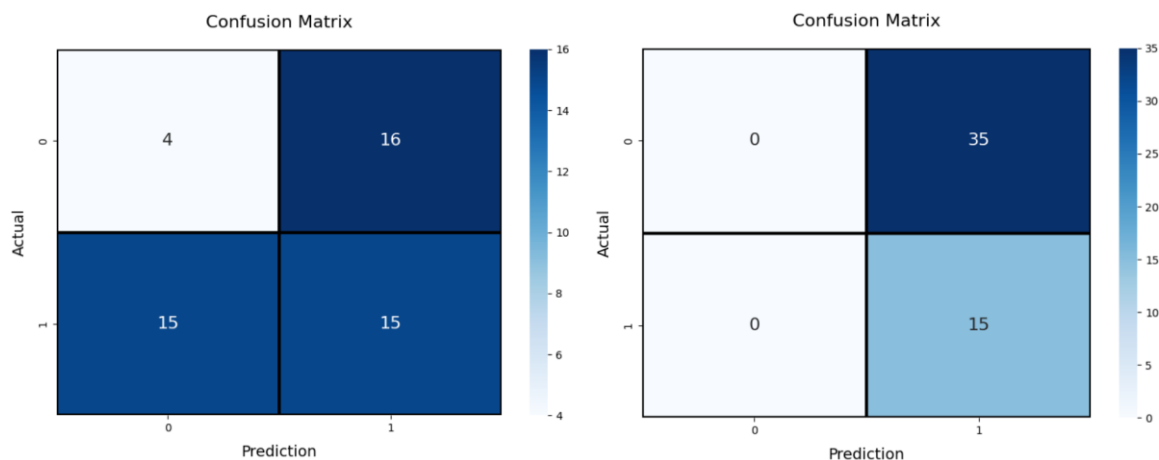
Κεφάλαιο 7: Βιβλιογραφία

- [1] K. B. N. D. P. Mauricio Flores Gustavo Glusman and L. Hood, "P4 Medicine: How Systems Medicine Will Transform the Healthcare Sector and Society," *Per Med*, vol. 10, no. 6, pp. 565–576, 2013, doi: 10.2217/pme.13.57.
- [2] Ventola CL., "Mobile devices and apps for health care professionals: uses and benefits," *PMC*, 2014.
- [3] T. Tao Jianhua and Tan, "Affective Computing: A Review," in *Affective Computing and Intelligent Interaction*, T. and P. R. W. Tao Jianhua and Tan, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 981–995.
- [4] F. Laamarti, M. Eid, and A. El Saddik, "An Overview of Serious Games," *International Journal of Computer Games Technology*, vol. 2014, no. 1, p. 358152, 2014, doi: <https://doi.org/10.1155/2014/358152>.
- [5] V. Wattanasoontorn, I. Boada, R. García, and M. Sbert, "Serious games for health," *Entertain Comput*, vol. 4, no. 4, pp. 231–247, Dec. 2013, doi: 10.1016/J.ENTCOM.2013.09.002.
- [6] K. Mitsis *et al.*, "A Multimodal Approach for Real Time Recognition of Engagement towards Adaptive Serious Games for Health," *Sensors*, vol. 22, no. 7, Apr. 2022, doi: 10.3390/s22072472.
- [7] Wikipedia contributors, "Serious game — Wikipedia, The Free Encyclopedia," Wikipedia. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Serious_game&oldid=1242838389
- [8] K. Mitsis, K. Zarkogianni, N. Bountouni, M. Athanasiou, and K. S. Nikita, "An Ontology-Based Serious Game Design for the Development of Nutrition and Food Literacy Skills," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 1405–1408. doi: 10.1109/EMBC.2019.8856604.
- [9] K. Mitsis, E. Kalafatis, K. Zarkogianni, G. Mourkousis, and K. S. Nikita, "Procedural content generation based on a genetic algorithm in a serious game for obstructive sleep apnea," in *2020 IEEE Conference on Games (CoG)*, 2020, pp. 694–697. doi: 10.1109/CoG47356.2020.9231785.
- [10] M. and M. A. and L. J. and D. W.-Z. L. and S. S. and S. A. and S. M. Graafland Maurits and Dankbaar, "How to Systematically Assess Serious Games Applied to Health Care," *JMIR Serious Games*, vol. 2, no. 2, p. e11, Nov. 2014, doi: 10.2196/games.3825.
- [11] V. Monteiro, L. Mata, and F. Peixoto, "Intrinsic Motivation Inventory: Psychometric Properties in the Context of First Language and Mathematics Learning," *Psicologia: Reflexão e Crítica*, vol. 28, no. 3, pp. 434–443, Jul. 2015, doi: 10.1590/1678-7153.201528302.
- [12] S. G. Hart, "Nasa-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, 2006, doi: 10.1177/154193120605000909.

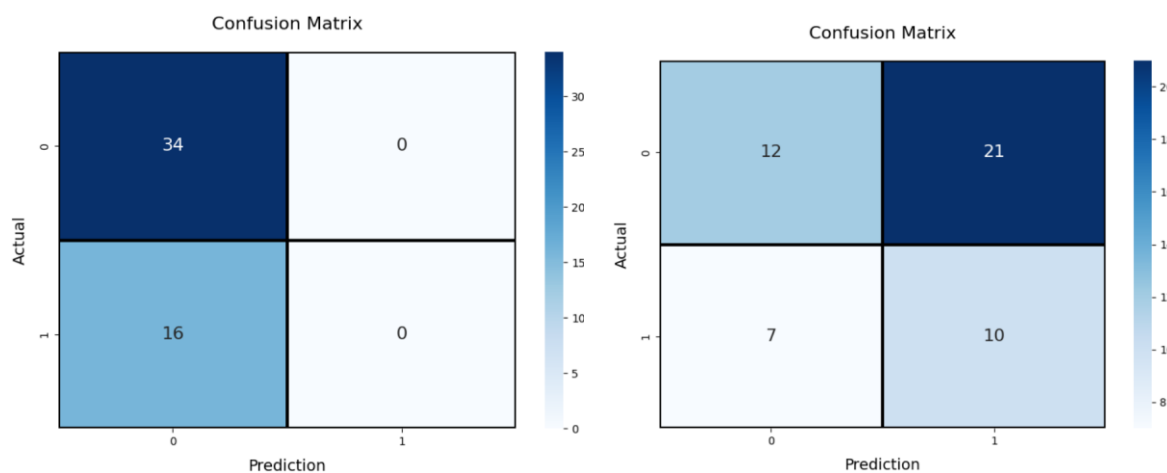
- [13] P. T. K. Aaron Bangor and J. T. Miller, "An Empirical Evaluation of the System Usability Scale," *Int J Hum Comput Interact*, vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.
- [14] R. Jääskeläinen, "Think-aloud protocol," *Handbook of translation studies*, vol. 1, pp. 371–374, 2010.
- [15] Kim Tae Kyun, "T test as a parametric statistic," *kja*, vol. 68, no. 6, pp. 540–546, 2015, doi: 10.4097/kjae.2015.68.6.540.
- [16] L. Strohle and S. Wold, "Analysis of variance (ANOVA)," *Chemometrics and Intelligent Laboratory Systems*, vol. 6, no. 4, pp. 259–272, 1989, doi: [https://doi.org/10.1016/0169-7439\(89\)80095-4](https://doi.org/10.1016/0169-7439(89)80095-4).
- [17] A. A. Tori, R. Tori, and F. de L. dos S. Nunes, "Serious Game Design in Health Education: A Systematic Review," *IEEE Transactions on Learning Technologies*, vol. 15, no. 6, pp. 827–846, 2022, doi: 10.1109/TLT.2022.3200583.
- [18] M. Graafland Maurits and Schijven, "How Serious Games Will Improve Healthcare," in *Digital Health: Scaling Healthcare to the World*, K. Rivas Homero and Wac, Ed., Cham: Springer International Publishing, 2018, pp. 139–157. doi: 10.1007/978-3-319-61446-5_10.
- [19] Anvita Saxena, Ashish Khanna, and Deepak Gupta, "Emotion Recognition and Detection Methods: A Comprehensive Survey," *Journal of Artificial Intelligence and Systems*, vol. 2, pp. 53–79, 2020.
- [20] A. Ezquerro, F. Agen, I. Rodríguez-Arteche, and I. Ezquerro-Romano, "Integrating Artificial Intelligence into Research on Emotions and Behaviors in Science Education," *EURASIA Journal of Mathematics, Science and Technology Education*, vol. 18, p. 4, 2022.
- [21] J. Zhang, Z. Yin, P. Chen, and S. Nichele, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," *Information Fusion*, vol. 59, pp. 103–126, Jul. 2020, doi: 10.1016/J.INFFUS.2020.01.011.
- [22] G. Assunção, B. Patrão, M. Castelo-Branco, and P. Menezes, "An Overview of Emotion in Artificial Intelligence," *IEEE Transactions on Artificial Intelligence*, vol. 3, no. 6, pp. 867–886, 2022, doi: 10.1109/TAI.2022.3159614.
- [23] F. Z. Canal *et al.*, "A survey on facial emotion recognition techniques: A state-of-the-art literature review," *Inf Sci (N Y)*, vol. 582, pp. 593–617, Jan. 2022, doi: 10.1016/J.INS.2021.10.005.
- [24] N. Alswaidan and M. E. B. Menai, "A survey of state-of-the-art approaches for emotion recognition in text," *Knowl Inf Syst*, vol. 62, no. 8, pp. 2937–2987, 2020, doi: 10.1007/s10115-020-01449-0.
- [25] N. S. Suhaimi, J. Mountstephens, and J. Teo, "EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities," *Comput Intell Neurosci*, vol. 2020, no. 1, p. 8875426, 2020, doi: <https://doi.org/10.1155/2020/8875426>.
- [26] M. Frutos-Pascual and B. G. Zapirain, "Review of the Use of AI Techniques in Serious Games: Decision Making and Machine Learning," *IEEE Trans Comput Intell AI Games*, vol. 9, no. 2, pp. 133–152, 2017, doi: 10.1109/TCIAIG.2015.2512592.

- [27] T.-L. Wang and Y.-H. B. Lien, "The power of using video data," *Qual Quant*, vol. 47, no. 5, pp. 2933–2941, 2013, doi: 10.1007/s11135-012-9717-0.
- [28] V. Sharma, M. Gupta, A. Kumar, and D. Mishra, "Video Processing Using Deep Learning Techniques: A Systematic Literature Review," *IEEE Access*, vol. 9, pp. 139489–139507, 2021, doi: 10.1109/ACCESS.2021.3118541.
- [29] M. Lauronen, "Ethical issues in topical computer vision applications," University of Jyväskylä, 2017. [Online]. Available: <https://jyx.jyu.fi/handle/123456789/55806>
- [30] P. Lopes, G. N. Yannakakis, and A. Liapis, "RankTrace: Relative and unbounded affect annotation," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 158–163. doi: 10.1109/ACII.2017.8273594.
- [31] T. Beyssolow II, "Convolutional Neural Networks (CNNs)," in *Introduction to Deep Learning Using R: A Step-by-Step Guide to Learning and Implementing Deep Learning Models Using R*, Berkeley, CA: Apress, 2017, pp. 101–112. doi: 10.1007/978-1-4842-2734-3_5.
- [32] E. Bisong, "Recurrent Neural Networks (RNNs)," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, Berkeley, CA: Apress, 2019, pp. 443–473. doi: 10.1007/978-1-4842-4470-8_36.

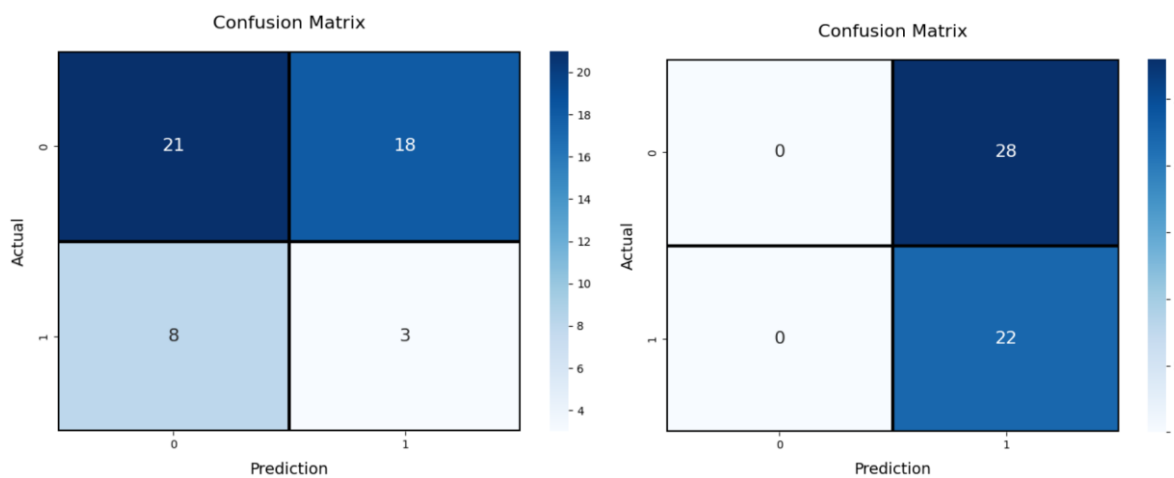
Παράρτημα Εικόνων



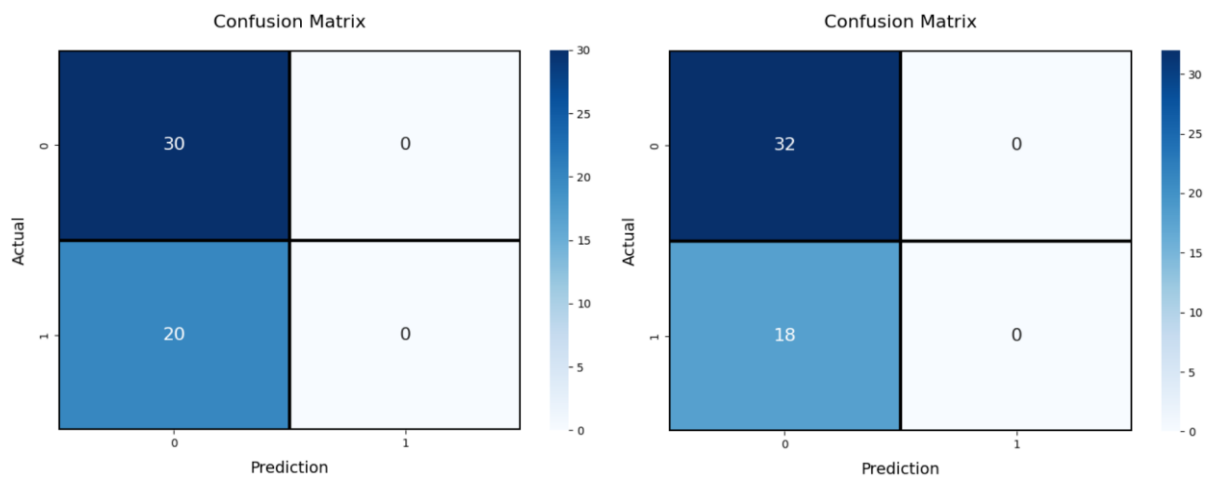
Εικόνα 24: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 1, 2



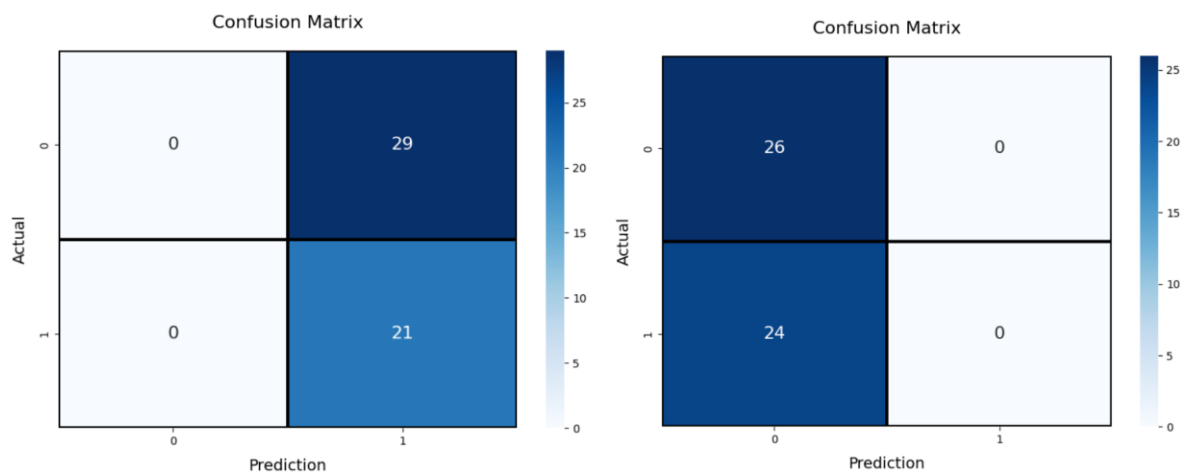
Εικόνα 25: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 7, 9



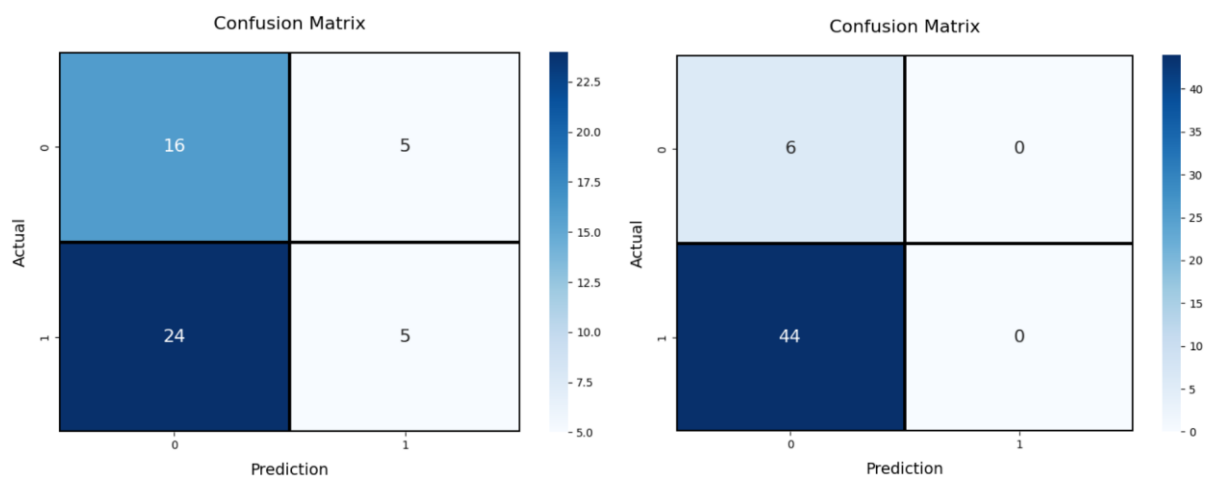
Εικόνα 26: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 11, 12



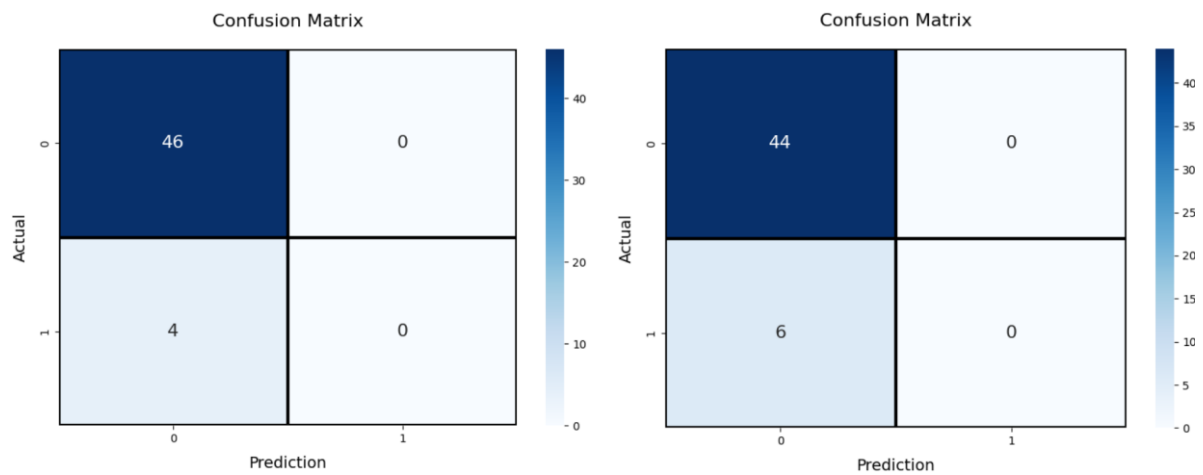
Εικόνα 27: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 14, 17



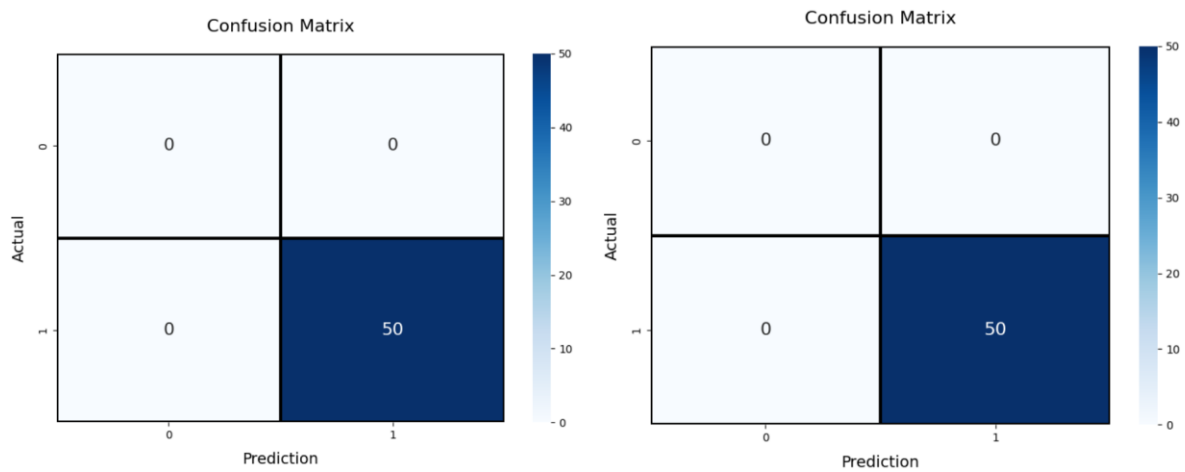
Εικόνα 28: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 21, 5



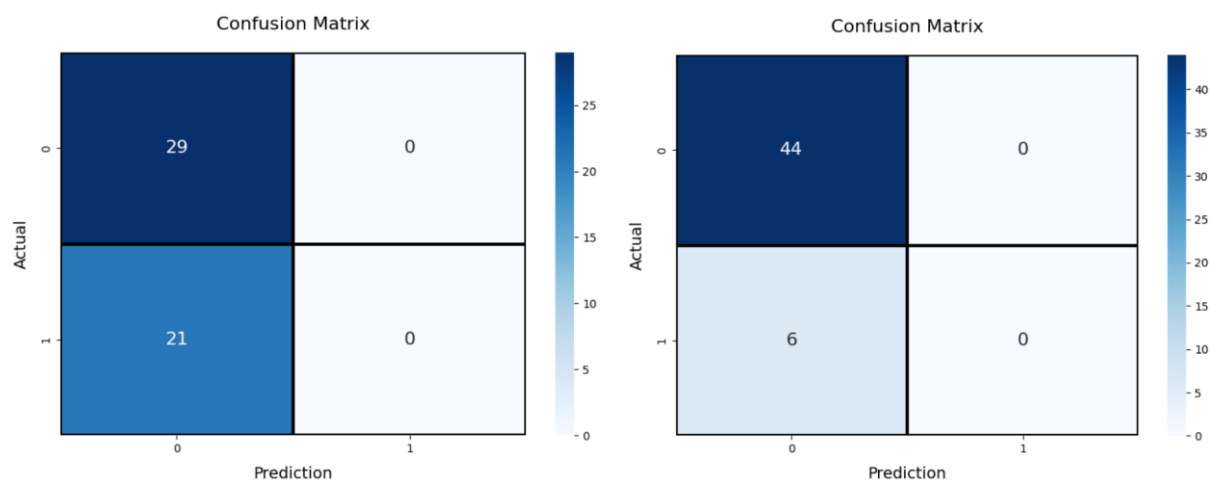
Εικόνα 29: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 22, 25



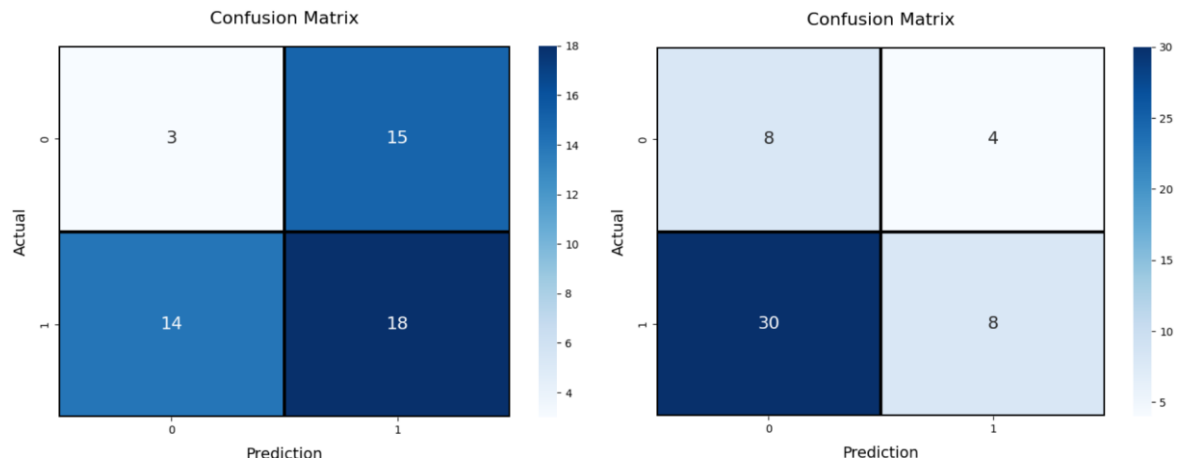
Εικόνα 30: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 3, 4



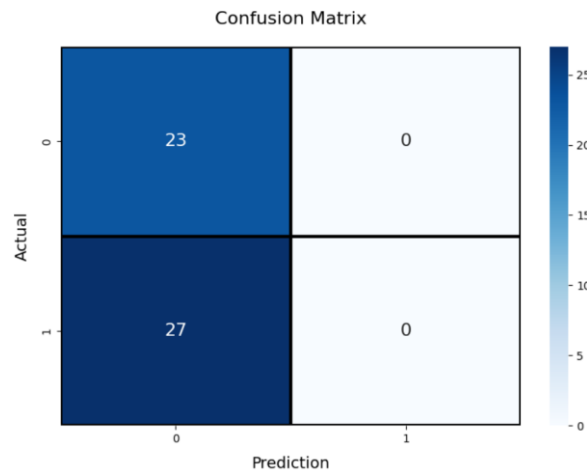
Εικόνα 31: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 8, 13



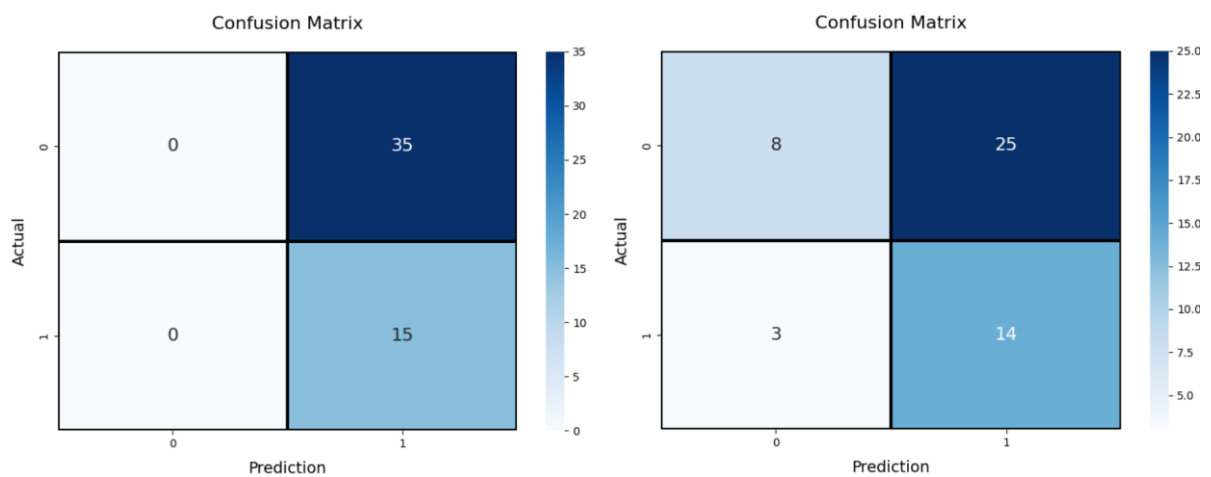
Εικόνα 32: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 16, 19



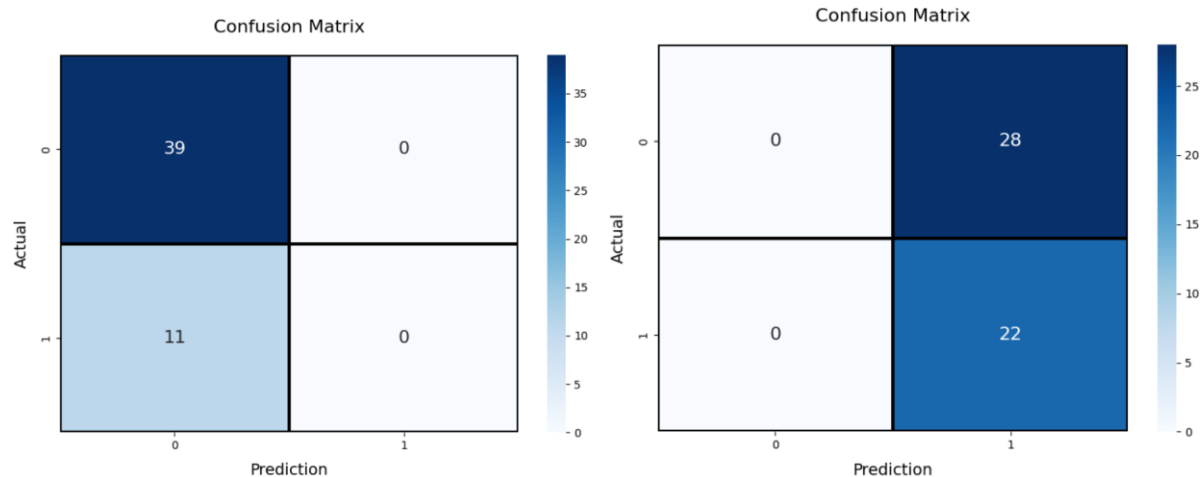
Εικόνα 33: Confusion Matrix για τα ατομικά μοντέλα (9) των συμμετεχόντων: 20, 23



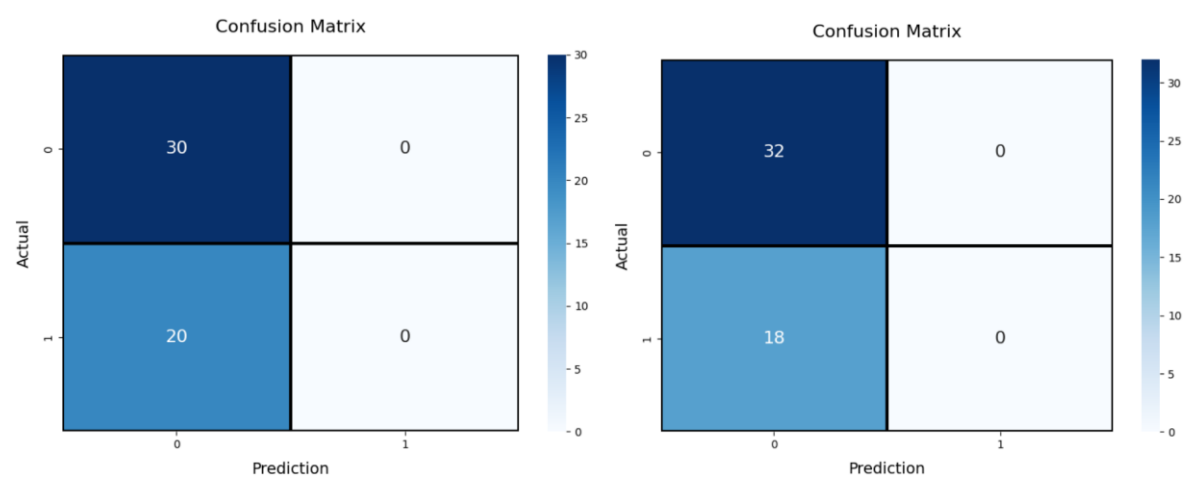
Εικόνα 34: Confusion Matrix για το ατομικό μοντέλο (9) του συμμετέχοντα: 24



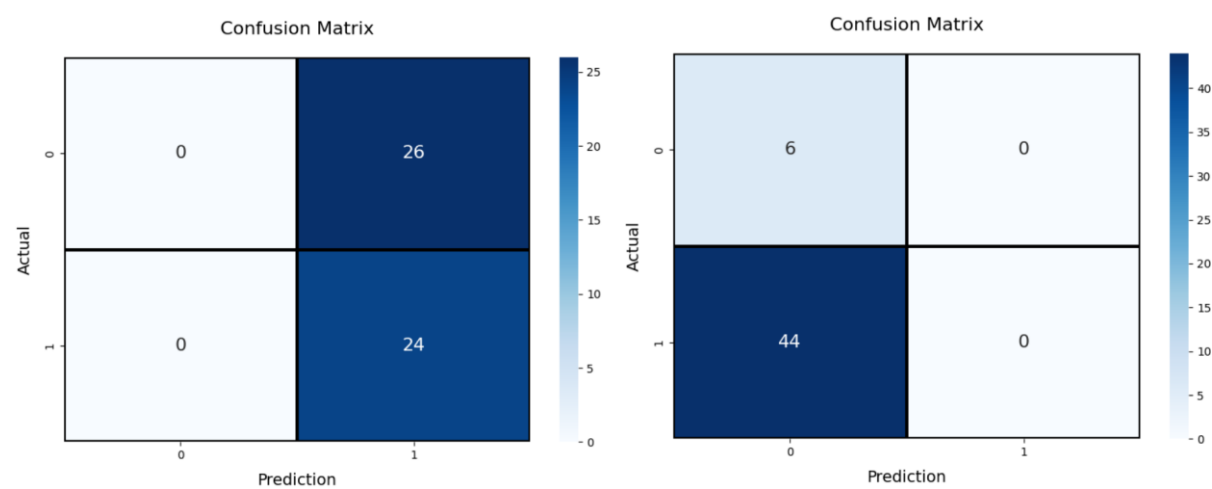
Εικόνα 35: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 2, 9



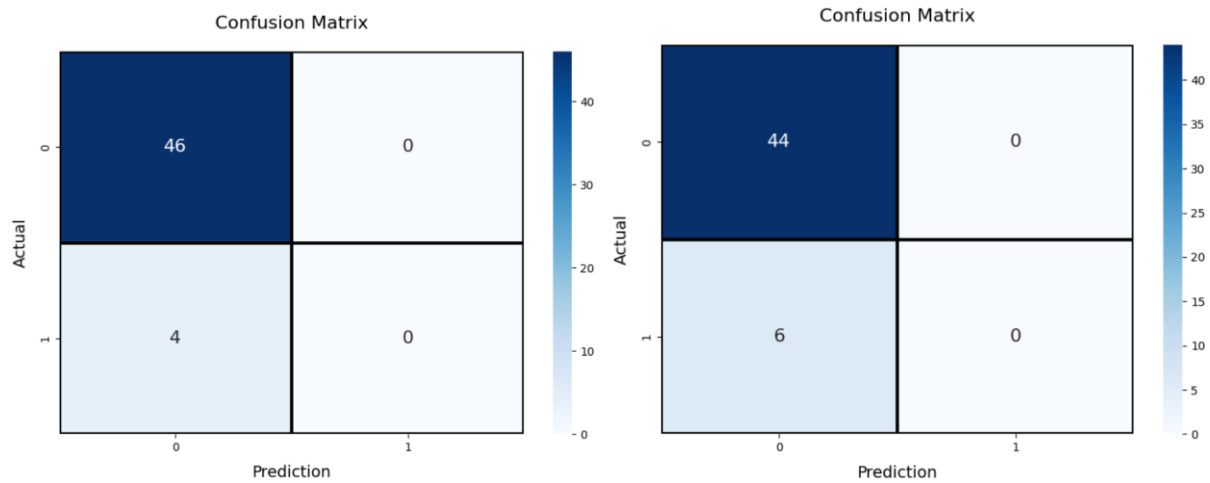
Εικόνα 36: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 11, 12



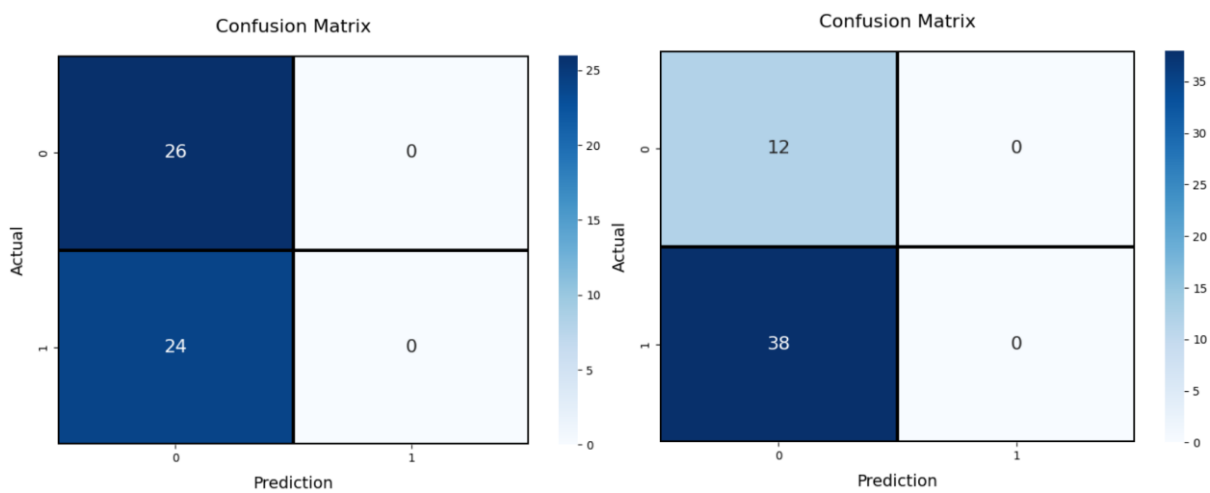
Εικόνα 37: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 14, 17



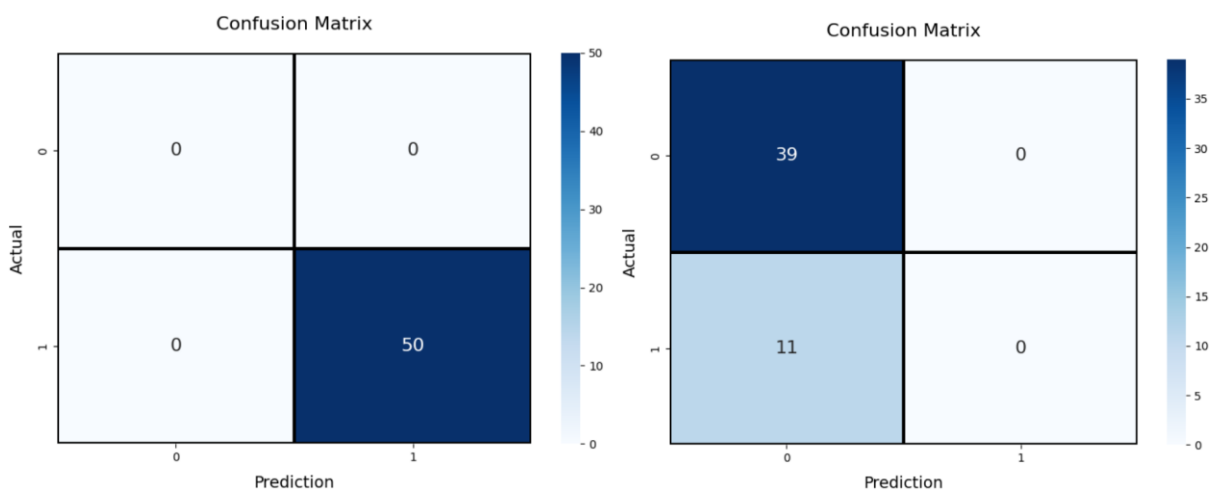
Εικόνα 38: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 18, 25



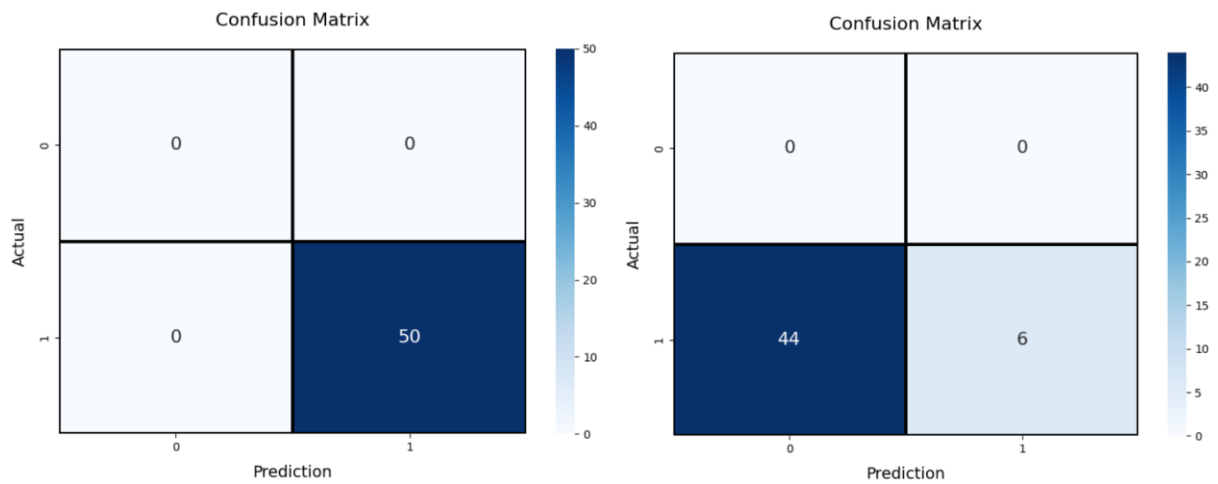
Εικόνα 36: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 3, 4



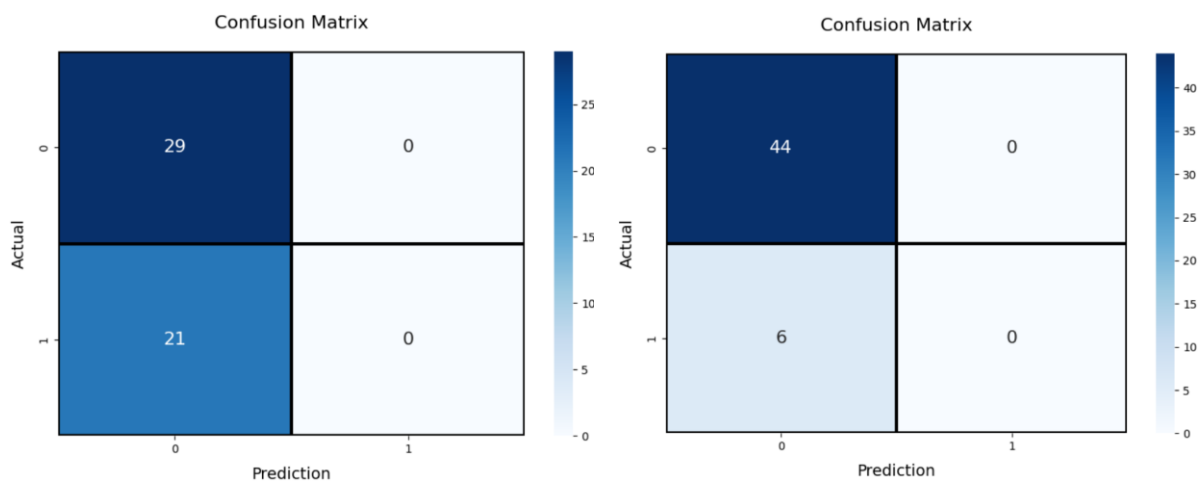
Εικόνα 37: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 5, 6



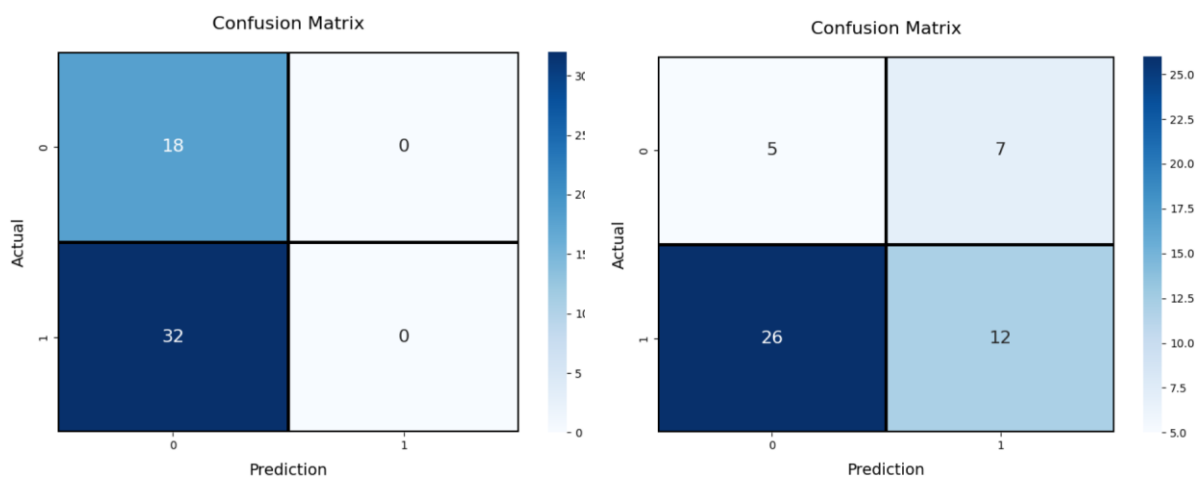
Εικόνα 38: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 8, 10



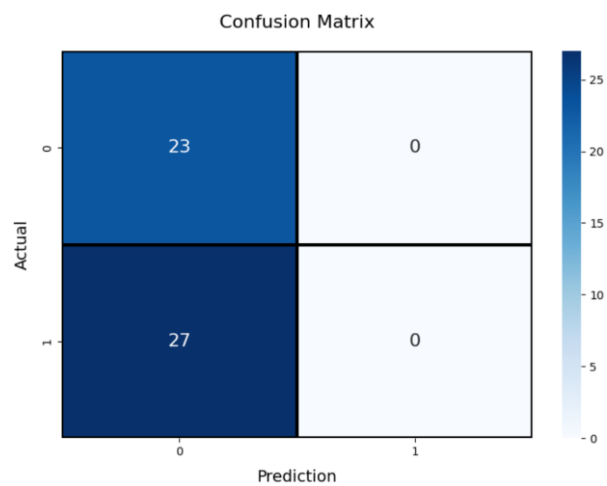
Εικόνα 39: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 13, 15



Εικόνα 40: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 16, 19



Εικόνα 41: Confusion Matrix για τα ατομικά μοντέλα (10) των συμμετεχόντων: 20, 23



Εικόνα 42: Confusion Matrix για το ατομικό μοντέλο (10) του συμμετέχοντα: 24