Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Εργαστήριο Επεξεργασίας Πληροφορίας και Υπολογισμών

Συνεχής έμμεση αυθεντικοποίηση χρηστών κινητού τηλεφώνου με συνδυασμό των δεδομένων πλοήγησης και συμπεριφοράς

Χρήστος Εμμανουήλ

Επιβλέπων Καθηγητής: **Ανδρέας Συμεωνίδης**

Επιβλέπων Υποψήφιος Διδάκτωρ: **Θωμάς Καρανικιώτης**

Περιεχόμενα

Κίνητρο

- Συνεχώς αυξανόμενος αριθμός χρηστών smartphones
- Παραγωγή και αποθήκευση προσωπικών και επαγγελματικών πληροφοριών
- Ανάγκη για την ασφάλεια των δεδομένων που αποθηκεύονται στην συσκευή.
- Προβληματισμοί για την επάρκεια των υφιστάμενων τρόπων αυθεντικοποίησης.
- Ανάγκη για υλοποίηση νέων μεθοδολογιών αυθεντικοποίησης.

Συνεχής – Έμμεση Αυθεντικοποίηση

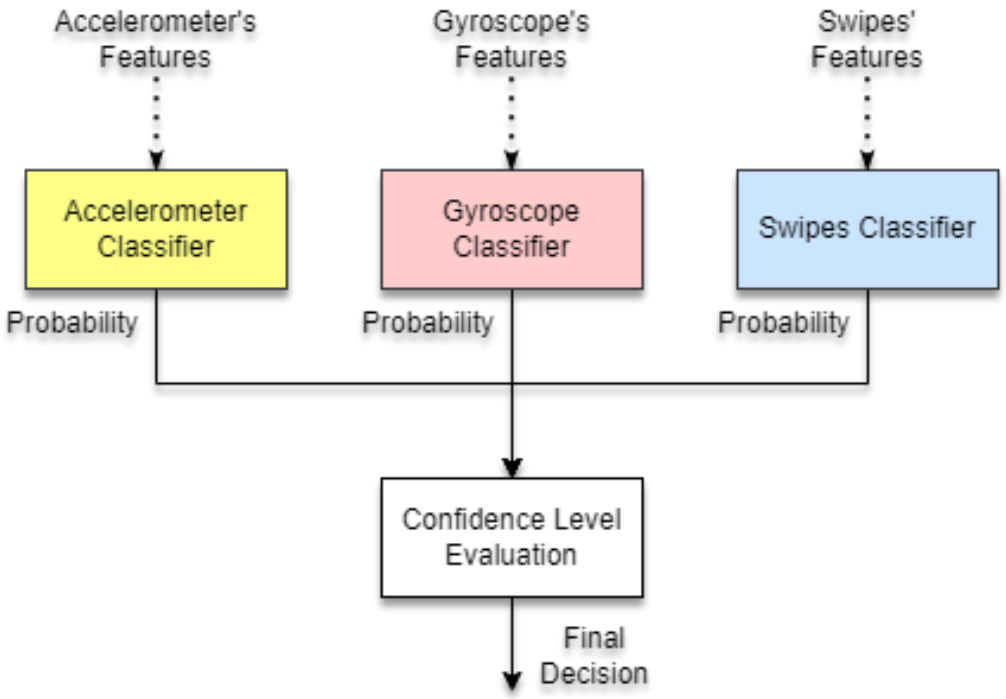
Πλεονεκτήματα:

- Ενισχυμένο σύστημα ασφάλειας
- Καλύτερη εμπειρία χρήστη
- Δυνατότητα εκμετάλλευσης συμπεριφορικών χαρακτηριστικών
  - Εύκολη προσαρμογή
  - Χαμηλό κόστος υλοποίησης
  - Προοπτικές εξέλιξης

Προβληματισμοί:

- Υψηλή ρυθμοί δειγματοληψίας
- Πόροι υψηλής κατανάλωσης ισχύος
- Δευτερεύων συσκευές (wearables)
- Ανεπαρκή αξιολόγηση
  - Μικρό πλήθος δεδομένων
  - Δεδομένα ‘εργαστηρίου’
  - ‘Λανθασμένες’ μετρικές
- Ανεπάρκεια δεδομένων κατά την εκτέλεση

Κεντρική Ιδέα



Στόχος:

- Ικανοποιητικά επίπεδα ασφάλειας και διαφάνειας
- Χρήση δεδομένων που παράγονται από το smartphone
- Ανθεκτικό σε σφάλματα ή/και ελλείψεις δεδομένων

Ερωτήματα:

- Σύνολο δεδομένων
- Εξαγωγή χαρακτηριστικών και προεπεξεργασία
- Δομή ταξινομητών
- Δομή υποσυστήματος εμπιστοσύνης
- Αντικειμενική αξιολόγηση

Μεθοδολογία

Σύνολο Δεδομένων

BrainRun:

- Σύνολο συμπεριφορικών δεδομένων
- Δεδομένα αισθητήρων κίνησης και χειρονομιών
- Εφαρμογή συλλογής δεδομένων (android & iOS)
- 5 διαφορετικά παιχνίδια, με διαφορετικά επίπεδα δυσκολίας

Χαρακτηριστικά:

- 2218 χρήστες
- 60% άντρες, 26% γυναίκες, 14% άγνωστα
- 90% android, 10% iOS

Παιχνίδια & Τελικά Σύνολα
(μετά την εφαρμογή κριτηρίων επιλογής):

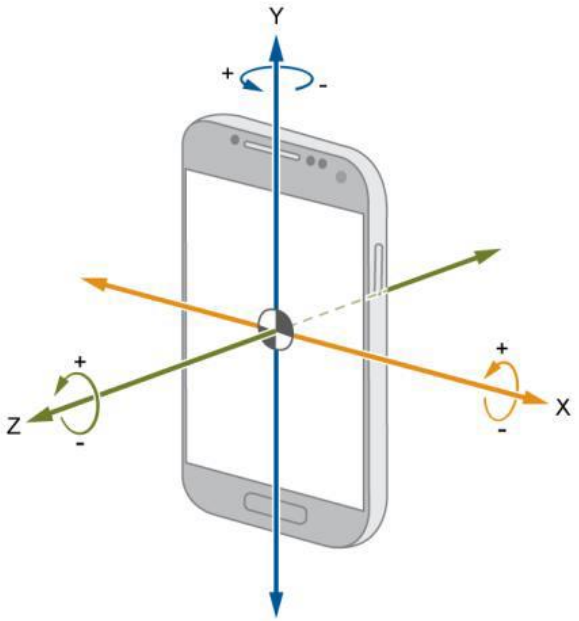
Παιχνίδια	Τύπος Δεδομένων	Αριθμός Χρηστών Εκπαίδευσης	Αριθμός Χρηστών Αξιολόγησης
Mathisis	Acc, Gyr, Swp	15	24
Focus	Acc, Gyr, Swp	15	30
Reacton	Acc, Gyr, Swp, Tap	15	45
Memoria	Acc, Gyr, Tap	15	44
Speedy	Acc, Gyr, Tap	15	45

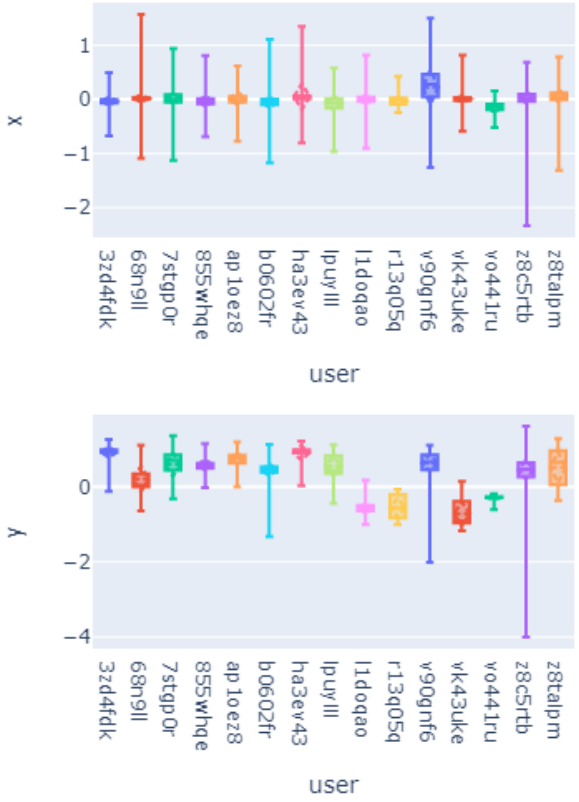
Acc: Επιταχυνσιόμετρο, Gyr: Γυροσκόπιο, Swp: Swipe

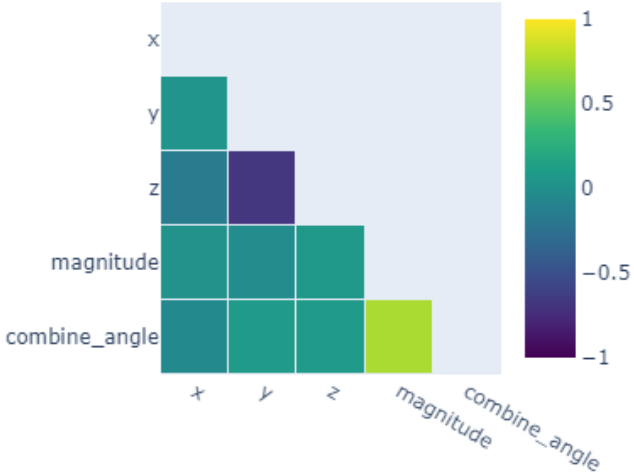
Μεθοδολογία

Εξαγωγή Χαρακτηριστικών

Επιταχυνσιόμετρο, Γυροσκόπιο (1)





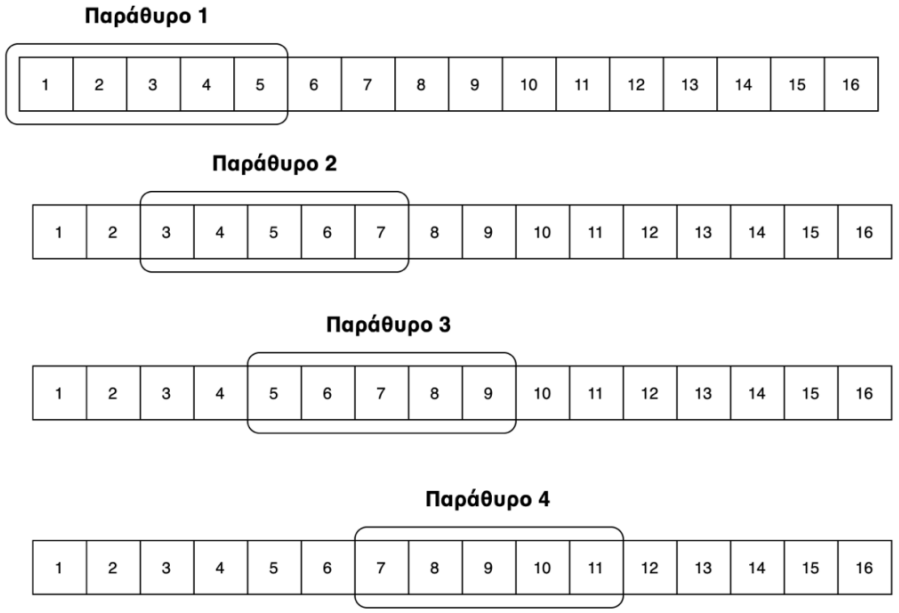


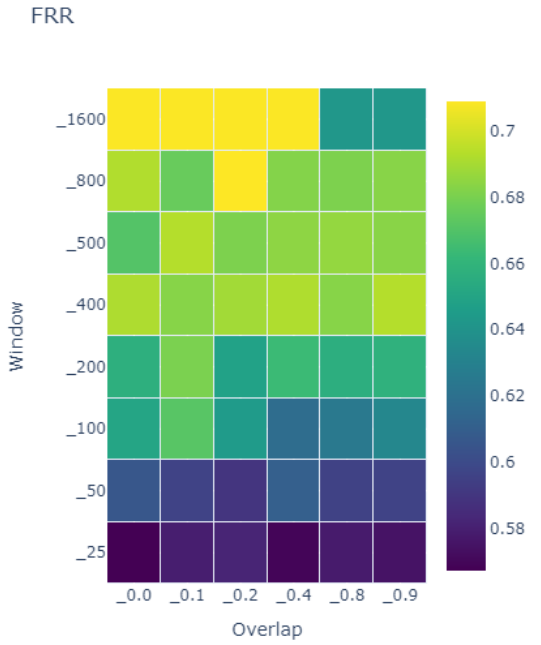
Επιλέχθηκαν:
x, y και magnitude

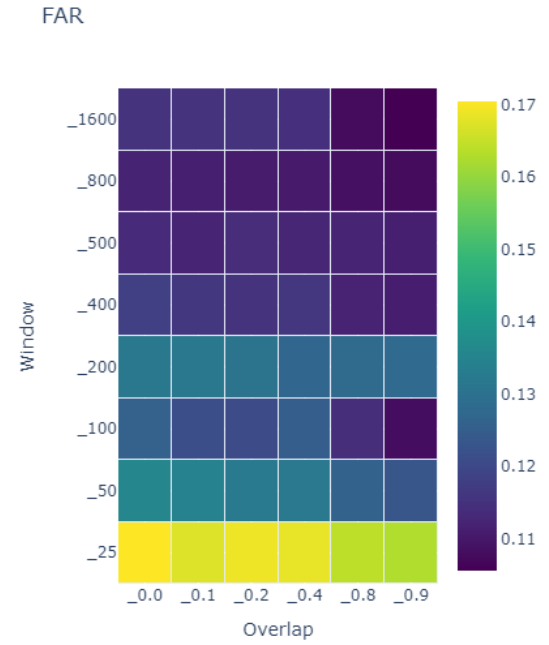
Μεθοδολογία

Εξαγωγή Χαρακτηριστικών

Επιταχυνσιόμετρο, Γυροσκόπιο (2)





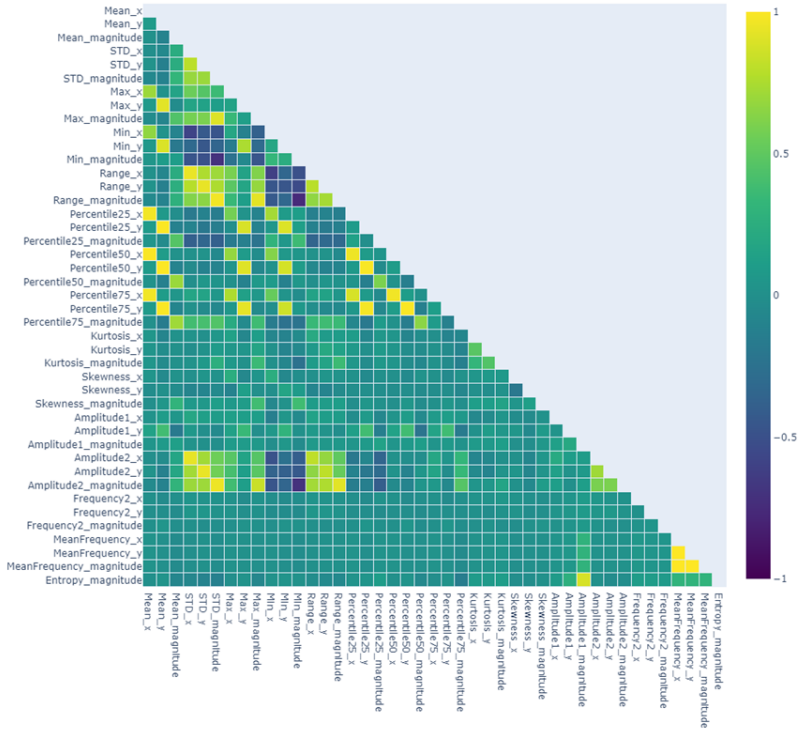


Επιλέχθηκαν:
Μέγεθος Παραθύρου: 50 δείγματα
Ποσοστό Επικάλυψης: 60%

Μεθοδολογία

Εξαγωγή Χαρακτηριστικών

Επιταχυνσιόμετρο, Γυροσκόπιο (3)

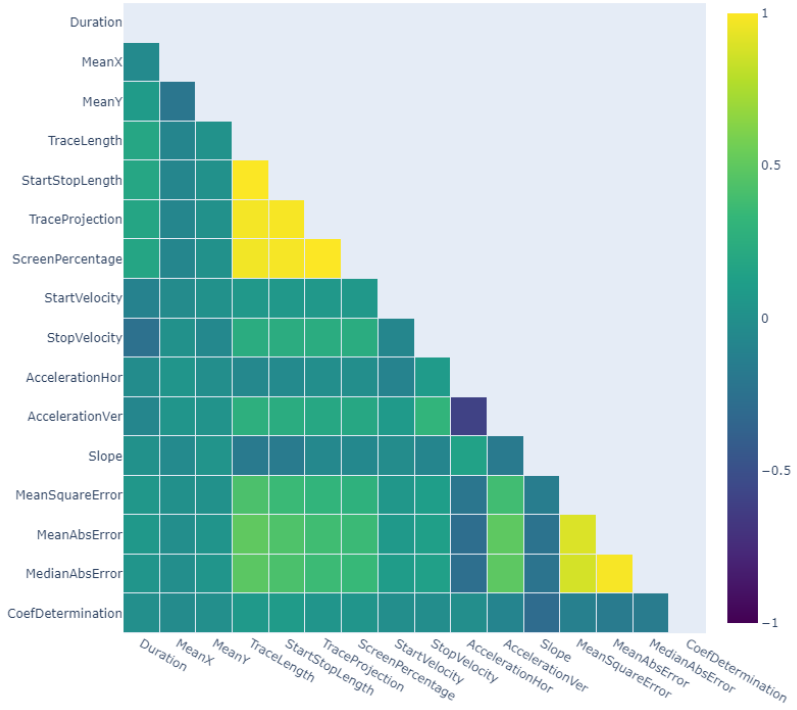


Αισθητήρας	Γνωρίσματα	Τελικά Χαρακτηριστικά
Επιταχυνσιόμετρο	x	Mean, STD, Max, Min, Percentile25, Percentile50, Percentile75, Kurtosis, Skewness, Amplitude1, Amplitude2, Frequency2, Mean Frequency
	y	Mean, STD, Max, Min, Percentile25, Percentile50, Percentile75, Kurtosis, Skewness, Amplitude1, Frequency2
	magnitude	Mean, STD, Max, Min, Percentile25, Percentile50, Percentile75, Kurtosis, Skewness, Amplitude, Frequency2
Γυροσκόπιο	x	Mean, Max, Min, Percentile75, Kurtosis, Skewness, Amplitude1, Frequency2, Mean Frequency
	y	Mean, Min, Kurtosis, Skewness, Frequency2
	magnitude	Mean, Min, Kurtosis, Skewness, Frequency2

Μεθοδολογία

Εξαγωγή Χαρακτηριστικών

Χειρονομίες



Είδος Gesture	Τελικά Χαρακτηριστικά
Tap	Duration
Swipe	Duration, Mean X, Mean Y, Trace Length, Trace Projection, Start Velocity, Stop Velocity, Horizontal Acceleration, Vertical Acceleration, Slope, Mean Square Error, Coefficient of Determination

Ταξινομητές

Τι γνωρίζουμε;

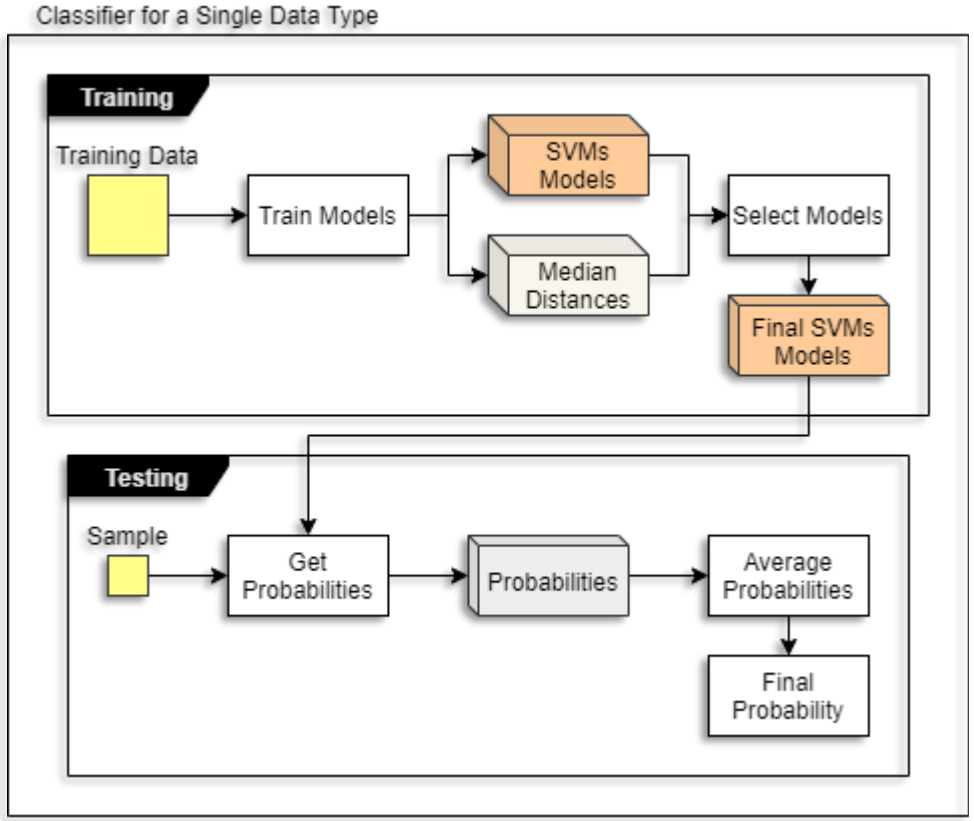
- Πρόβλημα ταξινόμησης μίας κλάσης
- Επίλυση με RBF-OCSVM
- Αδύνατη η χρήση ενός μοντέλου ανά ταξινομητή
- Οι παράμετροι (nu, gamma) επηρεάζουν τα RBF-OCSVMs

Τι προτείνουμε;

- Χρήση πολλαπλών RBF-OCSVMs, ανά ταξινομητή
- Χρήση εύρους τιμών για τις παραμέτρους
- Συλλογική τελική απόφαση

Ερωτήματα:

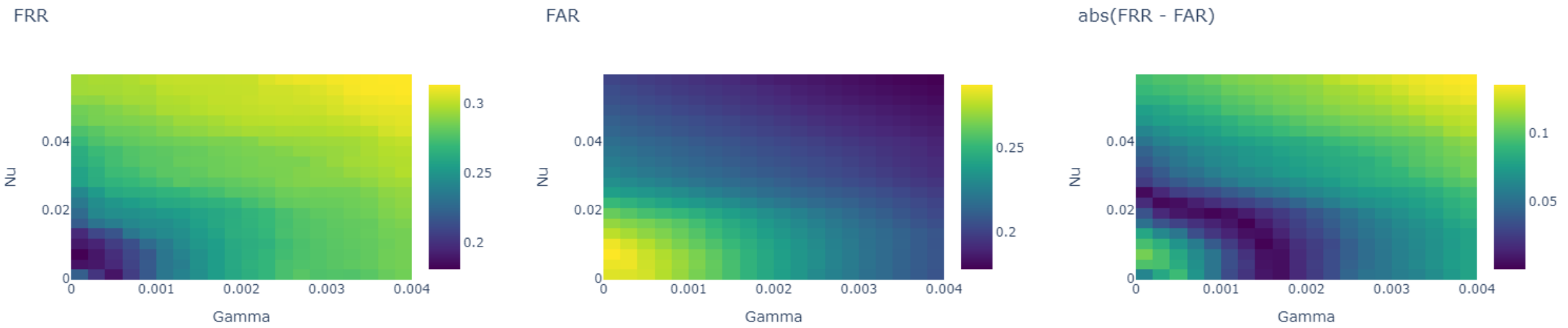
- Εύρος παραμέτρων
- Αριθμός μοντέλων που αποφασίζουν



Μεθοδολογία

Ταξινομητές

Εύρος Παραμέτρων

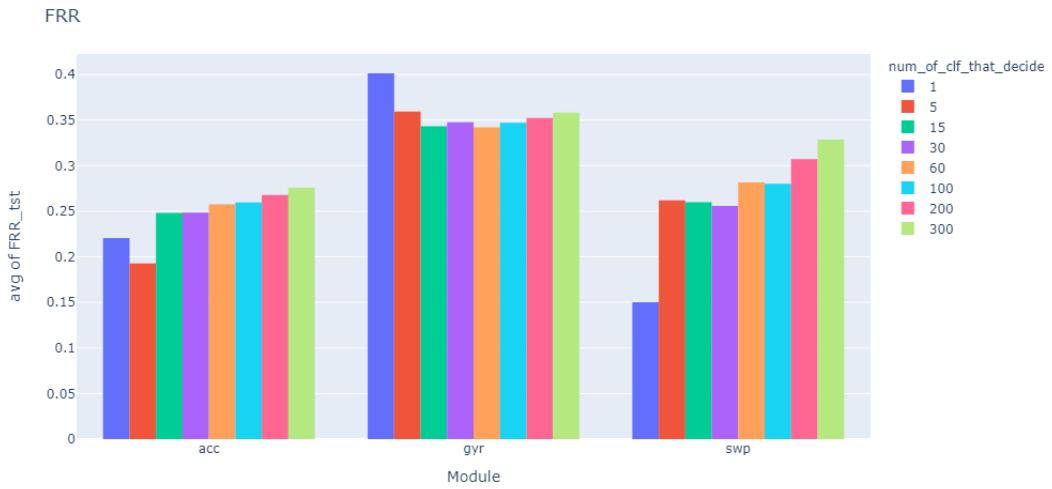


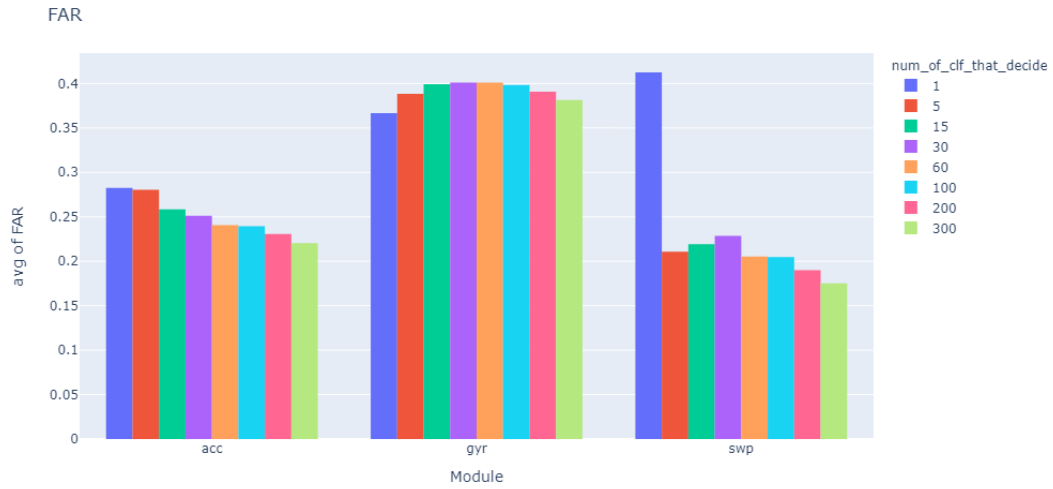
Κατηγορία	Nu			Gamma		
	Αρχική Τιμή	Τελική Τιμή	Βήμα	Αρχική Τιμή	Τελική Τιμή	Βήμα
Επιταχυνσιόμετρο	0.001	0.06	0.003	0.0001	0.004	0.0002
Γυροσκόπιο	0.11	0.31	0.01	0.001	0.04	0.002
Swipes	0.01	0.21	0.01	0.001	0.06	0.003
Taps	0.02	0.6	0.03	0.7	0.795	0.005

Μεθοδολογία

Ταξινομητές

Αριθμός Μοντέλων

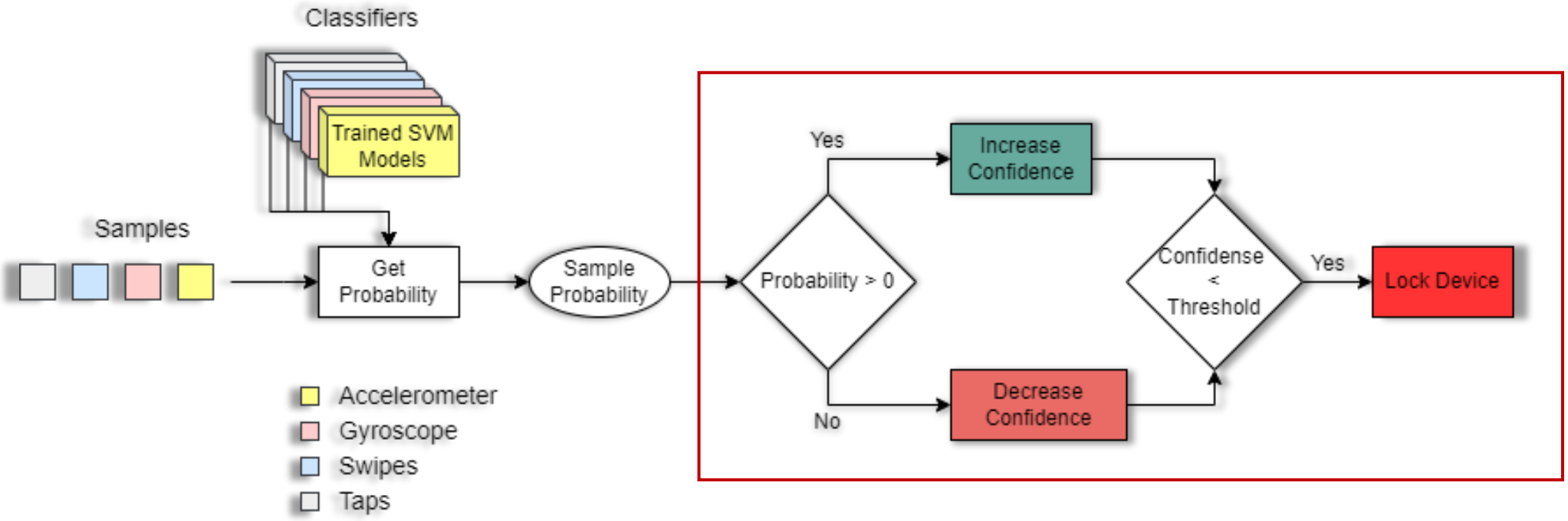




Κατηγορία	Βέλτιστος Αριθμός Μοντέλων
Επιταχυνσιόμετρο	30
Γυροσκόπιο	60
Swipes	60
Taps	60

Μεθοδολογία

Υποσύστημα Εμπιστοσύνης

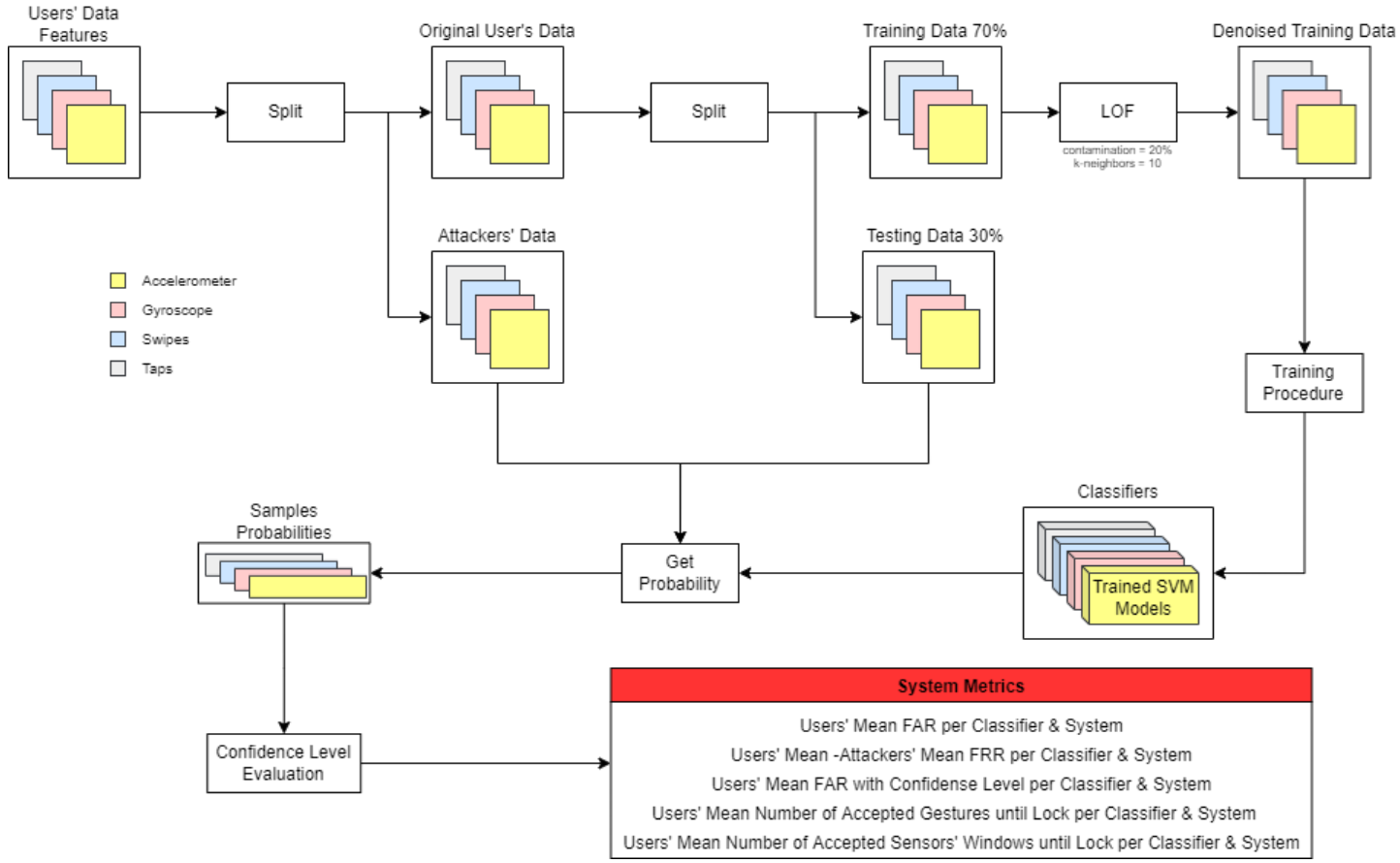


$$CL_n = \begin{cases} CL_{n-1} + PositiveStep(Game) * Weights(DataType) * abs(p), & p > 0 \\ CL_{n-1} + NegativeStep(Game) * Weights(DataType) * abs(p), & p \leq 0 \end{cases}$$

Initial Confidence Level			60		
Threshold			35		
	Mathisis	Focus	Reacton	Speedy	Memoria
Negative Step	-15	-15	-15	-15	-15
Positive Step	+10	+10	+10	+10	+10

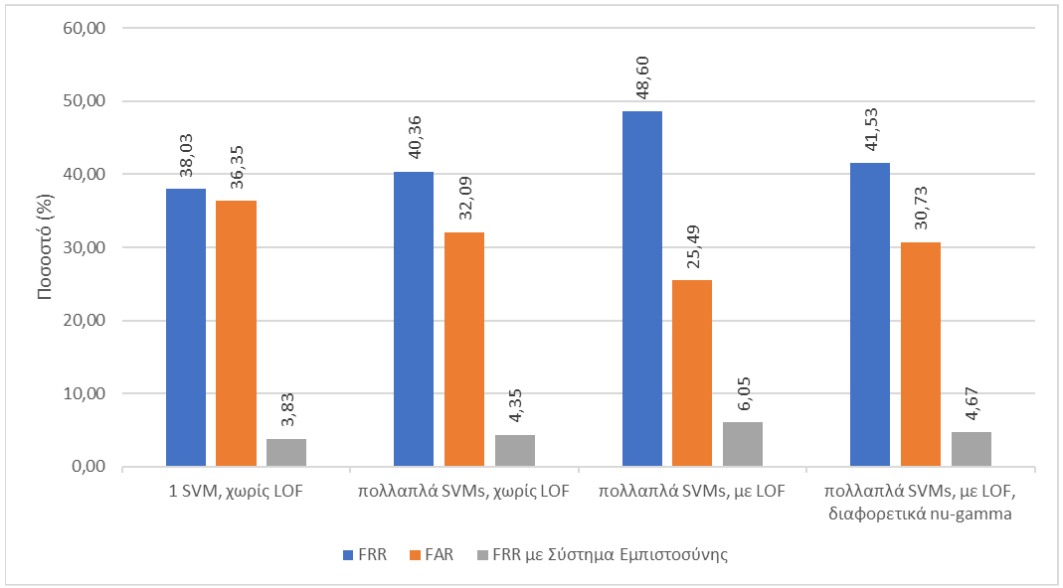
Μεθοδολογία

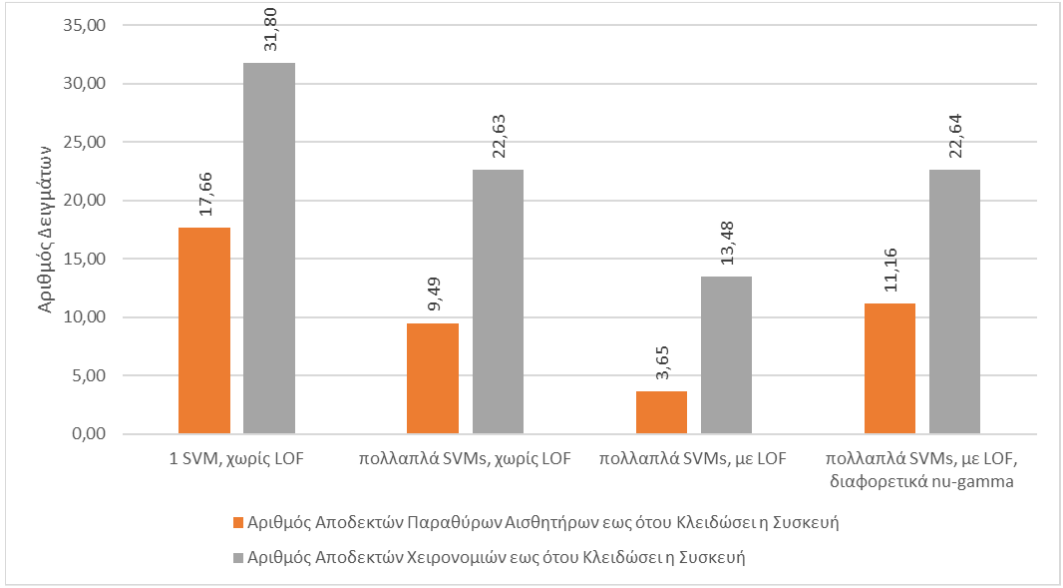
Σύνοψη Συστήματος – Δομή Τελικών Πειραμάτων



Αποτελέσματα

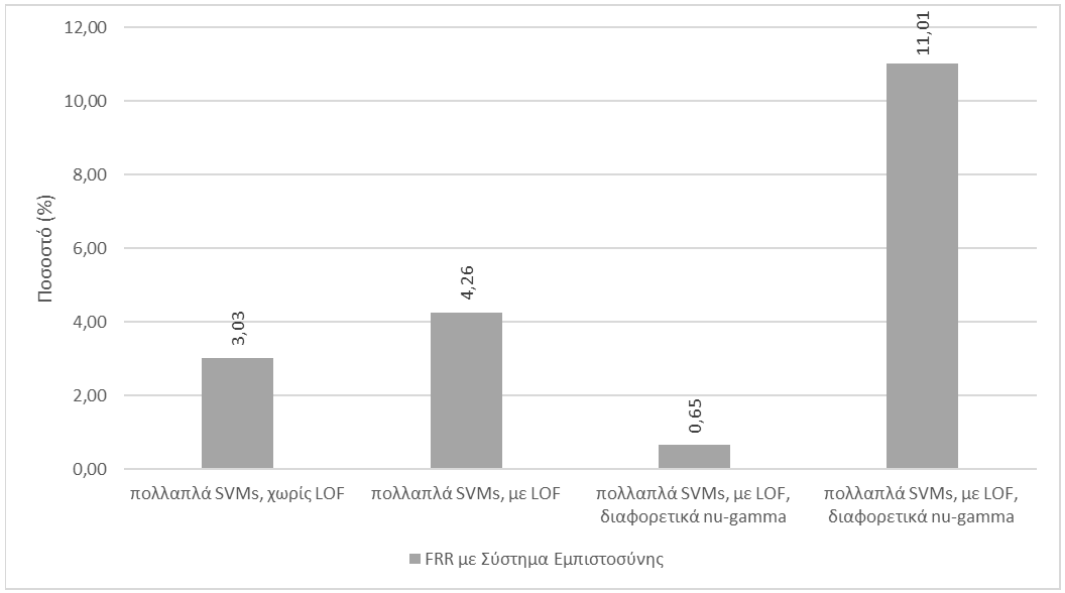
Σύστημα Εμπιστοσύνης – Πολλαπλά RBF-OCSVMs

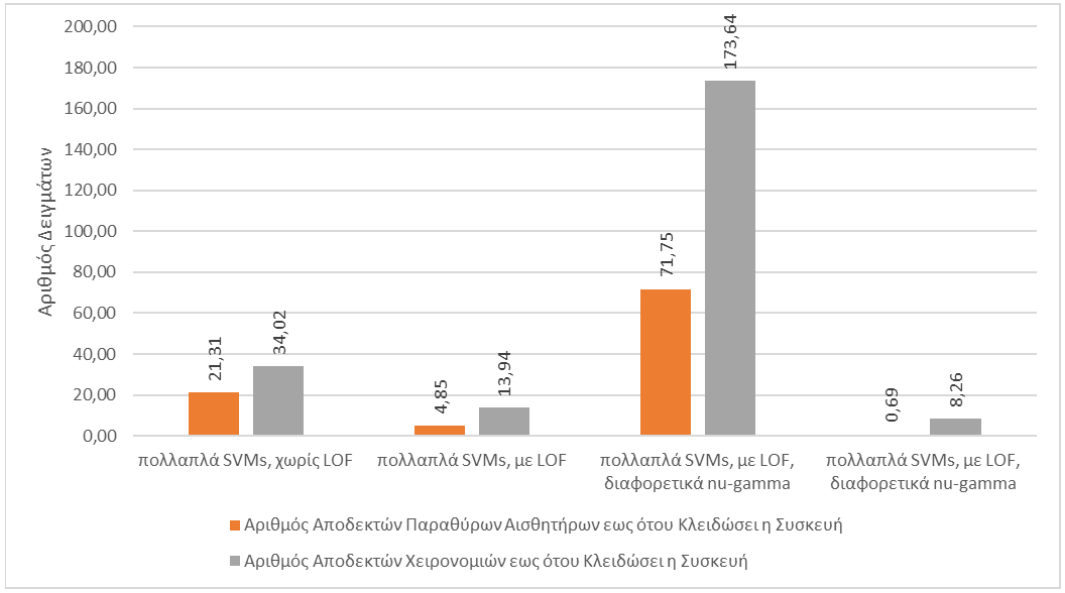




Αποτελέσματα

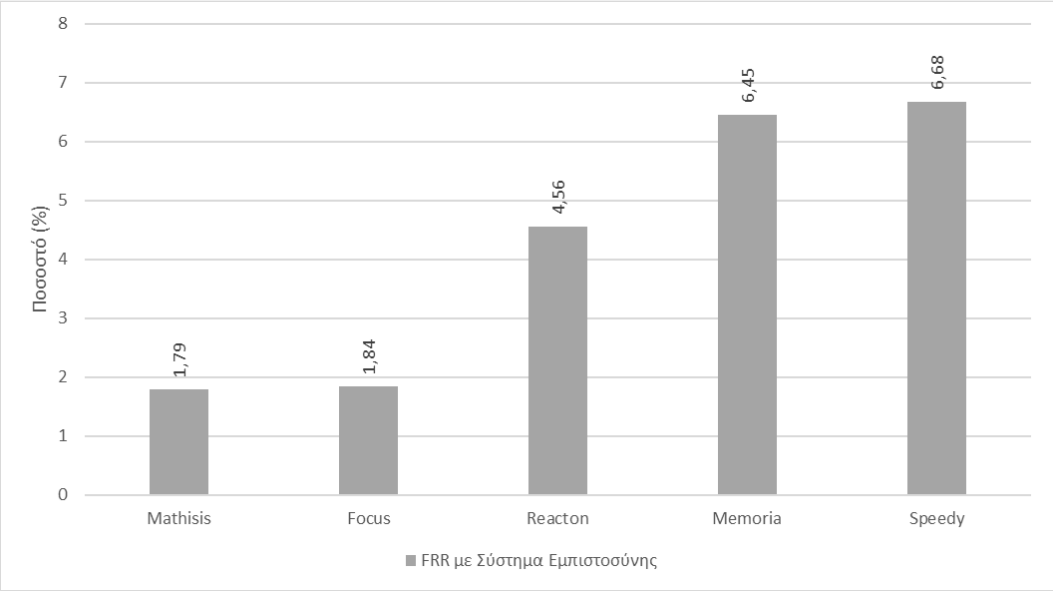
LOF – Περιοχές Nu-Gamma

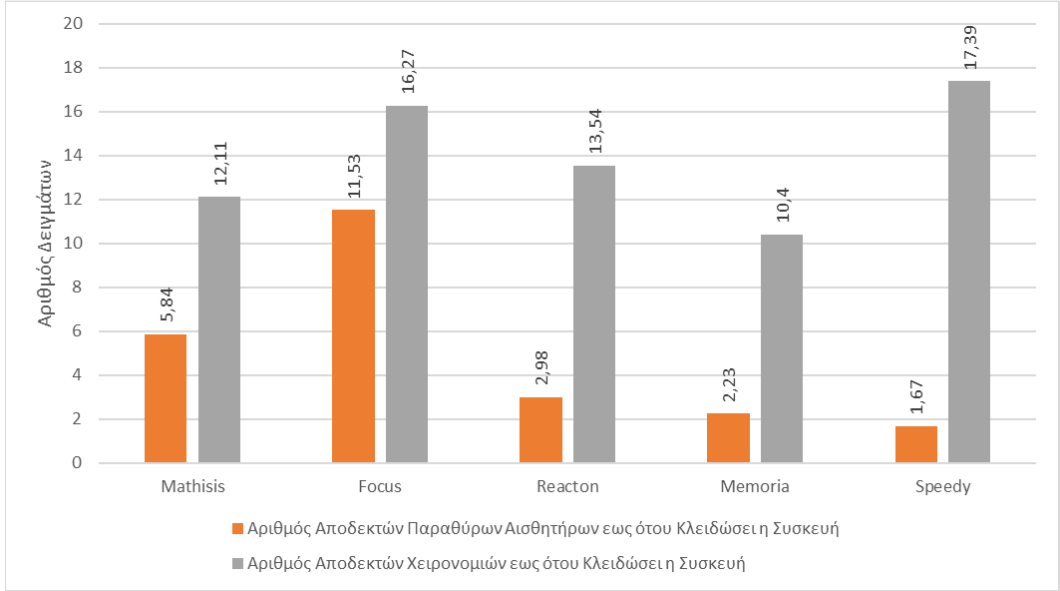




Αποτελέσματα

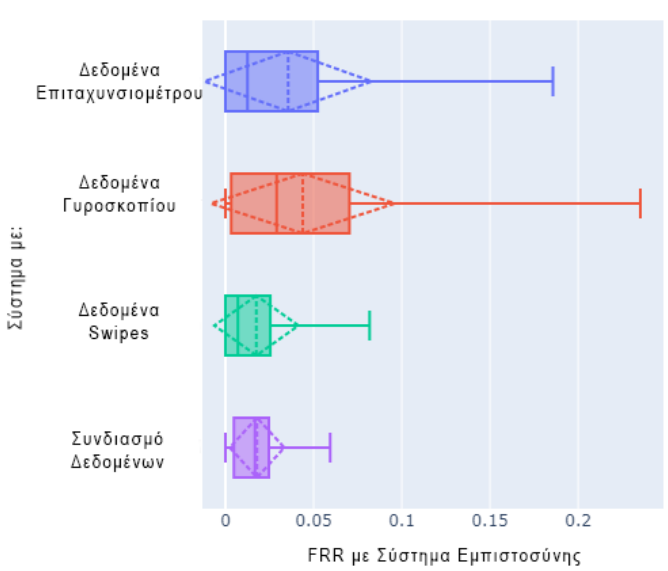
Ανά Παιχνίδι

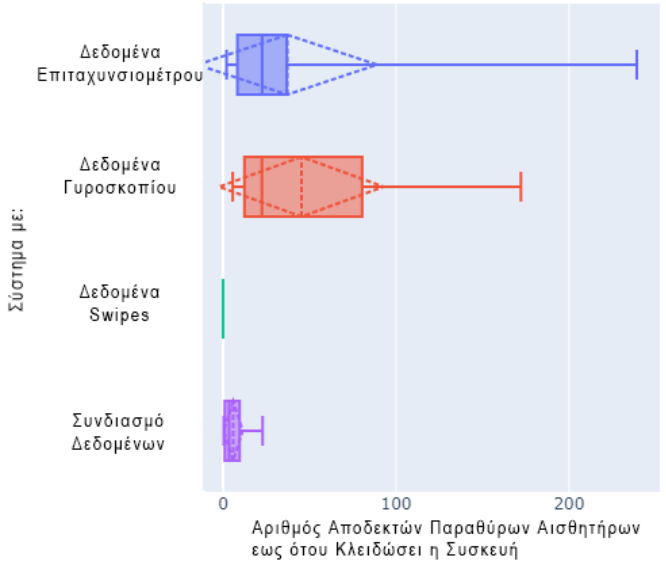


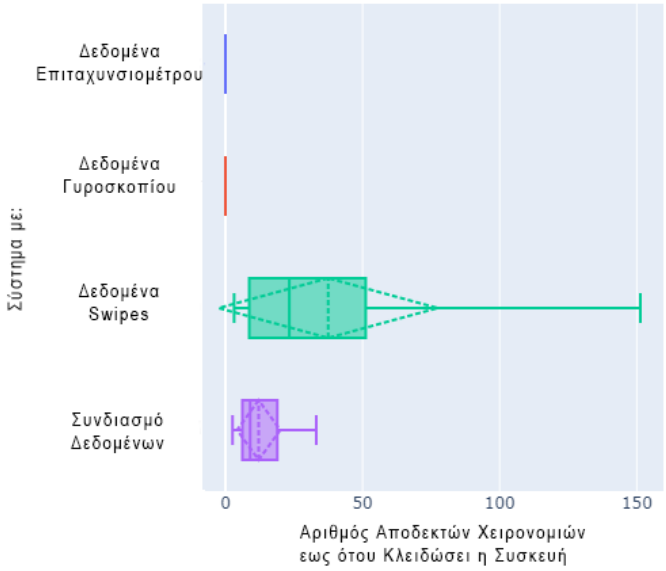


Αποτελέσματα

Συγκρίσεις (1)







Αποτελέσματα

Συγκρίσεις (2)

	Mathisis	Focus	Reacton	Memoria	Speedy
Σύστημα Αισθητήρων (FRR %)	5,20	6,00	4,30	5,70	5,70
Σύστημα Χειρονομιών (FRR με Σύστημα Εμπιστοσύνης %)	1,92	1,06	2,32 | 3,58 (Swipes | Taps)	3,44	0,065
Τρέχουσα Εργασία (FRR με Σύστημα Εμπιστοσύνης %)	1,79	1,84	4,56	6,45	6,68

	Mathisis	Focus	Reacton	Memoria	Speedy
Σύστημα Αισθητήρων (FAR %)	4,08	3,50	6,90	1,10	5,40
Σύστημα Χειρονομιών (Αριθμός Αποδεκτών Χειρονομιών)	1,70	3,92	8,08 | 11,37 (Swipes | Taps)	21,83	277,47
Τρέχουσα Εργασία (Αριθμός Αποδεκτών Δειγμάτων Αισθητήρων & Χειρονομιών)	5,84 & 12,11	11,53 & 16,27	2,98 & 13,54	2,23 & 10,40	1,67 & 17,39

Μέγεθος πακέτων αισθητήρων: 500 μετρήσεις

Μέγεθος πακέτων αισθητήρων: ~50 μετρήσεις

Συμπεράσματα

Μεθοδολογία & Τεχνικές

- Η χρήση πολλαπλών RBF-OCSVMs εξυπηρετεί την ασφάλεια του συστήματος.
- Το σύστημα εμπιστοσύνης βοηθάει στην διαμόρφωση ενός εύχρηστου συστήματος.
- Η αποθορυβοποίηση των δεδομένων εκπαίδευσης με LOF βελτιώσει την ασφάλεια.
- Οι παράμετροι nu και gamma των RBF-OCSVMs, παίζουν καθοριστικό ρόλο στην διασφάλιση ισορροπίας μεταξύ ασφάλειας και ευχρηστίας.

Σύστημα

- Ανθεκτικό σε σφάλματα μετρήσεων.
- Ικανοποιητικές μετρικές ασφάλειας και διαφάνειας.
- Γρήγορος έλεγχος
- Αντικειμενική αξιολόγηση

Μελλοντική Εργασία

Ιδέες

- Δυναμικά βάρη στους ταξινομητές.
- Δυνατότητα επιλογής περιοχών nu-gamma
- Συνδυασμός με τεχνικές επίγνωσης πλαισίου.
- Ικανότητα προσαρμογής στις αλλαγές συμπεριφοράς του ιδιοκτήτη.

Ευχαριστώ για την προσοχή σας!

Μηχανική Μάθηση

Εποπτευόμενη Μάθηση
(Supervised Learning)

Μη Εποπτευόμενη Μάθηση
(Unsupervised Learning)

Ημιεποπτευόμενη Μάθηση
(Semi-Supervised Learning)

Ενισχυτική Μάθηση
(Reinforcement Learning)

Ανίχνευση Ανωμαλιών – Ταξινόμηση Μίας Κλάσης

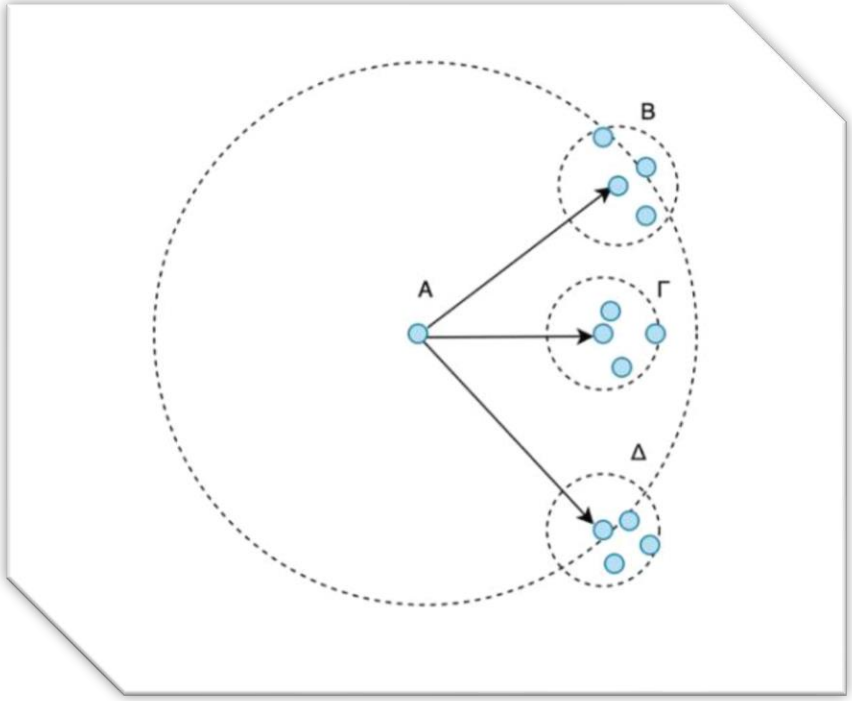
Ανίχνευση Ακραίων Δειγμάτων (Outlier Detection):

- Unsupervised
- Ανίχνευση Περιοχών Μεγάλης Πυκνότητας Δειγμάτων
- Αποθορυβοποίηση Δεδομένων
- Isolation Forest, Elliptic Envelope, Local Outlier Factor

Ανίχνευση Ασυνήθιστων Δειγμάτων (Novelty Detection):

- Semi-Supervised
- Οριοθέτηση Περιοχής Συνόλου Εκπαίδευσης
- Αποθορυβοποιημένα Σύνολα Εκπαίδευσης
- One Class Support Vector Machine

Local Outlier Factor (LOF)

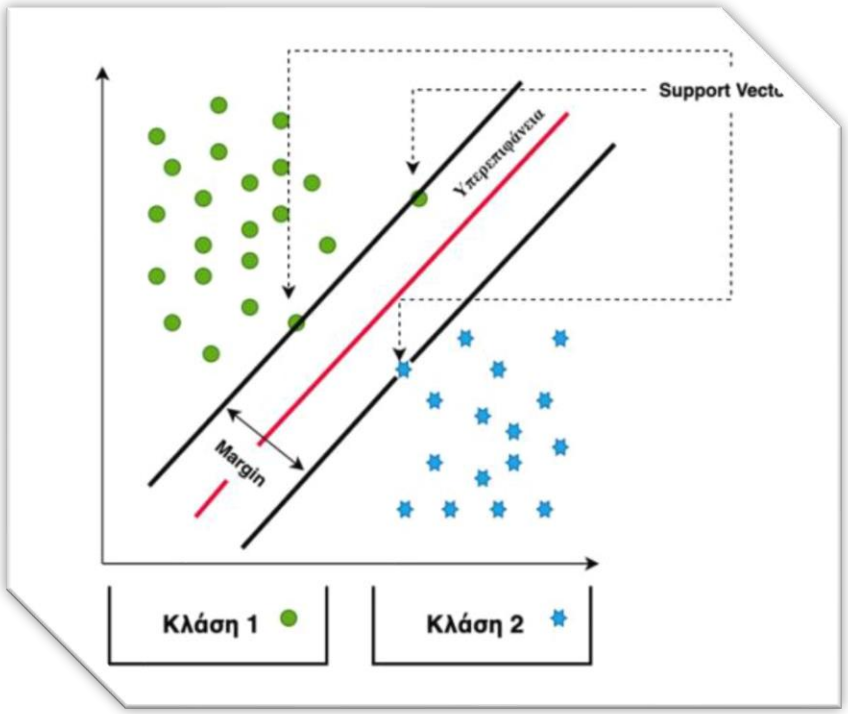


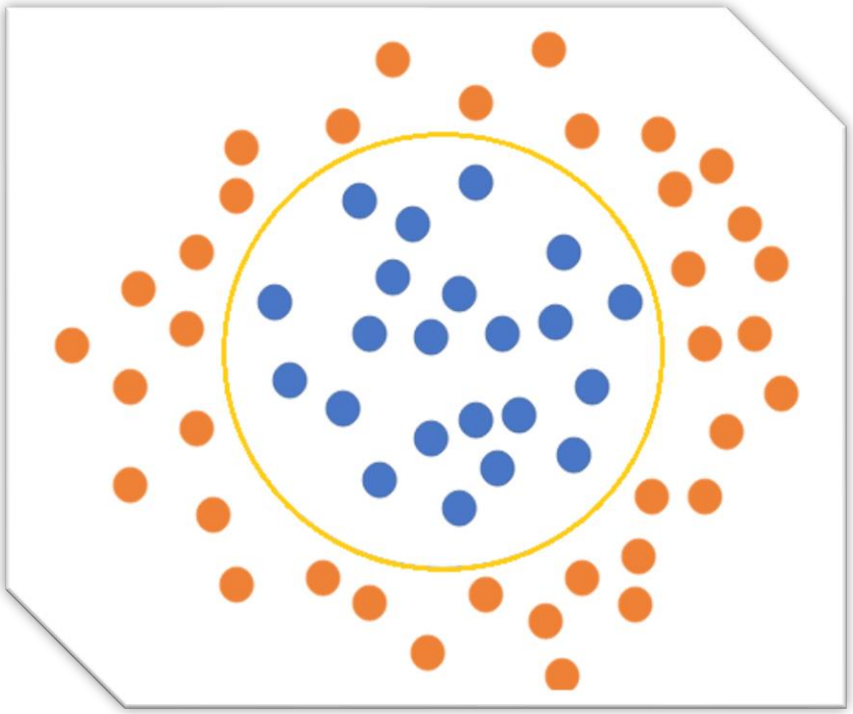
$$RD(Xi, Xj) = max(kDistance(Xj), Distance(Xi, Xj))$$

$$LDR_k(A) = \frac{1}{\sum_{Xj \in N_k(A)} \frac{RD(A, Xj)}{||\mathrm{N_k(A)}||}}$$

$$LOF_k(A) = \frac{\sum_{Xj \in N_k(A)} LRD_k(Xj)}{||\mathrm{N_k(A)}||} \times \frac{1}{LRD_k(A)}$$

One Class Support Vector Machine (OCSVM)





Μετρικές Αξιολόγησης

$$\text{False Rejection Rate} = \frac{FN}{TP + FN}$$

$$\text{False Acceptance Rate} = \frac{FP}{TN + FP}$$