



ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανάπτυξη Υβριδικού Μοντέλου ANFIS-Εξελικτικού
Αλγορίθμου για Ανίχνευση Εισβολών σε Δίκτυα

Του φοιτητή
Αριστείδη Τζήμκα
Αρ. Μητρώου: 2021168

Επιβλέπων
Παναγιώτης Αδαμίδης
Καθηγητής

Θεσσαλονίκη, Σεπτέμβριος 2026

Τίτλος Δ.Ε : Ανάπτυξη Υβριδικού Μοντέλου ANFIS-Εξελικτικού Αλγορίθμου για Ανίχνευση

Εισβολών σε Δίκτυα

Κωδικός Δ.Ε: 26122

Ονοματεπώνυμο φοιτητή: Αριστείδης Τζήμκας

Ονοματεπώνυμο εισηγητή: Παναγιώτης Αδαμίδης

Ημερομηνία ανάληψης Δ.Ε.: 22-02-2026

Ημερομηνία περάτωσης Δ.Ε.: 26-08-2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Αριστείδη Τζήμκα που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Σε όσους πιστεύουν στα όνειρά τους για ένα καλύτερο αύριο»

Πρόλογος

Η ιδέα της παρούσας διπλωματικής εργασίας πηγάζει από την ενασχόλησή μου με το μάθημα «Ευφυή Συστήματα» του τμήματος Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων της Πολυτεχνικής Σχολής του Διεθνούς Πανεπιστημίου της Ελλάδος. Οι εργασίες που εκπόνησα στο πλαίσιο του μαθήματος, μου προσέφεραν ένα ικανοποιητικό γνωστικό υπόβαθρο, το οποίο αξιοποίησα για την εκπόνηση της συγκεκριμένης διπλωματικής εργασίας. Τα Ασαφή Συστήματα και οι Εξελικτικοί Αλγόριθμοι μου κέντρισαν το ενδιαφέρον και αποφάσισα να συνδυάσω αυτούς τους δύο τομείς, με στόχο τη δημιουργία ενός υβριδικού συστήματος για την ανίχνευση εισβολών σε δίκτυα.

Περίληψη

Στη σύγχρονη εποχή, τα συνεχώς αυξανόμενα περιστατικά κυβερνοεπιθέσεων καθιστούν τα Συστήματα Ανίχνευσης Εισβολών (IDS) ζωτικής σημασίας για την ασφάλεια και τη λειτουργία των δικτύων. Παρόλο που τα σύγχρονα μοντέλα Μηχανικής Μάθησης προσφέρουν υψηλές επιδόσεις στην ανίχνευση απειλών, λειτουργούν συνήθως με περιορισμένη εγγενή ερμηνευσιμότητα, απαιτώντας πρόσθετες τεχνικές επεξήγησης. Η έλλειψη άμεσης διαφάνειας μειώνει την εμπιστοσύνη των αντίστοιχων αναλυτών ασφαλείας προς τα μοντέλα αυτά, καθώς δεν μπορούν να επαληθεύσουν τον λόγο πίσω από τις αποφάσεις τους.

Η παρούσα διπλωματική εργασία προτείνει ένα υβριδικό σύστημα ανίχνευσης εισβολών, το οποίο συνδυάζει την αρχιτεκτονική του Προσαρμοστικού Neuro-Ασαφούς Συστήματος Συμπερασμού (ANFIS) με τεχνικές βελτιστοποίησης μέσω Γενετικού Αλγορίθμου (GA). Κύριος στόχος είναι η δημιουργία ενός συστήματος που προσφέρει ικανοποιητική απόδοση με παράλληλη δυνατότητα ερμηνευσιμότητας. Το προτεινόμενο μοντέλο ANFIS-GA αξιολογήθηκε στο σύγχρονο σύνολο δεδομένων UNSW NB-15, τόσο σε δυαδική ταξινόμηση όσο και σε ταξινόμηση πολλαπλών κλάσεων. Για τη σωστή προεπεξεργασία των δεδομένων εφαρμόστηκαν καθιερωμένες τεχνικές, οι οποίες κατέστησαν τα δεδομένα έτοιμα για χρήση στη διαδικασία υλοποίησης του μοντέλου. Ο Γενετικός Αλγόριθμος έχει διττό ρόλο, καθώς χρησιμοποιείται για την επιλογή χαρακτηριστικών και για τη βελτιστοποίηση υπερπαραμέτρων του ANFIS. Τα πειραματικά αποτελέσματα καταδεικνύουν την επιτυχή εφαρμογή της προτεινόμενης αρχιτεκτονικής. Η λειτουργία του γενετικού αλγορίθμου οδήγησε σε σημαντική μείωση της διαστατικότητας των δεδομένων με παράλληλη αύξηση της απόδοσης. Το προτεινόμενο μοντέλο παρουσίασε μια πιο αυστηρή συμπεριφορά σε σύγκριση με τα υπόλοιπα πειραματικά μοντέλα, καταγράφοντας συγκριτικά μικρότερο πλήθος μη αναγνωρισμένων επιθέσεων. Επίσης, οι ασαφείς κανόνες που εξάγονται μέσω του ANFIS προσφέρουν διαφάνεια του τρόπου λήψης αποφάσεων του μοντέλου. Αυτό το πλεονέκτημα επιτρέπει την επαλήθευση και συνεχή αξιολόγηση της αξιοπιστίας του μοντέλου, γεγονός που μπορεί να διευκολύνει την αξιοποίησή του σε περιβάλλοντα όπου απαιτείται αυξημένη διαφάνεια στη λήψη αποφάσεων, όπως αυτό της κυβερνοασφάλειας. Έτσι, το προτεινόμενο μοντέλο αποτελεί μια υποσχόμενη επιλογή για την αξιοποίησή του ως ένα σύγχρονο IDS, προσφέροντας ανταγωνιστικές επιδόσεις ενώ παράλληλα συνδυάζει και τη δυνατότητα ερμηνευσιμότητας της λειτουργίας του.

Development of a Hybrid ANFIS-Evolutionary Algorithm Model for Intrusion Detection in Networks

Aristeidis Tzimkas

Abstract

In the modern era, the continuous rise in cyberattacks has made Intrusion Detection Systems (IDS) vital for network security and operation. Although modern Machine Learning models offer high performance in threat detection, they typically operate with limited interpretability, requiring additional explainability techniques. This lack of decision-making transparency diminishes security analysts' trust in these models, as they cannot verify the underlying reasoning behind the automated decisions.

This thesis proposes a hybrid intrusion detection system that combines the architecture of an Adaptive Neuro-Fuzzy Inference System (ANFIS) with optimization techniques utilizing a Genetic Algorithm (GA). The primary objective is to develop a system that delivers satisfactory performance alongside inherent interpretability. The proposed ANFIS-GA model was evaluated on the modern UNSW-NB15 dataset for both binary and multiclass classification tasks. Standard data preprocessing techniques were applied to prepare the dataset for the model implementation phase. The Genetic Algorithm serves a dual purpose, functioning both as a feature selection mechanism and as a hyperparameter optimizer for the ANFIS network. Experimental results demonstrate the successful application of the proposed architecture. The integration of the genetic algorithm led to a significant reduction in data dimensionality while simultaneously improving classification performance. The proposed model exhibited a stricter detection behavior compared to baseline models, recording a lower number of undetected attacks (false negatives). Furthermore, the fuzzy rules extracted by ANFIS provide transparency of the model's decision-making process. This advantage enables the verification and continuous assessment of the model's reliability, a feature that can facilitate its deployment in environments requiring heightened decision transparency, such as cybersecurity. Therefore, the proposed model represents a promising approach for modern IDS applications, offering competitive performance while maintaining the critical capability of operational interpretability.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή, κ. Παναγιώτη Αδαμίδα, για την πολύτιμη καθοδήγηση και τις γνώσεις που μου παρείχε για την ολοκλήρωση της παρούσας διπλωματικής εργασίας. Επίσης, θέλω να ευχαριστήσω την οικογένειά μου για τη στήριξη που μου έδειξε σε κάθε στάδιο της ζωής μου και, κατ' επέκταση, στη φοιτητική μου πορεία.

Περιεχόμενα

Πρόλογος	v
Περίληψη	vi
Abstract	vii
Ευχαριστίες	viii
Περιεχόμενα	ix
Κατάλογος Σχημάτων	xii
Κατάλογος Πινάκων	xii
Συντομογραφίες.....	xiii
Κεφάλαιο 1ο: Εισαγωγή.....	1
1.1 Κυβερνοασφάλεια και IDS	1
1.2 Προκλήσεις στα σύγχρονα IDS.....	1
1.3 Νευρο-Ασαφή Συστήματα και Εξελικτικοί Αλγόριθμοι	2
1.4 Σκοπός και στόχοι της Διπλωματικής Εργασίας	3
1.5 Συνεισφορά της Διπλωματικής Εργασίας.....	4
1.6 Δομή.....	4
Κεφάλαιο 2ο: Θεωρητικό υπόβαθρο	5
2.1 Συστήματα Ανίχνευσης Εισβολών (IDS)	5
2.2 Μηχανική Μάθηση σε Συστήματα Ανίχνευσης Εισβολών.....	6
2.3 Ασαφής Λογική και Ασαφή Συστήματα.....	7
2.4 Προσαρμοστικό Νευρο-Ασαφές Σύστημα Συμπερασμού (ANFIS).....	8
2.5 Εξελικτικοί Αλγόριθμοι.....	10
2.6 Στατιστική αξιολόγηση μοντέλων.....	12
Κεφάλαιο 3ο: Ανασκόπηση σχετικών εργασιών.....	14
3.1 Ανίχνευση εισβολών με Μηχανική Μάθηση	14
3.2 Ανίχνευση εισβολών με Βαθιά Μάθηση	14
3.3 Νευρο-Ασαφή Συστήματα για ανίχνευση εισβολών	14
3.4 Εξελικτική βελτιστοποίηση σε Νευρο-Ασαφή Συστήματα	15
3.5 Επιλογή χαρακτηριστικών σε Συστήματα Ανίχνευσης Εισβολών	15
3.6 Ερμηνευσιμότητα στην κυβερνοασφάλεια	16
3.7 Συζήτηση βιβλιογραφίας και ερευνητικό κενό	16
Κεφάλαιο 4ο: Περιγραφή και προεπεξεργασία δεδομένων	18
4.1 Ανάλυση του UNSW-NB15.....	18

4.1.1	Γενικές πληροφορίες του Συνόλου Δεδομένων.....	18
4.1.2	Κατηγορίες επιθέσεων	18
4.2	Διαχωρισμός δεδομένων.....	19
4.3	Τεχνικές προεπεξεργασίας δεδομένων	19
4.3.1	Καθαρισμός και επιλογή χαρακτηριστικών	19
4.3.2	One-Hot Encoding	20
4.3.3	Περιορισμός ακραίων τιμών και μετασχηματισμός.....	20
4.3.4	Αντιμετώπιση ανισορροπίας κλάσεων με SMOTE	21
4.3.5	Κανονικοποίηση δεδομένων	22
4.3.6	Πλαίσιο προεπεξεργασίας δεδομένων για δυαδική ταξινόμηση	22
4.3.7	Πλαίσιο προεπεξεργασίας δεδομένων για ταξινόμηση πολλαπλών κλάσεων.....	23
Κεφάλαιο 5ο:	Μεθοδολογία και υλοποίηση	25
5.1	Γενική αρχιτεκτονική συστήματος.....	25
5.2	Υλοποίηση του ANFIS.....	27
5.3	Υλοποίηση του Γενετικού Αλγορίθμου.....	28
5.4	Πειραματικό πρωτόκολλο για δυαδική ταξινόμηση.....	32
5.4.1	Μοντέλα αναφοράς.....	32
5.4.2	Υλοποίηση γενετικού αλγορίθμου για δυαδική ταξινόμηση.....	33
5.4.3	Σενάριο 1 – Βελτιστοποίηση υπερπαραμέτρων του ANFIS	34
5.4.4	Σενάριο 2 - Βελτιστοποίηση υπερπαραμέτρων και επιλογή χαρακτηριστικών.....	34
5.5	Πειραματικό πρωτόκολλο για ταξινόμηση πολλαπλών κλάσεων.....	35
5.5.1	Μοντέλα αναφοράς.....	36
5.5.2	Υλοποίηση γενετικού αλγορίθμου για ταξινόμηση πολλαπλών κλάσεων	36
5.5.3	Σενάριο 1 – Βελτιστοποίηση υπερπαραμέτρων.....	37
5.5.4	Σενάριο 2 – Βελτιστοποίηση υπερπαραμέτρων και επιλογή χαρακτηριστικών.....	37
5.6	Τελική διαδικασία εκπαίδευσης και αξιολόγησης για Σενάριο 1 & Σενάριο 2	37
Κεφάλαιο 6ο:	Πειραματικά αποτελέσματα και ανάλυση.....	39
6.1	Μετρικές αξιολόγησης.....	39
6.2	Αποτελέσματα δυαδικής ταξινόμησης	40
6.2.1	Συγκριτική αξιολόγηση μοντέλων.....	40
6.2.2	Ανάλυση επιλογής χαρακτηριστικών.....	42
6.2.3	Ανάλυση False Positives και False Negatives.....	42
6.3	Αποτελέσματα ταξινόμησης πολλαπλών κλάσεων	43
6.3.1	Συγκριτική αξιολόγηση μοντέλων.....	44
6.3.2	Ανάλυση επιλογής χαρακτηριστικών.....	46

6.3.3	Ανάλυση σπάνιων επιθέσεων	46
6.4	Στατιστική ανάλυση αποτελεσμάτων	48
6.4.1	Πολλαπλές εκτελέσεις πειραμάτων και περιγραφικά στατιστικά.....	48
6.4.2	Έλεγχος Friedman και μεταγενέστεροι στατιστικού έλεγχου	51
6.4.3	Σύγκριση αποτελεσμάτων δυαδικής ταξινόμησης και ταξινόμησης πολλαπλών κλάσεων 54	
6.5	Αξιολόγηση και επιλογή τελικού μοντέλου.....	56
6.6	Ερμηνεία και ανάλυση των κανόνων ANFIS	57
6.6.1	Ανάλυση Ασαφών Κανόνων	57
6.6.2	Συσχέτιση κανόνων με συμπεριφορές επιθέσεων	57
6.6.3	Ερμηνευσιμότητα του μοντέλου.....	60
Κεφάλαιο 7ο:	Συζήτηση αποτελεσμάτων	62
7.1	Συνολική αξιολόγηση του προτεινόμενου μοντέλου.....	62
7.2	Συμβιβασμοί μεταξύ απόδοσης και ερμηνευσιμότητας.....	62
7.3	Επίδραση της επιλογής χαρακτηριστικών (feature selection).....	62
7.4	Ανάλυση ψευδώς θετικών και ψευδώς αρνητικών αποτελεσμάτων	63
7.5	Συμπεριφορά σε σπάνιες επιθέσεις	64
7.6	Πιθανή ευελιξία σε Zero-day επιθέσεις.....	65
7.7	Περιορισμοί της εργασίας.....	65
Κεφάλαιο 8ο:	Συμπεράσματα.....	68
8.1	Συνολικά συμπεράσματα	68
8.2	Κύρια ερευνητικά ευρήματα.....	68
8.3	Μελλοντικές επεκτάσεις.....	69
ΒΙΒΛΙΟΓΡΑΦΙΑ.....		72
ΠΑΡΑΡΤΗΜΑ Α : Κανόνες IF-THEN		77

Κατάλογος Σχημάτων

Σχήμα 2.1: Γενική αρχιτεκτονική NIDS	5
Σχήμα 4.1: Διάγραμμα ροής προεπεξεργασίας δεδομένων.....	23
Σχήμα 5.1: Διάγραμμα ροής της αρχιτεκτονικής του προτεινόμενου συστήματος.....	25
Σχήμα 5.2: Αναπαράσταση των δύο προσεγγίσεων υβριδικού χρωμοσώματος για το Σενάριο 2	29
Σχήμα 5.3: Αναπαράσταση χρωμοσώματος για το Σενάριο 1	34
Σχήμα 6.1: Πίνακας σύγκρισης δυαδικής ταξινόμησης	43
Σχήμα 6.2: Πίνακας σύγκρισης ταξινόμησης πολλαπλών κλάσεων.....	48
Σχήμα 6.3: Στατιστική διακύμανση εκτελέσεων του μοντέλου για δυαδική ταξινόμηση	49
Σχήμα 6.4: Στατιστική διακύμανση εκτελέσεων του μοντέλου για ταξινόμηση πολλαπλών κλάσεων	51
Σχήμα 6.5: Post-Hoc έλεγχος Conover στη δυαδική ταξινόμηση	52
Σχήμα 6.6: Post-Hoc έλεγχος Conover στην ταξινόμηση πολλαπλών κλάσεων.....	53
Σχήμα 6.7: Γενική σύγκριση αποτελεσμάτων δυαδικής ταξινόμησης.....	54
Σχήμα 6.8: Γενική σύγκριση αποτελεσμάτων ταξινόμησης πολλαπλών κλάσεων	55

Κατάλογος Πινάκων

Πίνακας 3.1: Συγκριτικός πίνακας σχετικών εργασιών	17
Πίνακας 4.1: Εγγραφές συνόλου δεδομένων	18
Πίνακας 4.2: Διαχωρισμός δεδομένων σε υποσύνολα.....	19
Πίνακας 4.3: Εξισορρόπηση δεδομένων εκπαίδευσης για την ταξινόμηση πολλαπλών κλάσεων.....	24
Πίνακας 6.1: Υπερπαράμετροι ANFIS των πειραμάτων δυαδικής ταξινόμησης.....	40
Πίνακας 6.2: Αποτελέσματα δυαδικής ταξινόμησης (βέλτιστες εκτελέσεις).....	41
Πίνακας 6.3: Υπερπαράμετροι ANFIS των πειραμάτων ταξινόμησης πολλαπλών κλάσεων.....	45
Πίνακας 6.4: Αποτελέσματα ταξινόμησης πολλαπλών κλάσεων (βέλτιστες εκτελέσεις)	45
Πίνακας 6.5: Αναλυτικά στατιστικά αξιολόγησης για κάθε κλάση στο σενάριο 2 (Free 42).....	47
Πίνακας 6.6: Αποτελέσματα πολλαπλών εκτελέσεων του πειράματος δυαδικής ταξινόμησης Σενάριο 2 – 100 χαρακτηριστικών.....	49
Πίνακας 6.7: Αποτελέσματα πολλαπλών εκτελέσεων του πειράματος ταξινόμησης πολλαπλών κλάσεων Σενάριο 2 – 42 χαρακτηριστικών	50

Συντομογραφίες

IDS	Intrusion Detection System / Σύστημα Ανίχνευσης Εισβολών
SOC	Security Operations Center / Κέντρο Επιχειρήσεων Ασφαλείας
ANFIS	Adaptive Neuro-Fuzzy Inference System / Προσαρμοστικό Νευρο-Ασαφές Σύστημα Συμπερασμού
FCM	Fuzzy C-Means
GA	Genetic Algorithm / Γενετικός Αλγόριθμος
NIDS	Network-based Intrusion Detection System / Σύστημα Ανίχνευσης Εισβολών βασισμένο σε δίκτυο
CSOC	Cyber Security Operations Center / Κέντρα Επιχειρήσεων Κυβερνοασφάλειας
FIS	Fuzzy Inference System / Συστήματα Ασαφούς Συμπερασμού
TSK	Takagi-Sugeno-Kang
LSE	Least Square Estimation / Μέθοδος Ελαχίστων Τετραγώνων
STD	Standard Deviation / Τυπική Απόκλιση
SVM	Support Vector Machines / Μηχανές Διανοσμάτων Στήριξης
FGA	Fast Genetic Algorithm
PR-AUC	Precision-Recall Area Under Curve
SMOTE	Synthetic Minority Over-sampling Technique
ANN	Artificial Neural Network / Τεχνητό Νευρωνικό Δίκτυο
PSO	Particle Swarm Optimization
RF	Random Forest
SFS	Sequential Feature Selection

Κεφάλαιο 1ο: Εισαγωγή

1.1 Κυβερνοασφάλεια και IDS

Στη σύγχρονη εποχή, η ραγδαία εξέλιξη της τεχνολογίας και η ψηφιακή μετάβαση έχουν αλλάξει τον τρόπο λειτουργίας των κοινωνιών. Τεχνολογίες όπως το Υπολογιστικό Νέφος (Cloud Computing), το Διαδίκτυο των Πραγμάτων, τα δίκτυα υψηλής ταχύτητας και γενικά η ευρεία υιοθέτηση του διαδικτύου από την κοινωνία, οδήγησαν σε εκθετική αύξηση του όγκου δεδομένων που μεταφέρονται μέσω του διαδικτύου. Σε αυτό το σύγχρονο δικτυακό περιβάλλον, η αξία της μεταφερόμενης πληροφορίας έχει επίσης αυξηθεί ραγδαία. Η κυβερνοασφάλεια αποτελεί αναγκαιότητα για τη φυσιολογική λειτουργία κάθε οργανισμού, με σκοπό την αποτροπή απειλών κατά της ακεραιότητας και της ασφάλειας των δεδομένων.

Παράλληλα με την εξέλιξη της τεχνολογίας, οι κυβερνοεπιθέσεις έχουν εξελιχθεί στην πολυπλοκότητα, τη συχνότητα καθώς και την επίδραση που έχουν στους στόχους τους. Οι σύγχρονες απειλές, μέσω της χρήσης προηγμένων και συνεχώς μεταβαλλόμενων τεχνικών, παρακάμπτουν τα μέτρα ασφαλείας των δικτύων, στοχεύοντας στην υποκλοπή δεδομένων, στη διακοπή διαθεσιμότητας υπηρεσιών και γενικά σε κακόβουλες ενέργειες. Τα παραδοσιακά συστήματα ασφαλείας, όπως τα τείχη προστασίας (Firewalls), αποτελούν την πρώτη γραμμή άμυνας ενός περιβάλλοντος. Ωστόσο, αυτά τα συστήματα αποδεικνύονται συχνά ανεπαρκή για την αντιμετώπιση απειλών, καθώς επικεντρώνονται κυρίως στον έλεγχο πρόσβασης και όχι στην ανάλυση της συμπεριφοράς της δικτυακής κίνησης.

Αυτό το κενό ασφαλείας καλύπτεται με την ενσωμάτωση Συστημάτων Ανίχνευσης Εισβολών (Intrusion Detection Systems – IDS). Τα IDS λειτουργούν ως μηχανισμοί παρακολούθησης και η αποστολή τους είναι η έγκαιρη και έγκυρη ανίχνευση κακόβουλων δραστηριοτήτων ή οποιασδήποτε συμπεριφοράς που ενδεχομένως αποτελεί απειλή. Στα σύγχρονα δίκτυα, ο ρόλος των IDS είναι καθοριστικός, καθώς ανιχνεύουν οποιαδήποτε απειλή δεν έχουν καταφέρει να περιορίσουν τα παραδοσιακά συστήματα ασφαλείας. Στα σύγχρονα IDS, η ανίχνευση βασίζεται στην εκπαίδευση μοντέλων μέσω διαφόρων μεθόδων της Μηχανικής Μάθησης (Machine Learning). Τα IDS παρέχουν την απαραίτητη πληροφορία στην ομάδα του Κέντρου Επιχειρήσεων Ασφαλείας (Security Operations Center – SOC), προκειμένου οι αναλυτές να ανταποκριθούν άμεσα και αποτελεσματικά σε κρίσιμα περιστατικά ασφαλείας.

1.2 Προκλήσεις στα σύγχρονα IDS

Παρά τη σημαντικότητά τους, τα σύγχρονα IDS αντιμετωπίζουν διάφορες προκλήσεις που μπορεί να επηρεάσουν την επιχειρησιακή τους απόδοση. Καθώς ο όγκος των μεταφερόμενων δεδομένων αυξάνεται, παράλληλα εμφανίζονται περισσότερες και πιο εξελιγμένες απειλές. Αρχικά, ένα από τα σημαντικότερα προβλήματα σχετίζεται με την ποιότητα και την ισορροπία των δεδομένων. Τα IDS χρησιμοποιούν μοντέλα Μηχανικής Μάθησης με τα οποία πραγματοποιείται εκπαίδευση σε ήδη υπάρχοντα δεδομένα. Τα δεδομένα αυτά συχνά εμφανίζουν συγκεκριμένα χαρακτηριστικά τα οποία μειώνουν την ποιότητα και την αποτελεσματικότητα της εκπαίδευσης. Η δικτυακή κίνηση συχνά χαρακτηρίζεται από ακραία ανισορροπία κλάσεων (class imbalance), καθώς η πλειοψηφία των εγγράφων αφορά φυσιολογική δικτυακή κίνηση, ενώ αντίθετα οι κακόβουλες εγγραφές που αντιστοιχούν σε απειλές αποτελούν μειοψηφία. Αυτό το φαινόμενο είναι φυσιολογικό, ωστόσο ο μικρός όγκος πληροφορίας που σχετίζεται με τις απειλές δεν παρέχει την κατάλληλη πληροφορία για εκπαίδευση στα μοντέλα ώστε να γίνεται αποτελεσματικότερη ανίχνευση επιθέσεων από το σύστημα [1][2].

Ακόμα μία πρόκληση των IDS αποτελεί η διαχείριση των λανθασμένων συναγερμών (false positives). Οι πολλοί λανθασμένοι συναγερμοί μπορούν να προκαλέσουν «κόπωση» στην ομάδα SOC, όπου οι φυσιολογικές δικτυακές κινήσεις ανιχνεύονται λανθασμένα ως απειλές [1][2]. Η μείωση των εσφαλμένων συναγερμών μπορεί να επιτευχθεί με ανάλογες προσαρμογές, οι οποίες όμως θα έχουν ως αντίτιμο την αύξηση των μη αναγνωρισμένων επιθέσεων (false negatives). Αν και το θεμιτό αποτέλεσμα είναι η μείωση και των δύο, αυτό ουσιαστικά είναι εξαιρετικά δύσκολο στην υλοποίηση. Ο συμβιβασμός (trade-off) ανάμεσα στους λανθασμένους συναγερμούς και τις μη αναγνωρισμένες επιθέσεις αποτελεί μία πρόκληση η οποία εξαρτάται από τις ανάγκες του περιβάλλοντος που εφαρμόζεται το IDS. Στον τομέα της κυβερνοασφάλειας, η μείωση των μη αναγνωρισμένων επιθέσεων με αναγκαστική αύξηση των λανθασμένων συναγερμών αποτελεί ένα συνηθισμένο και γνωστό συμβιβασμό, καθώς η ελαχιστοποίηση των χαμένων επιθέσεων αποτελεί προτεραιότητα [3]-[5].

Η χρήση Μηχανικής Μάθησης δημιουργεί μια κρίσιμη πρόκληση που αφορά την έλλειψη ικανότητας επεξηγησιμότητας (explainability) από τα μοντέλα. Τα περισσότερα μοντέλα Μηχανικής Μάθησης παρέχουν πολύ υψηλές αποδόσεις, ωστόσο διαθέτουν περιορισμένη εγγενή ερμηνευσιμότητα (inherent interpretability). Η εσωτερική λογική τους δεν είναι άμεσα αναγνώσιμη σε μορφή ενός συνόλου κανόνων και συνήθως απαιτούν πρόσθετες μεταγενέστερες τεχνικές επεξήγησης (post-hoc explainability) για να αιτιολογήσουν τις αποφάσεις που λαμβάνουν. Η αδιαφάνεια αυτών των μοντέλων δημιουργεί ένα σημασιολογικό κενό (semantic gap) ανάμεσα στα αποτελέσματα του μοντέλου και την επαλήθευση ή παρακολούθηση της λογικής που χρησιμοποίησε το μοντέλο. Αυτό το κενό μπορεί να δημιουργήσει αρκετά προβλήματα. Όταν το IDS δεν μπορεί να τεκμηριώσει, με κατανοητό για τον άνθρωπο τρόπο, γιατί πήρε μια συγκεκριμένη απόφαση, δημιουργείται αυτόματα έλλειψη εμπιστοσύνης προς το σύστημα αφού ο εκάστοτε αρμόδιος ασφαλείας δεν μπορεί να ελέγξει την απόφαση αυτή. Η έλλειψη εμπιστοσύνης αποτελεί εμπόδιο στη λήψη αποφάσεων από την ομάδα SOC, αφού η αδιαφάνεια αποτρέπει την επαλήθευση μιας συμπεριφοράς για την οποία πρέπει να ληφθούν συγκεκριμένα μέτρα αντιμετώπισης. Στην περίπτωση των επιθέσεων μηδενικής ημέρας (zero-day), όπου εμφανίζονται πρωτόγνωρες απειλές οι οποίες δεν έχουν συμπεριληφθεί στην εκπαίδευση, είναι πολύ πιθανό να γίνει λάθος πρόβλεψη από το μοντέλο. Σε μια τέτοια λάθος ταξινόμηση επίθεσης, ο αναλυτής ασφαλείας δεν έχει τη δυνατότητα κατανόησης της δυσλειτουργίας του IDS λόγω της αδιαφάνειας. Ως αποτέλεσμα, δεν μπορεί να κάνει τις απαραίτητες διορθώσεις και προσαρμογές στο σύστημα ώστε να αποτραπεί μελλοντικά μια παρόμοια απειλή [2][6].

1.3 Νευρο-Ασαφή Συστήματα και Εξελικτικοί Αλγόριθμοι

Για την αντιμετώπιση του σημασιολογικού κενού που προκαλείται από τα μοντέλα «μαύρα κουτιά», η παρούσα εργασία στρέφεται στη χρήση των Νευρο-Ασαφών Συστημάτων (Neuro-Fuzzy Systems) και πιο συγκεκριμένα στο Προσαρμοστικό Νευρο-Ασαφές Σύστημα Συμπερασμού (Adaptive Neuro-Fuzzy Inference System – ANFIS). Το ANFIS συνδυάζει τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANNs) με την Ασαφή Λογική (Fuzzy Logic). Τα Τεχνητά Νευρωνικά Δίκτυα επιτρέπουν την εκπαίδευση σε πολύπλοκα και μη γραμμικά δεδομένα και η Ασαφής Λογική επιτρέπει τη μοντελοποίηση ασαφών προβλημάτων με τρόπο που προσεγγίζει τον ανθρώπινο τρόπο αντιμετώπισης των προβλημάτων αυτών. Το εξαιρετικά σημαντικό πλεονέκτημα που προσφέρει το ANFIS είναι η ερμηνευσιμότητα που παρέχει μέσω των ασαφών κανόνων, οι οποίοι φέρουν τη μορφή «ΑΝ-ΤΟΤΕ» (IF-THEN Rules) και παρέχουν μια περισσότερο άμεσα αναγνώσιμη αναπαράσταση της εσωτερικής λογικής του μοντέλου σε σύγκριση με μοντέλα περιορισμένης εγγενούς ερμηνευσιμότητας. Ωστόσο, το ANFIS παρουσιάζει κάποιες δυσκολίες απέναντι σε πολυδιάστατα δεδομένα όπως αυτά που αφορούν τη δικτυακή κίνηση. Συγκεκριμένα, η «κατάρρα της διαστατικότητας» (Curse of Dimensionality) στο

παραδοσιακό ANFIS εκδηλώνεται εντονότατα κατά την αρχικοποίηση του ασαφούς συστήματος με τη μέθοδο Διαμερισμού Πλέγματος (Grid Partitioning). Σε αυτή την περίπτωση, η αύξηση του αριθμού των χαρακτηριστικών (features) οδηγεί σε εκθετική αύξηση των παραγόμενων κανόνων. Μία τέτοια κατάσταση απαιτεί απαγορευτικά μεγάλη διαθεσιμότητα υπολογιστικών πόρων, καθώς το μοντέλο καλείται να επεξεργαστεί έναν τεράστιο και εξαιρετικά πολύπλοκο όγκο κανόνων.

Για την υπέρβαση αυτού του περιορισμού, καθίσταται αναγκαία η μείωση της πολυπλοκότητας. Αρχικά, για την αποτροπή της εκθετικής αύξησης των κανόνων, η αρχικοποίηση του συστήματος βασίζεται στην τεχνική ομαδοποίησης Fuzzy C-Means (FCM). Μέσω αυτής, η δημιουργία των κανόνων εξαρτάται αποκλειστικά από έναν προκαθορισμένο αριθμό συστάδων (clusters), επιτρέποντας στο μοντέλο να διατηρεί ένα διαχειρίσιμο πλαίσιο ανεξάρτητα από τη διαστατικότητα των δεδομένων εισόδου. Παράλληλα, για την αποτελεσματική πλοήγηση στον πολυδιάστατο χώρο αναζήτησης και τη βελτιστοποίηση του ίδιου του μοντέλου, το ANFIS συνδυάζεται με Εξελικτικό Αλγόριθμο (Evolutionary Algorithm – EA) και πιο συγκεκριμένα με Γενετικό Αλγόριθμο (Genetic Algorithm – GA). Οι Γενετικοί Αλγόριθμοι αποτελούν τεχνικές βελτιστοποίησης που χρησιμοποιούν διαδικασίες παρόμοιες με τη βιολογική εξέλιξη και φυσική επιλογή. Με αυτούς εξετάζονται πολλαπλές λύσεις μέσω μιας επαναληπτικής διαδικασίας επιλογής υποψήφιων λύσεων και αναπαραγωγής αυτών με στόχο τον εντοπισμό της συνολικά βέλτιστης λύσης. Στο πλαίσιο της συγκεκριμένης εργασίας, ο Γενετικός Αλγόριθμος έχει διττό ρόλο. Πρώτον, πραγματοποιεί επιλογή χαρακτηριστικών (feature selection), απορρίπτοντας την περιττή πληροφορία και κρατώντας μόνο τα απαραίτητα χαρακτηριστικά που αποδίδουν καλύτερα αποτελέσματα. Με αυτόν τον τρόπο μειώνεται σημαντικά η πολυπλοκότητα και οι επιπτώσεις του φαινομένου της «κατάρας της διαστατικότητας». Δεύτερον, πραγματοποιεί βελτιστοποίηση των υπερπαραμέτρων του ANFIS. Ο υβριδικός συνδυασμός ANFIS-GA επιτρέπει τη δημιουργία ενός συστήματος αυτοματοποιημένης βελτιστοποίησης απόδοσης, το οποίο εμφανίζει ταυτόχρονα περιορισμένη πολυπλοκότητα, υψηλή απόδοση και κυρίως διαθέτει ερμηνεύσιμους κανόνες, οι οποίοι είναι κατανοητοί από τα μέλη των ομάδων SOC.

1.4 Σκοπός και στόχοι της Διπλωματικής Εργασίας

Λαμβάνοντας υπόψη τις σύγχρονες προκλήσεις που εμφανίζουν τα IDS, ο σκοπός αυτής της εργασίας είναι η μελέτη και αξιολόγηση διάφορων προσεγγίσεων του υβριδικού ANFIS-GA μοντέλου για την ανίχνευση κυβερνοεπιθέσεων σε σύγχρονα δίκτυα. Βασικός σκοπός της παρούσας εργασίας είναι η ανάπτυξη ενός υβριδικού μοντέλου, το οποίο θα ανιχνεύει αποτελεσματικά τις κυβερνοεπιθέσεις διατηρώντας ταυτόχρονα την ερμηνευσιμότητα του τρόπου λειτουργίας του. Επιπλέον, επιδιώκεται η αξιολόγηση της συγκεκριμένης αρχιτεκτονικής σε ρεαλιστικές απειλές κυβερνοασφάλειας, καθώς και η ανάδειξη της συμβολής που έχει η επιλογή χαρακτηριστικών στην τελική απόδοση.

Για την επίτευξη αυτού του στόχου ορίστηκαν ορισμένοι επιμέρους στόχοι. Αρχικά επιδιώκεται η βελτιστοποίηση των υπερπαραμέτρων του ANFIS μέσω του Γενετικού Αλγορίθμου. Η βελτιστοποίηση των υπερπαραμέτρων του ANFIS στοχεύει στην αυτοματοποιημένη ρύθμιση των τιμών αυτών, χωρίς την ανάγκη χειροκίνητης και επαναληπτικής παραμετροποίησης. Η επιλογή χαρακτηριστικών επίσης μέσω του γενετικού αλγορίθμου στοχεύει στην αυτοματοποιημένη μείωση της διαστατικότητας των δεδομένων και τη μείωση του απαραίτητου υπολογιστικού φορτίου, μέσω της επιλογής συγκεκριμένων χαρακτηριστικών που προσφέρουν την κατάλληλη πληροφορία. Επίσης, στόχος αποτελεί η ικανοποιητική απόδοση του μοντέλου σε σπάνιες επιθέσεις, ανιχνεύοντάς τες σε μεγάλο ποσοστό. Ακόμα, επιδιώκεται η ικανότητα άμεσης ερμηνείας της λειτουργίας του μοντέλου, μέσω της εξαγωγής και ανάλυσης των παραγόμενων ασαφών κανόνων, καθιστώντας τη λειτουργία του μοντέλου διαφανή.

Τέλος, επιδιώκεται η συνολική αύξηση της απόδοσης και της καταλληλότητας του μοντέλου για προβλήματα κυβερνοασφάλειας σε σύγκριση με άλλες προσεγγίσεις.

1.5 Συνεισφορά της Διπλωματικής Εργασίας

Η συνεισφορά της συγκεκριμένης εργασίας πηγάζει από την προτεινόμενη υβριδοποίηση. Τα απλά νευρωνικά δίκτυα λειτουργούν ως «μαύρα κουτιά», ενώ το υβριδικό μοντέλο προσφέρει ερμηνευσιμότητα μέσω των ασαφών κανόνων αλλά και αυτοματοποιημένη εξέλιξη και βελτιστοποίηση μέσω του γενετικού αλγορίθμου. Στον τομέα της κυβερνοασφάλειας, τα σημερινά IDS παρουσιάζουν σημαντική έλλειψη ερμηνευσιμότητας των αποφάσεών τους. Η χρήση του υβριδικού μοντέλου μπορεί να προσφέρει την υψηλή απόδοση που παρέχουν τα αντίστοιχα κλασικά μοντέλα Μηχανικής Μάθησης, ενώ παράλληλα υπάρχει η κατάλληλη ερμηνευσιμότητα. Επίσης, σε αντίθεση με μεγάλο μέρος της βιβλιογραφίας, η παρούσα εργασία αξιολογεί την απόδοση του προτεινόμενου μοντέλου μέσω ενός πλαισίου αξιολόγησης με πολλαπλές εκτελέσεις και χρήση στατιστικών ελέγχων, ώστε να διασφαλιστεί η αξιοπιστία των αποτελεσμάτων. Τέλος, η σύγκριση του προτεινόμενου μοντέλου με τον αλγόριθμο Random Forest επιτρέπει την αξιολόγηση απέναντι σε ένα σημαντικά αποδοτικό μοντέλο.

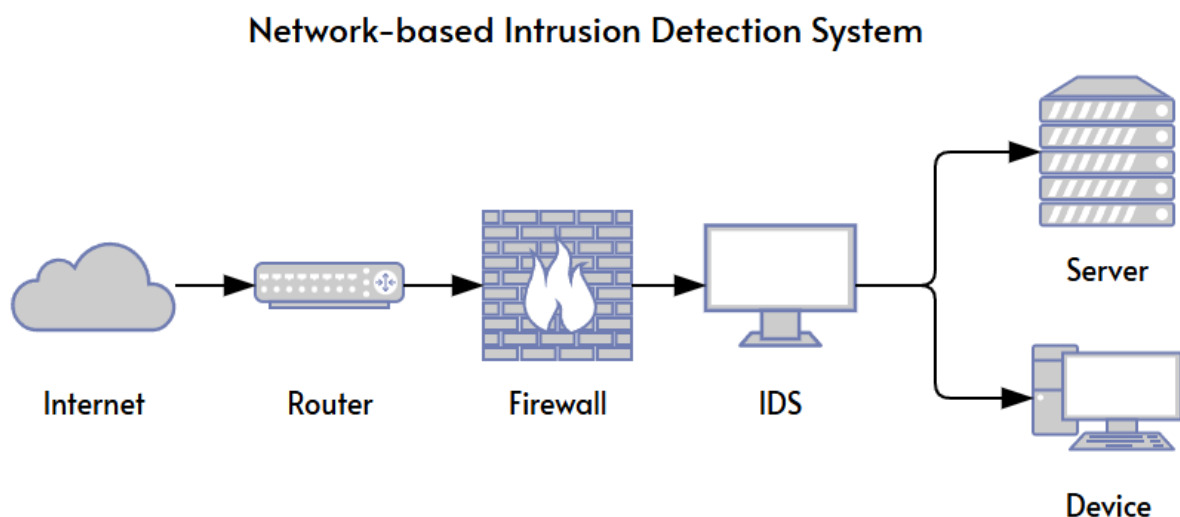
1.6 Δομή

Η παρούσα εργασία οργανώνεται σε οκτώ κεφάλαια. Το 1^ο Κεφάλαιο, αποτελεί το εισαγωγικό τμήμα της εργασίας όπου παρουσιάζεται ο γενικός στόχος. Στο 2^ο Κεφάλαιο, παρουσιάζεται το θεωρητικό υπόβαθρο, το οποίο αποτελεί βασικό τμήμα για την κατανόηση αλλά και την υλοποίηση του προτεινόμενου υβριδικού μοντέλου. Στο 3^ο Κεφάλαιο γίνεται μια ανασκόπηση σχετικών εργασιών με σχολιασμό ομοιοτήτων αλλά και διαφοροποιήσεων συγκριτικά με την παρούσα εργασία. Στο 4^ο Κεφάλαιο ξεκινά ουσιαστικά η υλοποίηση του μοντέλου με την προεπεξεργασία των δεδομένων. Σε αυτό, αναλύεται η διαδικασία προετοιμασίας των δεδομένων με στόχο τη χρήση τους στα επόμενα στάδια. Το 5^ο Κεφάλαιο αναλύει λεπτομερώς την αρχιτεκτονική του προτεινόμενου υβριδικού μοντέλου ANFIS-GA, τη σχεδίαση και τη διαδικασία υλοποίησής του. Στο 6^ο Κεφάλαιο παρουσιάζονται όλα τα πειραματικά αποτελέσματα, η στατιστική αξιολόγηση καθώς και η ανάλυση της ερμηνευσιμότητας των ασαφών κανόνων. Στο συγκεκριμένο κεφάλαιο γίνεται η σύγκριση της απόδοσης του προτεινόμενου μοντέλου με άλλες διαφορετικές προσεγγίσεις του πειραματικού πρωτοκόλλου, ενώ παράλληλα σχολιάζονται τα αποτελέσματα και ακολουθεί ένας γενικός απολογισμός ώστε να αναδειχθεί η βέλτιστη απόδοση μεταξύ των πειραμάτων. Στο 7^ο Κεφάλαιο πραγματοποιείται γενικός σχολιασμός των αποτελεσμάτων και η ανάδειξη της ανωτερότητας του προτεινόμενου μοντέλου μέσω της τεκμηριωμένης απόδοσης και ερμηνευσιμότητας που προσφέρει. Στο 8^ο και τελευταίο Κεφάλαιο, αναφέρονται τα συμπεράσματα της εργασίας καθώς και μερικές προτεινόμενες μελλοντικές επεκτάσεις.

Κεφάλαιο 2ο: Θεωρητικό υπόβαθρο

2.1 Συστήματα Ανίχνευσης Εισβολών (IDS)

Η αρχική ιδέα των συστημάτων ανίχνευσης εισβολών πρωτοεμφανίστηκε τη δεκαετία του 1960, ενώ η πρώτη γνωστή υλοποίηση πραγματοποιήθηκε το 1986 από την Dorothy Denning με το Intrusion Detection Expert System (IDES) [7]. Σήμερα, τα IDS χωρίζονται σε δύο κύριες αρχιτεκτονικές, οι οποίες είναι Host-based (HIDS) και Network-based (NIDS). Στην HIDS αρχιτεκτονική, η λειτουργία γίνεται σε μεμονωμένους υπολογιστές ή εξυπηρετητές (servers) με στόχο την επίβλεψη των εσωτερικών στοιχείων μόνο του εκάστοτε συστήματος. Αντίθετα, τα NIDS συστήματα ανίχνευσης αποτελούν ανεξάρτητες συσκευές ή λογισμικά τα οποία τοποθετούνται σε στρατηγικά σημεία ενός δικτύου για την επίβλεψη όλου του δικτύου. Συνήθως τοποθετούνται αμέσως μετά το Firewall και στη συνέχεια συνδέονται με switches που καταλήγουν στις συσκευές του δικτύου. Μια γενική απεικόνιση της αρχιτεκτονικής των NIDS παρουσιάζεται στο Σχήμα 2.1. Η παρούσα εργασία εντάσσεται πλήρως στο πλαίσιο της αρχιτεκτονικής NIDS, όπου αναλύεται το σύνολο εισερχόμενης και εξερχόμενης δικτυακής κίνησης προς τις συσκευές του δικτύου. Αναλύοντας τα πακέτα που μεταφέρονται, τα NIDS επιτρέπουν την ανίχνευση συμπεριφορών της δικτυακής κίνησης που παραπέμπουν σε απειλές κυβερνοασφάλειας.



Σχήμα 2.1: Γενική αρχιτεκτονική NIDS

Ανεξάρτητα με την αρχιτεκτονική τους, τα IDS κατηγοριοποιούνται περαιτέρω με βάση τη μεθοδολογία που ακολουθούν για την αναγνώριση απειλών. Η πρώτη κατηγορία είναι Signature-based IDS, όπου το σύστημα βασίζεται σε ήδη προϋπάρχουσες πληροφορίες, οι οποίες είναι γνωστές ως υπογραφές (signatures) και αντιστοιχούν σε συμπεριφορές απειλών. Μια τέτοια υπογραφή αντιστοιχεί σε συγκεκριμένο μοτίβο από χαρακτηριστικά όπως bytes, επικεφαλίδες πακέτων και άλλα τα οποία αθροιστικά αναγνωρίζονται ως επιθέσεις. Κατά τη λειτουργία του το Signature-based IDS συγκρίνει τη δικτυακή κίνηση με τα γνωστά μοτίβα επιθέσεων που γνωρίζει μέσω των πληροφοριών που κατέχει και αν εντοπίσει μια τέτοια συμπεριφορά την ταξινομεί κατάλληλα. Το πλεονέκτημα αυτής της κατηγορίας είναι η υψηλή ακρίβεια και ταχύτητα στον εντοπισμό γνωστών απειλών, παράγοντας σχετικά περιορισμένους λανθασμένους συναγερμούς (false positives). Ωστόσο, έχουν μειονέκτημα στη

δυνατότητα αναγνώρισης πρωτόγνωρων επιθέσεων (zero-day attacks), καθώς για αυτές δεν υπάρχουν οι κατάλληλες πληροφορίες για την αναγνώρισή τους, οδηγώντας σε λανθασμένες προβλέψεις [8].

Η δεύτερη κατηγορία είναι Anomaly-based IDS. Η συγκεκριμένη κατηγορία στηρίζεται στη δημιουργία ενός προφίλ που αντιστοιχεί στη φυσιολογική κίνηση και με την πάροδο του χρόνου προσαρμόζεται σε αυτό. Όταν μια εγγραφή-σύνδεση εμφανίζει διαφορετική συμπεριφορά από το φυσιολογικό προφίλ, ανιχνεύεται ως ανωμαλία και δημιουργείται αντίστοιχος συναγερμός ανίχνευσης απειλής. Αυτή η κατηγορία έχει πλεονέκτημα στην ανίχνευση zero-day επιθέσεων, όμως παρουσιάζει σημαντικά υψηλότερο πλήθος λανθασμένων συναγερμών. Επίσης, η λειτουργία των Anomaly-based IDS απαιτεί περισσότερη υπολογιστική ισχύ και μνήμη για την εκτέλεση πολύπλοκων αλγορίθμων. Ο συνδυασμός των Signature και Anomaly κατηγοριών αποτελεί συνηθισμένη πρακτική και η υβριδική IDS προσέγγιση παρέχει τα θετικά χαρακτηριστικά των δύο κατηγοριών[8].

Προκειμένου να ξεπεραστούν οι περιορισμοί των παραδοσιακών μεθόδων, το προτεινόμενο μοντέλο της παρούσας εργασίας λειτουργεί ως ένα ευφύες σύστημα που εξελίσσει τη λογική των Signature-based IDS. Αντί να βασίζεται σε στατικές υπογραφές, το μοντέλο εκπαιδεύεται σε ένα σύνολο δεδομένων προκειμένου να "μάθει" τα χαρακτηριστικά των επιθέσεων. Το πλεονέκτημα της προτεινόμενης αρχιτεκτονικής είναι η ενσωμάτωση της Ασαφούς Λογικής. Σε αντίθεση με τα παραδοσιακά συστήματα υπογραφών που αναζητούν αυστηρές και απόλυτες ταυτίσεις, η ασαφής λογική παρέχει στο μοντέλο την ικανότητα γενίκευσης (generalization), επιτρέποντάς του να αναγνωρίζει παραλλαγές γνωστών απειλών.

Τα σύγχρονα IDS καλούνται να εντοπίσουν σύγχρονες απειλές. Οι επιθέσεις στοχεύουν στην παραβίαση της Εμπιστευτικότητας (Confidentiality), της Ακεραιότητας (Integrity) και της Διαθεσιμότητας (Availability). Στις βασικές κατηγορίες κυβερνοεπιθέσεων συγκαταλέγονται οι επιθέσεις Denial-of-Service (DoS), οι οποίες υπερφορτώνουν το δίκτυο με στόχο τη μη διαθεσιμότητα των υπηρεσιών. Συχνά προηγούνται ενέργειες αναγνώρισης (Reconnaissance), με τις οποίες χαρτογραφείται το δίκτυο και εντοπίζονται ευπάθειες που χρησιμοποιούνται σε μεταγενέστερη στοχευμένη επίθεση. Γενικά, υπάρχουν πολλές κατηγορίες επιθέσεων που μπορούν να προκαλέσουν σημαντικά προβλήματα και περισσότερες πληροφορίες σχετικά με τις επιθέσεις του συνόλου δεδομένων που χρησιμοποιείται στην εργασία αναφέρονται σε επόμενο κεφάλαιο.

Για την αξιολόγηση των IDS χρησιμοποιούνται ορισμένες μετρικές που επιτρέπουν την αναγνώριση της αντικειμενικής τους απόδοσης. Οι μετρικές αυτές προέρχονται από τον τομέα της Μηχανικής Μάθησης. Μια γενική τεχνική αξιολόγησης αποτελεί το Confusion Matrix, με το οποίο αποτυπώνονται οι σωστές προβλέψεις του συστήματος. Επίσης, ακολουθούν μετρικές όπως Accuracy, Precision, Recall, F1-score και άλλες οι οποίες αναλύονται σε επόμενο κεφάλαιο.

2.2 Μηχανική Μάθηση σε Συστήματα Ανίχνευσης Εισβολών

Η αύξηση των απαιτήσεων κυβερνοασφάλειας στα σύγχρονα δίκτυα **κατέστησε** τις παραδοσιακές μεθόδους ανίχνευσης επιθέσεων ανεπαρκείς, καθιστώντας αναγκαία την υιοθέτηση τεχνικών Μηχανικής Μάθησης. Στο πλαίσιο των IDS, επικρατεί η προσέγγιση της Επιβλεπόμενης Μάθησης (Supervised Learning), όπου το μοντέλο εκπαιδεύεται χρησιμοποιώντας ένα σύνολο δεδομένων που περιέχει τις κατάλληλες πληροφορίες για την υπόδειξη φυσιολογικής ή κακόβουλης συμπεριφοράς. Μέσω αυτής της διαδικασίας, η λειτουργία του IDS μετατρέπεται σε ένα πρόβλημα ταξινόμησης (classification).

Η ταξινόμηση μπορεί να είναι είτε δυαδική, όπου διαχωρίζεται η φυσιολογική κίνηση από τις επιθέσεις, είτε να είναι ταξινόμηση πολλαπλών κλάσεων όπου το μοντέλο καλείται να αναγνωρίζει επιπλέον και

συγκεκριμένους τύπους επιθέσεων. Το σύνολο δεδομένων πρέπει να έχει την κατάλληλη αναπαράσταση της πληροφορίας με χαρακτηριστικά (features) τα οποία αντιστοιχούν σε διάφορα στοιχεία και πληροφορίες. Ένα κρίσιμο πρόβλημα που εμφανίζουν τα περισσότερα σύνολα δεδομένων είναι η ανισορροπία κλάσεων (class imbalance), όπου η πλειοψηφία των εγγραφών του συνόλου αποτελείται από φυσιολογική δικτυακή κίνηση το οποίο είναι απόλυτα φυσιολογικό και αντικατοπτρίζει την πραγματικότητα. Ωστόσο, σε αυτή την κατάσταση οι επιθέσεις λαμβάνουν ένα μικρό ποσοστό του συνόλου δεδομένων. Αυτό το πρόβλημα μπορεί να οδηγήσει σε ελλιπή όγκο πληροφορίας για την εκπαίδευση του συστήματος σε αυτές τις επιθέσεις και ακόμα μεγαλύτερο πρόβλημα υπάρχει για τις σπάνια εμφανιζόμενες επιθέσεις του ήδη περιορισμένου πλήθους εγγραφών επιθέσεων. Για την επίλυση αυτού του προβλήματος υπάρχουν λύσεις όπως η χρήση υπερδειγματοληψίας για τη δημιουργία επιπλέον εγγραφών που θα επιτρέψουν την αύξηση της ωφέλιμης πληροφορίας.

Για την επίλυση του προβλήματος ταξινόμησης των IDS έχουν προταθεί πολλοί αλγόριθμοι Μηχανικής Μάθησης. Ανάμεσα στις πιο δημοφιλείς προσεγγίσεις βρίσκονται οι Μηχανές Διανυσμάτων Στήριξης (Support Vector Machines – SVM), τα Δέντρα Απόφασης (Decision Trees) και το Τυχαίο Δάσος (Random Forest). Στο πλαίσιο του πειραματικού πρωτοκόλλου της εργασίας, χρησιμοποιήθηκε μόνο το Random Forest, καθώς φημίζεται για την υψηλή ακρίβεια που παρέχει με πολυδιάστατα δεδομένα και την αντοχή του στην υπερεκπαίδευση (overfitting), ενώ επίσης μπορεί να εφαρμοστεί τόσο για δυαδική ταξινόμηση όσο και για ταξινόμηση πολλαπλών κλάσεων [9]. Έτσι το Random Forest αποτελεί έναν γνωστό αλγόριθμο στον τομέα και χρησιμοποιείται στην παρούσα εργασία ως ένα ισχυρό μοντέλο αναφοράς (baseline model) για τη συγκριτική αξιολόγηση με το προτεινόμενο υβριδικό μοντέλο.

Παρά τις υψηλές αποδόσεις των κλασικών αλγορίθμων Μηχανικής Μάθησης, η εφαρμογή τους σε σύγχρονα Κέντρα Επιχειρήσεων Κυβερνοασφάλειας (Cyber Security Operations Center – CSOC) αντιμετωπίζει προβλήματα που σχετίζονται με τη λειτουργία των μοντέλων ως «μαύρα κουτιά» (black-box models). Τα περισσότερα κλασικά μοντέλα Μηχανικής Μάθησης λαμβάνουν αποφάσεις χωρίς την παροχή διαφάνειας και αιτιολόγησης των ενεργειών τους λόγω της υψηλής πολυπλοκότητας της λειτουργίας τους. Στον τομέα της κυβερνοασφάλειας η Ερμηνευσιμότητα (Interpretability) και η Επεξηγησιμότητα (Explainability) αποτελούν αναγκαιότητα, καθώς οι αναλυτές πρέπει να κατανοούν το πλαίσιο απόφασης του συστήματος ώστε να υπάρχει εμπιστοσύνη προς αυτό.

2.3 Ασαφής Λογική και Ασαφή Συστήματα

Η Ασαφής Λογική (Fuzzy Logic) έκανε την πρώτη εμφάνιση το 1920 από τους Lukasiewicz και Post όμως καθιερώθηκε το 1965 από τον Zadeh, ο οποίος παρουσίασε το μαθηματικό πλαίσιο στο οποίο δεν συμπεριλαμβάνονται αυστηρά μόνο τιμές της δίτιμης λογικής [10]. Η Ασαφής Λογική παρομοιάζεται με τον ανθρώπινο τρόπο λήψης αποφάσεων, όπου γίνεται συμβιβασμός ανάμεσα στη σημασία και την ακρίβεια. Στη δίτιμη λογική, μια λογική πρόταση μπορεί να είναι αληθής ή ψευδής, αποκλείοντας κάποια τρίτη λύση. Η Ασαφής Λογική είναι μια επέκταση της Δίτιμης Λογικής, η οποία όμως σχετίζεται με τη σημαντικότητα της ακρίβειας. Σε αυτή, μπορούν να υπάρχουν λύσεις ανάμεσα από το αληθές ή ψευδές, μοντελοποιώντας την κοινή ανθρώπινη λογική σκέψη. Η Ασαφής λογική προσφέρει εύκολη κατανόηση καθώς βασίζεται στη φυσική γλώσσα, είναι ευέλικτη, λειτουργεί με ακριβή και μη ακριβή δεδομένα και επιτρέπει τη μοντελοποίηση πολύπλοκων ασαφών προβλημάτων.

Η υλοποίηση της Ασαφούς Λογικής πραγματοποιείται με τα Συστήματα Ασαφούς Συμπερασμού (Fuzzy Inference Systems - FIS). Αυτά, στηρίζονται σε κανόνες ορισμένους από κάποιον ειδικό ο οποίος έχει τις κατάλληλες γνώσεις και ορίζει τις κατάλληλες παραμέτρους για το αντίστοιχο πρόβλημα παρέχοντας ένα πλαίσιο σωστής λειτουργία του συστήματος. Η χρήση Ασαφών Συνόλων (Fuzzy Sets),

διαφοροποιεί αυτά τα συστήματα από τα κλασικά. Στα κλασικά σύνολα, η μετάβαση ενός στοιχείου σε κάποιο σύνολο είναι απόλυτη. Για παράδειγμα, ένα bit μπορεί να ανήκει μόνο στο σύνολο 0 ή μόνο στο σύνολο 1. Στα ασαφή σύνολα, η μετάβαση αυτή είναι σταδιακή και ποσοτικοποιείται μέσω των Συναρτήσεων Συμμετοχής (Membership Functions), οι οποίες αποδίδουν σε κάθε στοιχείο έναν Βαθμό Συμμετοχής (Degree of Membership). Ένα παράδειγμα που μπορεί να περιγράψει αυτό το πλαίσιο είναι η αίσθηση της θερμοκρασίας. Στη κλασική λογική, ένα σύστημα θα μπορούσε να θέσει ένα αυστηρό και απόλυτο κατώφλι, με αυτό για παράδειγμα να είναι 20°C, κάτω από το οποίο η θερμοκρασία θα χαρακτηριζόταν ως «κρύα» και πάνω από αυτό ως «ζεστή». Αντίθετα, στην ασαφή λογική μια θερμοκρασία 25°C μπορεί να ανήκει ταυτόχρονα στο ασαφές σύνολο «κρύο» με ένα συγκεκριμένο βαθμό συμμετοχής και στο ασαφές σύνολο «ζεστό» με έναν άλλο βαθμό συμμετοχής. Για έναν άνθρωπο οι 25°C μπορεί να θεωρηθούν περισσότερο ως «ζέστη», ενώ για έναν άλλον μπορεί να θεωρηθούν περισσότερο ως «μέτρια» θερμοκρασία, επομένως η ταξινόμηση της θερμοκρασίας σε μια κατηγορία είναι υποκειμενική και τα ασαφή σύνολα με τις συναρτήσεις συμμετοχής και τους βαθμούς συμμετοχής αποτελούν λύση σε ένα τέτοιο πρόβλημα. Είναι σημαντικό να αναφερθεί πως ο βαθμός συμμετοχής εκφράζει την κατανομή δυνατότητας (possibility distribution) και όχι πιθανότητας. Ενώ η πιθανότητα αφορά τη στατιστική συχνότητα ενός γεγονότος, η δυνατότητα περιγράφει το κατά πόσο μια συγκεκριμένη τιμή ικανοποιεί την εννοιολογική περιγραφή ενός συνόλου.

Η διαδικασία της Ασαφοποίησης (Fuzzification) αποτελεί το αρχικό βήμα σε ένα FIS. Σε αυτό το στάδιο, οι αριθμητικές τιμές εισόδου μετατρέπονται σε ασαφείς όρους με συναρτήσεις συμμετοχής και βαθμούς συμμετοχής. Τα ασαφή σύνολα στα οποία αντιστοιχίζονται ονομάζονται Λεκτικοί Όροι (Linguistic Terms), όπως για παράδειγμα «χαμηλός», «μεσαίος» ή «υψηλός». Στη συνέχεια δημιουργούνται οι Ασαφείς Κανόνες (Fuzzy Rules) στη διαδικασία Ασαφούς Συνεπαγωγής (Fuzzy Implication), με τους κανόνες να έχουν μορφή AN-TOTE (IF-THEN). Οι κανόνες αυτοί επιτρέπουν στο σύστημα να συνδυάζει πολλαπλές παραμέτρους εισόδου. Τα συμπεράσματα όλων των κανόνων συναθροίζονται στο στάδιο Ασαφούς Συνάθροισης (Fuzzy Aggregation), οδηγώντας στο τελικό ασαφές συμπέρασμα. Τέλος, το παραγόμενο ασαφές συμπέρασμα μετατρέπεται ξανά σε μια σαφή, αριθμητική τιμή εξόδου μέσω της διαδικασίας της Αποασαφοποίησης (Defuzzification), η οποία αποτελεί και την τελική πρόβλεψη-απόφαση του συστήματος.

Τα Συστήματα Ασαφούς Συμπερασμού διακρίνονται σε δύο κύριες κατηγορίες, οι οποίες είναι το μοντέλο Mamdani και το μοντέλο Takagi-Sugeno-Kang (TSK). Το μοντέλο Mamdani προτάθηκε το 1974 για τη δημιουργία ενός συστήματος με σκοπό τον έλεγχο ατμομηχανών σε εργοστάσια [11]. Σε αυτό, κάθε κανόνας αποτελείται από συνθήκες ασαφών συνόλων και η έξοδος αποτελείται επίσης από ένα ασαφές σύνολο, έτσι η λογική παραμένει πλήρως γλωσσική με λεκτικούς όρους από την αρχή μέχρι το τέλος. Το μοντέλο TSK προτάθηκε το 1985 και διατηρεί την ίδια ασαφή λογική στο τμήμα συνθηκών των κανόνων αλλά διαφοροποιείται στην έξοδο του συστήματος [12]. Αντί για ένα ασαφές σύνολο, η έξοδος του συστήματος είναι μια γραμμική μαθηματική συνάρτηση των εισόδων. Αυτή η μαθηματική φύση κάνει τη διαδικασία αποασαφοποίησης εξαιρετικά γρήγορη και αποδοτική. Το μεγάλο πλεονέκτημα του TSK είναι ότι η μαθηματική του δομή επιτρέπει την ομαλή ενσωμάτωσή του με μεθόδους Μηχανικής Μάθησης.

2.4 Προσαρμοστικό Νευρο-Ασαφές Σύστημα Συμπερασμού (ANFIS)

Παρά την ικανότητα των Συστημάτων Ασαφούς Συμπερασμού (FIS) να μοντελοποιούν την ανθρώπινη λογική, η πρακτική τους εφαρμογή συναντά ένα σημαντικό εμπόδιο, γνωστό στη βιβλιογραφία ως το «πρόβλημα προσδιορισμού της δομής» (structure identification problem). Σε ένα παραδοσιακό FIS, η επιλογή των συναρτήσεων συμμετοχής, το πλήθος τους και η συγγραφή των ασαφών κανόνων

βασίζονται αποκλειστικά στη γνώση ενός ανθρώπου-ειδικού. Αυτή η προσέγγιση είναι χρονοβόρα, προϋποθέτει την ύπαρξη ενός κατάλληλου ειδικού και καθίσταται αδύνατη η εφαρμογή της στα σύγχρονα και πολυδιάστατα προβλήματα, όπως η ανάλυση της δικτυακής κίνησης για τη συγκεκριμένη εργασία. Σε μεγάλα σύνολα δεδομένων, τα μοτίβα που εμφανίζουν οι κυβερνοεπιθέσεις δεν είναι εύκολα αντιληπτά από τον άνθρωπο, καθιστώντας τον χειροκίνητο ορισμό των ασαφών κανόνων αδύνατο.

Η λύση του παραπάνω προβλήματος, προτάθηκε το 1993 από τον Jang με το Προσαρμοστικό Νευρο-Ασαφές Σύστημα Συμπερασμού (Adaptive Neuro-Fuzzy Inference System – ANFIS) [13]. Το ANFIS αποτελεί μια πρωτοποριακή υβριδική προσέγγιση, η οποία συνδυάζει τα συστήματα ασαφούς συμπερασμού με την ικανότητα μάθησης των Τεχνητών Νευρωνικών Δικτύων (Adaptive Neural Networks – ANNs). Ουσιαστικά, το ANFIS υλοποιεί ένα σύστημα ασαφούς συμπερασμού τύπου TSK μέσα στη δομή ενός τεχνητού νευρωνικού δικτύου. Αυτό του επιτρέπει να εξάγει αυτόματα τους ασαφείς κανόνες από τα δεδομένα και να ρυθμίζει δυναμικά τις παραμέτρους του χωρίς την ανάγκη ανθρώπινης παρέμβασης.

Η αρχιτεκτονική του ANFIS αποτελείται από πέντε διαδοχικά επίπεδα (layers) επεξεργασίας τα οποία αναφέρονται παρακάτω:

- **1^ο Επίπεδο Ασαφοποίησης (Fuzzification Layer):** Στο πρώτο επίπεδο, οι καθαρές αριθμητικές τιμές κάθε εισόδου μετατρέπονται σε ασαφείς λεκτικούς όρους (πχ. «λίγο», «μεσαίο»). Αυτή η διαδικασία γίνεται μέσω παραμετροποιήσιμων συναρτήσεων συμμετοχής με τις οποίες μετατρέπονται οι τιμές σε βαθμούς συμμετοχής. Οι συναρτήσεις συμμετοχής χρησιμοποιούνται για τον ορισμό ασάφειας (fuzziness), η οποία επηρεάζει τον βαθμό συμμετοχής μιας εισόδου σε ένα λεκτικό όρο. Δηλαδή, για μια είσοδο X , υπολογίζεται ο βαθμός συμμετοχής στα αντίστοιχα ασαφή σύνολα. Υπάρχουν διάφοροι τύποι συναρτήσεων συμμετοχής, όπως Γκαουσιανή, Τριγωνική, Τραπεζοειδής και άλλοι.
- **2^ο Επίπεδο Κανόνων (Rule Layer):** Σε αυτό το επίπεδο υπολογίζεται ο βαθμός ενεργοποίησης του κάθε κανόνα. Ο βαθμός ενεργοποίησης εξαρτάται από τους βαθμούς συμμετοχής που δημιουργήθηκαν στο πρώτο επίπεδο.
- **3^ο Επίπεδο Κανονικοποίησης (Normalization Layer):** Στο επίπεδο αυτό, ο βαθμός ενεργοποίησης κάθε κανόνα κανονικοποιείται σε σχέση με το άθροισμα των βαθμών ενεργοποίησης όλων των κανόνων. Με αυτήν την τεχνική διασφαλίζεται πως τα κανονικοποιημένα βάρη αθροίζουν σε 1, αλλά δεν είναι απαραίτητα ίσα μεταξύ τους.
- **4^ο Επίπεδο Συμπεράσματος (Defuzzification Layer):** Σε αυτό το επίπεδο υπολογίζεται η έξοδος κάθε κανόνα με τις γραμμικές εξισώσεις του μοντέλου TSK. Η τιμή της εξόδου είναι επηρεασμένη από τα βάρη που υπολογίστηκαν στο τρίτο επίπεδο.
- **5^ο Επίπεδο Εξόδου (Output Layer):** Στο τελευταίο επίπεδο, τα επιμέρους αποτελέσματα του τέταρτου επιπέδου συναθροίζονται για την παραγωγή της τελικής αριθμητικής πρόβλεψης του συστήματος.

Πριν την έναρξη εκπαίδευσης του ANFIS προηγείται η αρχικοποίηση. Με τη χρήση της παραδοσιακής μεθόδου Grid Partitioning, το πλήθος των ασαφών κανόνων αυξάνεται εκθετικά σε σχέση με το πλήθος των εισόδων, δημιουργώντας το πρόβλημα «κατάρας της διαστατικότητας». Για τον λόγο αυτό, πλέον στα σύγχρονα συστήματα με πολυδιάστατα δεδομένα είναι περισσότερο συνηθισμένη η χρήση του αλγορίθμου Fuzzy C-Means (FCM). Ο συγκεκριμένος αλγόριθμος με βάση τα δεδομένα εκπαίδευσης, δημιουργεί ομαδοποιήσεις (clusters) των οποίων τα κέντρα μεταφράζονται σε ασαφείς κανόνες. Κάθε cluster αποτελεί έναν ασαφή κανόνα και η αποκέντρωση αντιστοιχεί σε μικρότερο βαθμό συμμετοχής στον συγκεκριμένο κανόνα. Η λειτουργία του FCM είναι παρόμοια με τον πολύ γνωστό αλγόριθμο του τομέα της Μηχανικής Μάθησης, τον K-Means. Στοιχειώδης διαφορά μεταξύ των δύο αποτελεί το γεγονός πως στον K-Means, ένα στοιχείο μπορεί να ανήκει αποκλειστικά μόνο σε ένα cluster, ενώ στον FCM ένα στοιχείο μπορεί να ανήκει ταυτόχρονα σε πολλαπλά clusters αλλά με διαφορετικούς βαθμούς

συμμετοχής. Στην παρούσα εργασία, η δομή των κανόνων προκύπτει μέσω του αλγορίθμου FCM. Ειδικότερα, τα clusters που εξάγονται χρησιμοποιούνται για την αρχικοποίηση της βάσης κανόνων του ANFIS, ενώ στη συνέχεια, οι παράμετροι των κανόνων προσαρμόζονται και βελτιστοποιούνται δυναμικά κατά τη φάση της εκπαίδευσης [14].

Η εκπαίδευση του ANFIS ακολουθεί την προτεινόμενη υβριδική προσέγγιση από τον Jang στο αρχικό άρθρο για το ANFIS [13]. Ο συγκεκριμένος υβριδικός αλγόριθμος εκπαίδευσης (hybrid learning algorithm) συνδυάζει τη μέθοδο Ελαχίστων Τετραγώνων (Least Squares Estimation – LSE) με τη μέθοδο Καθοδικής Κλίσης (Gradient Descent – GD) [15]. Αυτές οι δύο μέθοδοι λειτουργούν σε δύο ξεχωριστές φάσεις, με την πρώτη να είναι το Εμπρόσθιο Πέρασμα (Forward Pass). Σε αυτή τη φάση, το σύστημα διατηρεί σταθερές τις παραμέτρους των συναρτήσεων συμμετοχής του πρώτου επιπέδου (layer). Έπειτα, οι παράμετροι αυτές φτάνουν μέχρι και το τέταρτο επίπεδο, όπου υπολογίζονται και βελτιστοποιούνται μέσω της μεθόδου LSE. Η επόμενη φάση είναι το Οπίσθιο Πέρασμα (Backward Pass). Σε αυτή τη φάση το σύστημα υπολογίζει το σφάλμα της τελικής του πρόβλεψης. Οι παράμετροι των συναρτήσεων συμμετοχής του τέταρτου επιπέδου που υπολογίστηκαν στην προηγούμενη φάση παραμένουν σταθερές και το σύστημα ανάλογα με το σφάλμα προσαρμόζει τις παραμέτρους των συναρτήσεων συμμετοχής του πρώτου επιπέδου [16]. Αυτή η υβριδική προσέγγιση εξασφαλίζει την ταχύτατη εκπαίδευση του συστήματος και μειώνει τον κίνδυνο εγκλωβισμού της εκπαίδευσης σε τοπικά βέλτιστα, το οποίο είναι ένα συχνό πρόβλημα που εμφανίζεται σε αλγόριθμους Μηχανικής Μάθησης. Η διαδικασία της εκπαίδευσης ακολουθεί μια επαναληπτική διαδικασία στην οποία κάθε μία επανάληψη αποκαλείται εποχή.

Η υβριδική αρχιτεκτονική του ANFIS προσφέρει σημαντικά πλεονεκτήματα. Η ικανότητα εκμάθησης και προσαρμογής των Τεχνητών Νευρωνικών Δικτύων σε συνδυασμό με τη λογική αναπαράσταση των Ασαφών Συστημάτων προσφέρει τα πλεονεκτήματα των δύο αυτών πεδίων. Ωστόσο, το σημαντικότερο πλεονέκτημα του ANFIS απέναντι σε άλλα κλασικά μοντέλα μηχανικής μάθησης, αποτελεί η ικανότητα ερμηνευσιμότητας που προσφέρει μέσω των ασαφών κανόνων. Με αυτή την ικανότητα, παρέχει αυξημένη ερμηνευσιμότητα συγκριτικά με μοντέλα «μαύρα κουτιά» και προσφέρει διαφάνεια στον τρόπο λειτουργίας του, το οποίο το κάνει ένα μοντέλο που συχνά χαρακτηρίζεται ως «λευκό κουτί» (white-box model). Το ANFIS δεν παράγει μόνο προβλέψεις, αλλά με τη διαφάνεια των ασαφών κανόνων προσφέρει τη δυνατότητα στα μέλη των ομάδων SOC να εξετάσουν και να επαληθεύσουν τα αποτελέσματα του μοντέλου. Αυτό, επιτρέπει τη δημιουργία εμπιστοσύνης προς το μοντέλο και επαλήθευση των παραγόμενων συναγεμίων επιθέσεων. Βέβαια, υπάρχουν και ορισμένοι περιορισμοί που πρέπει να ληφθούν υπόψιν. Ο σημαντικότερος περιορισμός είναι η υψηλή υπολογιστική ισχύ που απαιτεί όταν εφαρμόζονται πολύπλοκα και πολυδιάστατα δεδομένα (high-dimensional data), δηλαδή όταν εμφανίζεται το πρόβλημα «κατάρτας της διαστατικότητας». Επιπλέον, το ANFIS μπορεί να εμφανίσει κίνδυνο υπερεκπαίδευσης, γεγονός που καθιστά επιτακτική την αυστηρή παραμετροποίηση και αξιολόγηση του μοντέλου.

2.5 Εξελικτικοί Αλγόριθμοι

Η εξέλιξη, σύμφωνα με τον Δαρβίνο, ακολουθεί τη φυσική επιλογή (natural selection) και προκύπτει από τα άτομα ενός είδους τα οποία έχουν τη δυνατότητα να προσαρμόζονται και να επιβιώνουν στο εκάστοτε περιβάλλον. Τα άτομα που δεν παρουσιάζουν προσαρμογή στο περιβάλλον τους, έχουν λίγες πιθανότητες να επιβιώσουν και κατά συνέπεια να παράξουν απογόνους (offspring). Αντίθετα, τα άτομα που έχουν την ικανότητα προσαρμογής, έχουν περισσότερες πιθανότητες να επιβιώσουν και να παράξουν απογόνους. Αυτό σημαίνει πως η εξέλιξη ακολουθεί μια φυσιολογική πορεία στην οποία επιβιώνουν μόνο τα καλύτερα άτομα και η καταλληλότητα αυτών αυξάνεται συνεχώς. Οι Εξελικτικοί

Αλγόριθμοι είναι εμπνευσμένοι από τη συγκεκριμένη θεωρία του Δαρβίνου και τα πρώτα ουσιαστικά βήματα υλοποίησης, πραγματοποιήθηκαν τη δεκαετία του 1960. Επιστήμονες όπως οι Lawrance Fogel, John Holland και Hans-Paul Schwefel έθεσαν τα θεμέλια του τομέα των Εξελικτικών Αλγορίθμων. Έπειτα ο τομέας των Εξελικτικών Αλγορίθμων εξελίχθηκε παρουσιάζοντας διάφορες υποενοότητες και κατηγορίες. Μια από τις διασημότερες κατηγορίες των Εξελικτικών Αλγορίθμων αποτελούν οι Γενετικοί Αλγόριθμοι, οι οποίοι προτάθηκαν από τον John Holland αρχικά το 1975 αλλά καθιερώθηκαν σε επίσημη δημοσίευση το 1992. Από τότε, έχει αποδειχθεί πως οι Γενετικοί Αλγόριθμοι μπορούν να ξεπεράσουν τις κλασικές μη-γενετικές προσεγγίσεις σε πολύπλοκα προβλήματα αναζήτησης βέλτιστης λύσης, ωστόσο το απαιτούμενο, ανάλογα πάντα την περίπτωση, υπολογιστικό κόστος των Γενετικών αλγορίθμων αποτελεί γνωστό μειονέκτημα αυτών [17].

Η βασική ιδέα πίσω από τους Γενετικούς Αλγορίθμους είναι η προσομοίωση του εξελικτικού μηχανισμού της φύσης για την επίλυση πολύπλοκων προβλημάτων βελτιστοποίησης ή/και αναζήτησης στα οποία η εύρεση λύσης είναι περίπλοκη και χρονοβόρα. Η γενική αρχιτεκτονική των Γενετικών Αλγορίθμων στηρίζεται στην ύπαρξη ενός αρχικού «πληθυσμού» (population) από τυχαίες πιθανές λύσεις του προβλήματος. Έπειτα, οι λύσεις αυτές αξιολογούνται μέσω μιας μαθηματικής συνάρτησης καταλληλότητας (fitness function), με στόχο τον ορισμό μιας τιμής που καθορίζει την ποιότητα του ατόμου. Στη συνέχεια, τα άτομα με τις καλύτερες τιμές καταλληλότητας, επιλέγονται ως γονείς για να «αναπαραχθούν». Αυτοί οι γονείς μέσω διαδικασιών, μιμούνται τη βιολογική αρχιτεκτονική αναπαραγωγής, όπου γίνεται ανταλλαγή γονιδίων των γονέων και με τυχαίες γενετικές αλλαγές δημιουργούν νέα άτομα απογόνους. Αυτοί οι απόγονοι, αποτελούν μια νέα γενιά λύσεων του προβλήματος, η οποία αντικαθιστά την παλιά γενιά-πληθυσμό ατόμων. Η συγκεκριμένη διαδικασία πραγματοποιείται επαναληπτικά, μέχρι να βρεθεί ένα άτομο-λύση το οποίο ικανοποιεί τις προϋποθέσεις του εκάστοτε προβλήματος. Η γενική δομή της εξελικτικής διαδικασίας του γενετικού αλγορίθμου παρουσιάζεται παρακάτω με περισσότερες πληροφορίες:

- **Αναπαράσταση Χρωμοσώματος:** Το πρώτο βήμα είναι η μετάφραση του προβλήματος σε μορφή κατανοητή για τον αλγόριθμο. Κάθε πιθανή λύση-άτομο του προβλήματος αναπαρίσταται από ένα Χρωμόσωμα (Chromosome). Το κάθε χρωμόσωμα αποτελείται από δύο θεωρητικά τμήματα, τον Γονότυπο (Genotype) και τον Φαινότυπο (Phenotype). Ο Γονότυπος αποτελεί την εσωτερική και κωδικοποιημένη αναπαράσταση της λύσης, σε μορφή κατανοητή από τον γενετικό αλγόριθμο η οποία αποτελείται συνήθως από ψηφία του δυαδικού συστήματος αρίθμησης τα οποία ονομάζονται γονίδια (genes). Ο Φαινότυπος είναι αυτό που αντιπροσωπεύει το περιεχόμενο του Γονότυπου. Δηλαδή, ο Φαινότυπος αποτελεί την ερμηνεία του Γονότυπου και αυτά τα δύο μαζί αποτελούν ένα Χρωμόσωμα. Το περιεχόμενο του χρωμοσώματος είναι το πλαίσιο στο οποίο πραγματοποιούνται αλλαγές στα επόμενα στάδια, καθώς και η τελική λύση που αναζητείται. Διευκρινίζεται πως οι έννοιες άτομο, λύση και χρωμόσωμα έχουν την ίδια σημασιολογική έννοια.
- **Συνάρτηση Καταλληλότητας και Επιλογή:** Αρχικά πραγματοποιείται η αρχικοποίηση του πληθυσμού χρωμοσωμάτων με την οποία ξεκινά η αξιολόγησή τους μέσω της συνάρτησης καταλληλότητας. Η συνάρτηση καταλληλότητας μπορεί να προσαρμοστεί με τέτοιο τρόπο ώστε να υποδείξει την επιθυμητή συμπεριφορά από τα χρωμοσώματα. Αφού γίνει ο υπολογισμός της καταλληλότητας όλων των χρωμοσωμάτων, ακολουθεί το στάδιο της επιλογής. Ο αλγόριθμος επιλέγει τα χρωμοσώματα-γονείς δίνοντας προτεραιότητα σε αυτά τα οποία παρουσίασαν μεγαλύτερες τιμές καταλληλότητας. Ο στόχος της επιλογής γονέων είναι να επιλεγούν τα καλύτερα χρωμοσώματα τα οποία θα μεταβιβάσουν το γενετικό τους υλικό (γονίδια) στην επόμενη γενιά χρωμοσωμάτων, προσομοιώνοντας τη διαδικασία της βιολογικής κληρονομικότητας. Η διαδικασία επιλογής έχει διάφορους τύπους υλοποίησης, όπως Roulette Wheel Selection, Tournament Selection, Rank Selection, Elitism. Για την υλοποίηση του προτεινόμενου μοντέλου χρησιμοποιήθηκε η μέθοδος επιλογής Tournament Selection. Σε αυτή, επιλέγεται ένα προκαθορισμένο από τον χρήστη X πλήθος τυχαίων χρωμοσωμάτων του

πληθυσμού. Τα χρωμοσώματα αυτά στη συνέχεια με βάση την καταλληλότητά τους, συγκρίνονται μεταξύ τους και επιλέγεται το καλύτερο χρωμόσωμα μέσω μιας διαδικασίας τουρνουά. Αυτή η διαδικασία γίνεται μέχρι να ικανοποιηθεί ο επιθυμητός πληθυσμός γονέων που θα παράξουν νέα χρωμοσώματα για την επόμενη γενιά. Με την Tournament Selection, υπάρχει η δυνατότητα επανατοποθέτησης ενός ήδη επιλεγμένου χρωμοσώματος-γονέα σε ένα νέο τουρνουά, οπότε εκείνα τα χρωμοσώματα που έχουν υψηλή τιμή καταλληλότητας μπορεί να εκλεχθούν ως νικητήρια σε πολλαπλά τουρνουά.

- **Διασταύρωση (Crossover):** Μετά την επιλογή των γονέων, ακολουθεί η διασταύρωση αυτών για την παραγωγή νέων απογόνων. Η διασταύρωση αποτελεί την πρώτη από τις δύο φάσεις παραγωγής απογόνων. Σε αυτή, δύο χρωμοσώματα-γονείς ανταλλάσσουν μεταξύ τους γονίδια με διάφορους τρόπους, με στόχο τη δημιουργία νέων χρωμοσωμάτων-απογόνων. Αυτή η διαδικασία επιτρέπει τον συνδυασμό των καλύτερων γονιδίων κάθε γονέα, που μπορεί να οδηγήσει στην παραγωγή ενός νέου βελτιωμένου χρωμοσώματος.
- **Μετάλλαξη (Mutation):** Η δεύτερη φάση της παραγωγής νέων απογόνων είναι η μετάλλαξη. Σε αυτό το στάδιο, μερικά γονίδια του νεοπαραγόμενου απογόνου λαμβάνουν νέες και τυχαίες τιμές εντός ενός προκαθορισμένου εύρους, με στόχο την περαιτέρω αναζήτηση βελτιωμένων λύσεων-χρωμοσωμάτων. Η μετάλλαξη επιτρέπει την αποφυγή τοπικών βέλτιστων (local optima), στα οποία το θεωρητικό εύρος εξέλιξης έχει φτάσει στο ανώτατο όριο. Έτσι, επιτρέπεται η εξερεύνηση νέων χώρων λύσεων σε αντίθεση με τη διασταύρωση όπου ο απόγονος κληρονομεί γονίδια αποκλειστικά από τους γονείς του.
- **Αντικατάσταση, Επανάληψη και Τερματισμός:** Οι νέοι απόγονοι ενσωματώνονται στον πληθυσμό της επόμενης γενιάς, η οποία αντικαθιστά πλήρως ή μερικώς την προηγούμενη. Έπειτα, η διαδικασία της συνάρτησης καταλληλότητας, επιλογής γονέων και δημιουργίας απογόνων επαναλαμβάνεται διαδοχικά μέχρις ότου ικανοποιηθεί κάποιο κριτήριο τερματισμού. Αυτό το κριτήριο μπορεί να είναι κάποιο κατώτερο όριο τιμής καταλληλότητας με το οποίο κηρύσσεται ένα χρωμόσωμα ως βέλτιστο για το εκάστοτε πρόβλημα ή η συμπλήρωση ενός προκαθορισμένου μέγιστου πλήθους γενεών του γενετικού αλγορίθμου.

Οι γενετικοί αλγόριθμοι, λόγω της αρχιτεκτονικής αλλά και της λειτουργικής διαδικασίας που ακολουθούν, επιτρέπουν τη χρήση τους για την επιλογή χαρακτηριστικών (feature selection) ενός συνόλου δεδομένων. Η επιλογή συγκεκριμένων χαρακτηριστικών από σύνολα δεδομένων σε μορφή πίνακα (tabular data), επιτρέπει τη μείωση της πλεονάζουσας ή μη χρήσιμης πληροφορίας, η οποία έχει ως συνέπεια και τη μείωση της διαστατικότητας των δεδομένων. Η επαναληπτική διαδικασία εύρεσης του βέλτιστου χρωμοσώματος μέσω του γενετικού αλγορίθμου αυτοματοποιεί την εύρεση των χρησιμότερων χαρακτηριστικών που παρέχουν τις κατάλληλες πληροφορίες για τη λύση του εκάστοτε προβλήματος.

2.6 Στατιστική αξιολόγηση μοντέλων

Η αξιολόγηση ενός μοντέλου Μηχανικής Μάθησης δεν μπορεί να βασιστεί αποκλειστικά σε μια μεμονωμένη εκτέλεση, καθώς η φύση των αλγορίθμων μπορεί να οδηγήσει σε παραπλανητικά αποτελέσματα. Η στατιστική αξιολόγηση των μοντέλων καθίσταται επιτακτική, ειδικά για τον τομέα της κυβερνοασφάλειας και των IDS, όπου η αξιοπιστία της απόδοσης των μοντέλων έχει πρωταρχικό ρόλο. Σε αυτούς τους τομείς, είναι εξαιρετικά σημαντικό να αποδεικνύεται πως η σταθερότητα απόδοσης προκύπτει από τη μεθοδολογία του μοντέλου και όχι από κάποιο τυχαίο ευνοϊκό αποτέλεσμα.

Για τη διασφάλιση της σταθερότητας απόδοσης των μοντέλων επιβάλλεται η πραγματοποίηση πολλαπλών ανεξάρτητων εκτελέσεων κάθε μοντέλου. Από αυτές τις πολλαπλές εκτελέσεις προκύπτουν η μέση τιμή (mean) και η τυπική απόκλιση (Standard Deviation – STD) των μετρικών αξιολόγησης, οι οποίες προσφέρουν τη γενική εικόνα σχετικά με τη σταθερότητα και την απόκλιση των αποτελεσμάτων κάθε μοντέλου. Προκειμένου να συγκριθούν πολλαπλά μοντέλα και να ερευνηθεί αν υπάρχει συνολική διαφοροποίηση στις αποδόσεις αυτών, χρησιμοποιείται ο στατιστικός έλεγχος Friedman. Ο

συγκεκριμένος έλεγχος δημιουργήθηκε τη δεκαετία του 1940 και υποδεικνύει μόνο εάν τουλάχιστον ένα από τα μοντέλα διαφέρει στατιστικά από τα υπόλοιπα [18]. Ωστόσο, δεν προσδιορίζει ποιο μοντέλο είναι αυτό που εμφανίζει τη διαφορά. Για την περαιτέρω ανάλυση αυτού του ζητήματος, αν αρχικά ο έλεγχος Friedman έχει επιδείξει πως υπάρχει στατιστική διαφορά, μπορεί να ακολουθήσει η χρήση μεταγενέστερων ελέγχων (post-hoc tests), όπως οι έλεγχοι Nemenyi και Conover. Οι έλεγχοι αυτοί πραγματοποιούν αυστηρές συγκρίσεις ανά ζεύγη, εντοπίζοντας με ακρίβεια μεταξύ ποιων συγκεκριμένων μοντέλων εντοπίζονται οι διαφορές. Όταν και τα στατιστικά αποτελέσματα αυτών των μεταγενέστερων ελέγχων υποδείξουν στατιστικά σημαντική διαφορά μεταξύ δύο μοντέλων, τότε αυτό σημαίνει πως η πιθανότητα της διαφοράς αυτής να οφείλεται σε τυχαίους παράγοντες είναι εξαιρετικά μικρή έως μηδενική. Επομένως, με αυτή τη μεθοδολογία αποδεικνύεται μαθηματικά πως υπάρχει διαφορά μεταξύ των μοντέλων. Ωστόσο, αυτοί οι έλεγχοι δεν υποδεικνύουν ποιο μοντέλο είναι ανώτερο, αλλά μόνο πως υπάρχει στατιστική διαφορά μεταξύ αυτών. Το αν είναι ανώτερο ή όχι ένα μοντέλο είναι αποτέλεσμα της συνεκτίμησης με τη φύση των αποτελεσμάτων και την ερμηνεία τους με βάση το πρόβλημα.

Κεφάλαιο 3ο: Ανασκόπηση σχετικών εργασιών

3.1 Ανίχνευση εισβολών με Μηχανική Μάθηση

Η χρήση των κλασικών τεχνικών Μηχανικής Μάθησης, όπως SVM, Random Forest και Decision Trees, αποτέλεσε την πρώτη προσπάθεια αυτοματοποίησης της ανίχνευσης εισβολών. Έρευνες έχουν δείξει πως τέτοια μοντέλα μπορούν να επιτύχουν εξαιρετικά υψηλά ποσοστά ανίχνευσης εισβολών [19]. Η χρήση του Random Forest, επιτρέπει την ανίχνευση επιθέσεων σε δίκτυα με ποσοστό επιτυχίας πάνω ακόμα και από 99%, τεκμηριώνοντας τις υψηλές αποδόσεις του μοντέλου [20]. Για την περαιτέρω βελτίωση της απόδοσης έχουν προταθεί και συνδυαστικά μοντέλα, στα οποία ενσωματώνονται βελτιστοποιητικοί αλγόριθμοι [21]. Αυτά τα μοντέλα έχουν ισχυρή παρουσία στον τομέα των IDS και συνεχίζουν να προτιμώνται σε περιπτώσεις που ενδείκνυται υψηλό ποσοστό ανίχνευσης εισβολών.

3.2 Ανίχνευση εισβολών με Βαθιά Μάθηση

Η Βαθιά Μάθηση (Deep Learning), αποτελεί έναν τομέα της Μηχανικής Μάθησης. Σε αυτόν, χρησιμοποιούνται πολλαπλά επίπεδα νευρωνικών δικτύων που προσομοιώνουν τον ανθρώπινο τρόπο λήψης αποφάσεων, χωρίς ωστόσο να παρέχεται η αντίστοιχη ερμηνευσιμότητα λειτουργίας του μοντέλου. Η Βαθιά Μάθηση ενδείκνυται για προβλήματα με μεγάλη διαστατικότητα και σύνθετη φύση, όπως ο τομέας των IDS. Συγκεκριμένα, η Βαθιά Μάθηση γίνεται ολοένα και πιο δημοφιλής λύση για τα IDS, καθώς επιτρέπει την αναγνώριση των σύνθετων δεδομένων, όπως αυτά της δικτυακής κίνησης, και υποστηρίζει τόσο επιβλεπόμενες όσο και μη-επιβλεπόμενες μεθόδους μάθησης [22]. Οι συγγραφείς του [23] προτείνουν τη συνδυαστική χρήση Βαθιάς Μάθησης και κλασικών μοντέλων Μηχανικής Μάθησης. Σε αυτό το πλαίσιο συνδυάζεται ο Μη-Συμμετρικός Βαθύς Αυτοκωδικοποιητής (Non-Symmetric Deep Auto-Encoder – NDAE), με το μοντέλο Random Forest. Έτσι, η ικανότητα της Βαθιάς Μάθησης απέναντι σε πολυδιάστατα δεδομένα και η ταχύτητα και ακρίβεια του Random Forest, δημιουργούν ένα υβριδικό μοντέλο για την εφαρμογή του σε ένα Network-based IDS.

3.3 Νευρο-Ασαφή Συστήματα για ανίχνευση εισβολών

Η ανάγκη για ερμηνευσιμότητα και προσαρμογή σε διάφορες επιθέσεις έχει στρέψει το ενδιαφέρον στη χρήση των Νευρο-Ασαφών Συστημάτων και συγκεκριμένα στο ANFIS για την εφαρμογή του σε IDS. Σε αυτό το πλαίσιο, το ANFIS μπορεί να αξιοποιηθεί για την ανίχνευση επιθέσεων, όπως οι DoS, αναλύοντας τα πακέτα σε ασύρματα δίκτυα [24]. Τα αποτελέσματα δείχνουν πως το ANFIS-IDS μοντέλο, μπορεί να μειώσει τον χρόνο καθυστέρησης (delay) εντοπισμού μιας επίθεσης DoS υπό υψηλό φορτίο πακέτων και η εφαρμογή του ANFIS επιτρέπει την προσαρμογή σε διάφορα δίκτυα και διαφορετικά μοτίβα επιθέσεων. Παράλληλα, οι Al-Turaiqi και Al-Mutaiifi παρουσίασαν ένα σύστημα ανίχνευσης ανωμαλιών (Anomaly-based IDS) βασισμένο στο ANFIS, το οποίο, μέσα από εκτενείς πειραματικές αξιολογήσεις, απέδειξε ότι το Νευρο-Ασαφές Σύστημα επιτυγχάνει ιδιαίτερα υψηλά ποσοστά ακρίβειας στον εντοπισμό κακόβουλης δραστηριότητας, αποτελώντας μια ισχυρή εναλλακτική στις παραδοσιακές μεθόδους ταξινόμησης [25]. Ο Uqaili μελέτησε την ενσωμάτωση του ANFIS με αλγορίθμους Μηχανικής και Βαθιάς Μάθησης, καταλήγοντας στο συμπέρασμα ότι τέτοια βελτιστοποιημένα υβριδικά πλαίσια διαχειρίζονται αποδοτικότερα την αβεβαιότητα των σύγχρονων δικτυακών δεδομένων και υπερτερούν στην αναγνώριση πολύπλοκων δικτυακών επιθέσεων [26].

3.4 Εξελικτική βελτιστοποίηση σε Νευρο-Ασαφή Συστήματα

Παρόλο που το ANFIS διαθέτει την αυτοματοποιημένη παραγωγή κανόνων, πρέπει να προκαθοριστούν ορισμένες υπερπαραμέτροι πριν ξεκινήσει η διαδικασία της εκπαίδευσης. Τον προκαθορισμό αυτό, πρέπει να τον κάνει χειροκίνητα ο χρήστης, το οποίο σημαίνει πως η εύρεση των βέλτιστων υπερπαραμέτρων απαιτεί πολλαπλές εκτελέσεις ή ειδικές γνώσεις σχετικά με τα δεδομένα. Για την επίλυση αυτού του θέματος, στη βιβλιογραφία υπάρχει πληθώρα εργασιών που χρησιμοποιούν τους Εξελικτικούς και Γενετικούς Αλγορίθμους για την εύρεση των βέλτιστων υπερπαραμέτρων του ANFIS για τα εκάστοτε πεδία χρήσης. Οι Haznedar και Kalinli προτείνουν την ενσωμάτωση τεχνικών Εξελικτικής Βελτιστοποίησης όπως οι Γενετικοί Αλγόριθμοι, για τη βελτιστοποίηση των εσωτερικών παραμέτρων (υπερπαραμέτρων) του ANFIS, ξεπερνώντας τις δυσκολίες εύρεσης των βέλτιστων τιμών υπερπαραμέτρων [27]. Οι Γενετικοί Αλγόριθμοι προσφέρουν την ικανότητα ρύθμισης των παραμέτρων όπως οι συναρτήσεις συμμετοχής, ο αριθμός των εποχών και παράμετροι σχετικοί με τους κανόνες του συστήματος, με σημαντική μείωση του υπολογιστικού σφάλματος, όπως ανέδειξε το [28], για προβλήματα ταξινόμησης.

Η αποτελεσματικότητα του συνδυασμού ANFIS-GA για την υπερπαραμετροποίηση του ANFIS, επιβεβαιώνεται από την πολυάριθμη εφαρμογή του στη βιβλιογραφία σε πολύπλοκα προβλήματα. Ενδεικτικά, ο συγκεκριμένος συνδυασμός ANFIS-GA αξιοποιήθηκε επιτυχώς από τους Haznedar και Kilinc για την ακριβή πρόβλεψη ρευμάτων ποταμών, χρησιμοποιώντας Γενετικό Αλγόριθμο για την υπερπαραμετροποίηση του ANFIS με στόχο τη βελτιστοποίηση απόδοσής του [29]. Αντίστοιχα, παρόμοιες εφαρμογές του συνδυασμού παρουσιάζονται για τη βελτιστοποίηση και σταθεροποίηση απόδοσης ανεμογεννητριών αλλά και για την πρόβλεψη της μεταβολής τιμών σε αγορές όπως αυτή του χαλκού [30][31].

3.5 Επιλογή χαρακτηριστικών σε Συστήματα Ανίχνευσης Εισβολών

Η επιλογή χαρακτηριστικών επιτρέπει τη μείωση του υπολογιστικού κόστους αλλά και τον καθαρισμό των δεδομένων από αχρείαστες πληροφορίες που προκαλούν «θόρυβο» στην εκπαίδευση των μοντέλων. Στη σχετική βιβλιογραφία, η συνδυαστική χρήση Γενετικών Αλγορίθμων και ANFIS έχει εφαρμοστεί επιτυχώς για την εύρεση των απαραίτητων χαρακτηριστικών, με στόχο την αποδοτικότερη ανίχνευση spam μηνυμάτων ηλεκτρονικού ταχυδρομείου [32]. Επιπλέον, εκτενείς έρευνες σε διάφορα σύνολα δεδομένων του τομέα των IDS αξιολογούν την απόδοση της προσέγγισης ANFIS-GA, επιβεβαιώνοντας τη χρησιμότητα του Γενετικού Αλγορίθμου στην αναγνώριση των βέλτιστων χαρακτηριστικών [33]. Οι παραπάνω βιβλιογραφικές αναφορές, αναφέρουν πως με τη χρήση του Γενετικού Αλγορίθμου για την επιλογή χαρακτηριστικών, μειώθηκε σημαντικά το υπολογιστικό κόστος, ενώ παράλληλα αυξήθηκε σε σημαντικό βαθμό η απόδοση των συστημάτων για την ανίχνευση των αντίστοιχων εισβολών.

Οι συγγραφείς του [34] προτείνουν μια διαφορετική προσέγγιση όπου η επιλογή χαρακτηριστικών δεν γίνεται μέσω γενετικού αλγορίθμου αλλά μέσω του κέρδους πληροφορίας (Information Gain) κάθε χαρακτηριστικού. Ο υπολογισμός σημαντικότητας κάθε χαρακτηριστικού πραγματοποιείται με βάση ορισμένους κανόνες που όρισαν οι συγγραφείς και κατατάσσουν τα χαρακτηριστικά σε τρεις κατηγορίες σημαντικότητας. Μετά τον υπολογισμό, χρησιμοποιούνται μόνο οι δύο από τις τρεις κατηγορίες, οι οποίες περιέχουν τα σημαντικότερα χαρακτηριστικά. Έπειτα, με αυτά τα χαρακτηριστικά εκπαιδεύεται μοντέλο τύπου Decision Trees για την ανίχνευση επιθέσεων στον τομέα των IDS.

Σε άλλες μελέτες προτείνονται ευρύτερες προσεγγίσεις βελτιστοποίησης. Για παράδειγμα, στο [34] ο γενετικός αλγόριθμος αναλαμβάνει διττό ρόλο, εκτελώντας ταυτόχρονα την επιλογή χαρακτηριστικών

και την υπερπαραμετροποίηση των μοντέλων. Από την άλλη πλευρά, στην εργασία των Bakir και Ceviz [36], ακολουθείται μια διαφορετική συνδυαστική προσέγγιση όπου χρησιμοποιείται μια υβριδική μέθοδος για την εύρεση των βέλτιστων χαρακτηριστικών, ενώ ο γενετικός αλγόριθμος εφαρμόζεται αποκλειστικά για τη βελτιστοποίηση των υπερπαραμέτρων των μοντέλων. Αν και αυτές οι προσεγγίσεις ενσωματώνουν γενετικούς αλγόριθμους και τεχνικές επιλογής χαρακτηριστικών, έχοντας δηλαδή αρκετά κοινά στοιχεία με την παρούσα εργασία, παρουσιάζουν σημαντικές διαφορές στον τρόπο υλοποίησης. Κυριότερη διαφορά αποτελεί το γεγονός ότι δεν κάνουν χρήση του ANFIS, με αποτέλεσμα οι αρχιτεκτονικές τους να υστερούν σημαντικά στην ερμηνευσιμότητα και τη διαφάνεια που προσφέρει ένα Νευρο-Ασαφές Σύστημα.

3.6 Ερμηνευσιμότητα στην κυβερνοασφάλεια

Στον τομέα της κυβερνοασφάλειας, η ερμηνευσιμότητα αποτελεί έναν από τους κρισιμότερους παράγοντες για την υιοθέτηση χρήσης του ANFIS. Υπενθυμίζεται πως το ANFIS αποτελεί ένα μοντέλο, το οποίο προσφέρει διαφάνεια του τρόπου λήψης αποφάσεων του μέσω των ασαφών κανόνων. Όπως τονίζουν οι συγγραφείς στο [37], στον τομέα της κυβερνοασφάλειας και του Βιομηχανικού Διαδικτύου των Πραγμάτων (Industrial Internet of Things – IIoT) η ερμηνευσιμότητα έχει σημαντικό ρόλο καθώς προσφέρει διαφάνεια μέσω των λεκτικών κανόνων που διαθέτει το Ασαφές Σύστημα που χρησιμοποιεί, καθιστώντας το ένα αξιόπιστο μοντέλο για τον τομέα του IoT. Αντίστοιχα, η αναγκαιότητα της ερμηνευσιμότητας στον τομέα της κυβερνοασφάλειας υπογραμμίζεται έντονα στη σχετική βιβλιογραφία [38]. Ειδικότερα, η διαφάνεια του μοντέλου ANFIS επιτρέπει την επαλήθευση των ενεργειών του συστήματος. Το χαρακτηριστικό αυτό κρίνεται ζωτικής σημασίας σε επιχειρησιακά περιβάλλοντα που χαρακτηρίζονται από υψηλή αβεβαιότητα (uncertainty) και απαιτούν αυξημένο επίπεδο ελέγχου και προσοχής [38].

3.7 Συζήτηση βιβλιογραφίας και ερευνητικό κενό

Η βιβλιογραφική ανασκόπηση των προηγούμενων ενοτήτων αναδεικνύει την ευρεία εφαρμογή της Μηχανικής Μάθησης στον τομέα της κυβερνοασφάλειας και των IDS. Τα κλασικά μοντέλα Μηχανικής Μάθησης προσφέρουν υψηλά ποσοστά ακρίβειας, αλλά υστερούν σε διαφάνεια λειτουργίας, καθιστώντας τα «μαύρα κουτιά». Σημαντική μείωση αυτού του προβλήματος προσφέρουν τα Νευρο-Ασαφή Συστήματα, τα οποία παρέχουν τη ζητούμενη διαφάνεια. Παρά τα πλεονεκτήματα του ANFIS, η απόδοσή του εξαρτάται από την ποιότητα των δεδομένων εισόδου και τις υπερπαραμέτρους του. Στη βιβλιογραφία, οι Γενετικοί Αλγόριθμοι χρησιμοποιούνται ευρέως για να αντιμετωπίσουν αυτούς τους δύο περιορισμούς. Οι Γενετικοί Αλγόριθμοι αναλαμβάνουν δύο διακριτούς ρόλους. Ο πρώτος είναι η υπερπαραμετροποίηση του ANFIS ή του εκάστοτε μοντέλου και ο δεύτερος είναι η επιλογή χαρακτηριστικών.

Με βάση τις σχετικές εργασίες που εντοπίστηκαν και μελετήθηκαν, παρατηρείται περιορισμένη διερεύνηση όσον αφορά την ταυτόχρονη αντιμετώπιση των προβλημάτων υπερπαραμετροποίησης του ANFIS και της επιλογής χαρακτηριστικών. Η διαθέσιμη βιβλιογραφία στοχεύει κατά κύριο λόγο μεμονωμένα σε ένα από τα δύο προβλήματα. Ακόμη και στις σπάνιες περιπτώσεις όπου γίνεται συνδυαστική προσέγγιση για την αντιμετώπιση και των δύο προβλημάτων, εφαρμόζονται σε διαφορετικές αρχιτεκτονικές και μοντέλα, όπως Decision Trees ή SVM, και όχι σε Νευρο-Ασαφή Συστήματα. Η παρούσα εργασία επιχειρεί να συνεισφέρει στο συγκεκριμένο ερευνητικό πεδίο, βελτιστοποιώντας τις υπερπαραμέτρους του ANFIS ενώ παράλληλα πραγματοποιείται επιλογή χαρακτηριστικών μέσω ενός κοινού πλαισίου Γενετικού Αλγορίθμου. Για την καλύτερη κατανόηση της θέσης της παρούσας εργασίας στο ευρύτερο ερευνητικό πεδίο, ο Πίνακας 3.1 συνοψίζει τα

χαρακτηριστικά επιλεγμένων σχετικών μελετών που συζητήθηκαν. Μέσα από τον συγκριτικό αυτό πίνακα, καθίσταται σαφές ποιοι συνδυασμοί μεθόδων έχουν ήδη εξεταστεί στη βιβλιογραφία και σε ποιους άξονες διαφοροποιείται και πλεονεκτεί η προτεινόμενη μεθοδολογία, συνδυάζοντας την ερμηνευσιμότητα, την εξελικτική βελτιστοποίηση και την αυστηρή στατιστική αξιολόγηση.

Πίνακας 3.1: Συγκριτικός πίνακας σχετικών εργασιών

Εργασία	Μοντέλο	Σύνολο δεδομένων	Επιλογή χαρακτηριστικών	Βελτιστοποίηση υπερπαραμέτρων	Ερμηνευσιμότητα
20	SVM, RF, ELM	NSL-KDD	Όχι	Όχι	Όχι
21	SVM-kNN Ensemble	KDD99	Όχι	Ναι (PSO)	Όχι
23	NDAE-Random Forest	KDD99-NSLKDD	Ναι (unsupervised feature learning)	Όχι	Όχι
24	ANFIS	Ιδιοπαραγόμενα	Όχι	Όχι	Ναι
25	ANFIS-Decision Trees	NSL-KDD	Ναι (Pearson correlation analysis)	Όχι	Ναι
26	ANFIS	NSL-KDD	Ναι	Ναι (CSA)	Ναι
27	ANFIS-GA	Ιδιοπαραγόμενα	Όχι	Ναι (GA)	Ναι
28	MLP, RBF, ANFIS	Πολλαπλά σύνολα	Όχι	Ναι (GA)	Ναι (ANFIS)
29	ANFIS-GA	Πραγματικά δεδομένα	Όχι	Ναι (GA)	Ναι
30	ANFIS-GA	Πραγματικά δεδομένα	Όχι	Ναι (GA)	Ναι
31	ANFIS-GA	Πραγματικά δεδομένα	Όχι	Ναι (GA)	Ναι
32	ANFIS-GA	Προσαρμοσμένα δεδομένα	Ναι (GA)	Όχι	Ναι
33	ANFIS-GA	Πολλαπλά σύνολα	Ναι (GA)	Όχι	Ναι
34	Decision Trees	NSL-KDD	Ναι (Information Gain)	Όχι	Ναι (J48)
35	Decision Trees, RF, ANN	Westermo network traffic	Ναι (GA)	Ναι (GA)	Όχι
36	XGBoost-RF	CICIDS2017	Ναι (Mutual Information-SFS)	Ναι (GA)	Όχι

Κεφάλαιο 4ο: Περιγραφή και προεπεξεργασία δεδομένων

4.1 Ανάλυση του UNSW-NB15

4.1.1 Γενικές πληροφορίες του Συνόλου Δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν στην παρούσα εργασία προέρχονται από το σύνολο δεδομένων UNSW-NB15 [39]-[43]. Το συγκεκριμένο δημιουργήθηκε από τους Moustafa και Slay στη Καμπέρα Αυστραλίας στο εργαστήριο Cyber Range Lab του Αυστραλιανού Κέντρου Κυβερνοασφάλειας (ACCS), το οποίο ανήκει στο Πανεπιστήμιο της Νέας Νότιας Ουαλίας (UNSW). Το συγκεκριμένο επιλέχθηκε καθώς παλαιότερα σύνολα δεδομένων, όπως το KDDCUP 99 και NSLKDD, θεωρούνται πλέον απαρχαιωμένα και ελλιπή, καθώς δεν περιέχουν δεδομένα σύγχρονων κυβερνοεπιθέσεων.

Οι δημιουργοί χρησιμοποίησαν το εργαλείο IXIA PerfectStorm με το οποίο συνέλεξαν τα ακατέργαστα δεδομένα, τα οποία στην πορεία επεξεργάστηκαν και κατέληξαν σε ένα σύνολο 2,54 εκατομμυρίων εγγραφών. Το σύνολο αποτελείται από 49 χαρακτηριστικά (features) με τα 2 τελευταία να αποτελούν τη δυαδική ετικέτα κλάσης και την κατηγορία επίθεσης. Η δυαδική ετικέτα κλάσης λαμβάνει δύο μόνο τιμές, 0 για φυσιολογική δικτυακή κίνηση και 1 για εγγραφή που αντιστοιχεί σε επίθεση. Η κατηγορία επίθεσης απλά περιέχει την ονομασία της επίθεσης, προϋποθέτοντας ότι έχει τιμή 1 η ετικέτα κλάσης. Τα πρώτα 47 χαρακτηριστικά αποτελούνται από διάφορες κατηγορίες πληροφοριών της κίνησης του δικτύου, όπως για παράδειγμα το μέγεθος δεδομένων που μεταφέρθηκαν, αποτυχημένες αποστολές δεδομένων, πρωτόκολλο επικοινωνίας, χρόνο σύνδεσης, πλήθος πακέτων και άλλα.

4.1.2 Κατηγορίες επιθέσεων

Στο σύνολο υπάρχουν συνολικά 9 κατηγορίες επιθέσεων και η φυσιολογική κατάσταση (normal). Οι κατηγορίες, οι οποίες είναι Analysis, Backdoors, DoS, Exploits, Fuzzers, Generic, Reconnaissance, Shellcode και Worms, καλύπτουν ένα μεγάλο εύρος οικογενειών κυβερνοεπιθέσεων. Η κατανομή τους στο σύνολο είναι ανισόρροπη, αντικατοπτρίζοντας πραγματικές συνθήκες δικτύων. Οι σπάνιες επιθέσεις όπως οι Worms έχουν μόνο 174 εμφανίσεις, ενώ αντίθετα οι πιο συχνές επιθέσεις τύπου Generic εμφανίζονται σε πάνω από 200 χιλιάδες εγγραφές. Παρακάτω στον Πίνακα 4.1, παρουσιάζονται οι εγγραφές των 9 τύπων επιθέσεων και της κανονικής κίνησης (Normal).

Πίνακας 4.1: Εγγραφές συνόλου δεδομένων

A/A	Τύπος	Πλήθος εγγραφών	Ποσοστό
1	Normal	2.218.761	87,36 %
2	Fuzzers	24.246	0,95 %
3	Analysis	2.677	0,10 %
4	Backdoors	2.329	0,09 %
5	Dos	16.353	0,65 %
6	Exploits	44.525	1,75 %
7	Generic	215.481	8,48 %
8	Reconnaissance	13.987	0,55 %
9	Shellcode	1.511	0,06 %
10	Worms	174	0,007 %
Σύνολο Επιθέσεων		321.283	12,64 %
Σύνολο		2.540.044	100 %

4.2 Διαχωρισμός δεδομένων

Το αρχικό σύνολο των περίπου 2,54 εκατομμυρίων εγγραφών χωρίστηκε αρχικά σε σύνολο εκπαίδευσης, σύνολο επικύρωσης και τελικό σύνολο ελέγχου, με αναλογία 60%-20%-20% αντίστοιχα. Κατά την εξελικτική βελτιστοποίηση, το σύνολο επικύρωσης διαιρέθηκε περαιτέρω σε εσωτερικό υποσύνολο επικύρωσης, το οποίο χρησιμοποιήθηκε για τον πρόωρο τερματισμό της εκπαίδευσης, και σε υποσύνολο αξιολόγησης καταλληλότητας, το οποίο χρησιμοποιήθηκε για τον υπολογισμό της τιμής καταλληλότητας κάθε υποψήφιας λύσης.

Είναι σημαντικό να σημειωθεί ότι κατά τον διαχωρισμό εφαρμόστηκε στρωματοποιημένη δειγματοληψία (stratified sampling), διασφαλίζοντας ότι η αναλογία των κλάσεων (φυσιολογική δικτυακή κίνηση έναντι επιθέσεων) διατηρήθηκε σταθερή σε κάθε υποσύνολο. Επιπλέον, τα δημιουργηθέντα υποσύνολα παρέμειναν απολύτως σταθερά σε όλα τα πειράματα, και τα ίδια ακριβώς δεδομένα χρησιμοποιήθηκαν για την εκπαίδευση και αξιολόγηση όλων των συγκρινόμενων μεθόδων (τόσο του προτεινόμενου μοντέλου όσο και των μοντέλων αναφοράς), εξασφαλίζοντας την αντικειμενικότητα της συγκριτικής αξιολόγησης. Στον Πίνακα 4.2 παρουσιάζονται συνοπτικά τα ποσοστά που αναλογούν στο κάθε υποσύνολο.

Πίνακας 4.2: Διαχωρισμός δεδομένων σε υποσύνολα

Υποσύνολο		Πλήθος / Ποσοστό	Χρήση
Εκπαίδευσης (Training)		~1.524.000 / 60%	Εκπαίδευση του ANFIS
Επικύρωσης (Validation)	Εσωτερικής επικύρωσης (Internal validation)	~254.000 / 10%	Early stopping
	Αξιολόγησης καταλληλότητας (Fitness evaluation)	~254.000 / 10%	Fitness του εξελικτικού αλγορίθμου
Τελικού ελέγχου (Final Test)		~508.000 / 20%	Τελική αξιολόγηση

4.3 Τεχνικές προεπεξεργασίας δεδομένων

4.3.1 Καθαρισμός και επιλογή χαρακτηριστικών

Τα 49 χαρακτηριστικά παρέχουν πολλά και χρήσιμα δεδομένα, όμως ορισμένα υστερούν σε προγνωστική αξία (predictive value) για την αναγνώριση των επιθέσεων του συνόλου δεδομένων και αν δεν αφαιρεθούν ενδέχεται να προκαλέσουν αθέμιτο θόρυβο και τελικά κατώτερα αποτελέσματα. Αρχικά, αφαιρέθηκαν τα χαρακτηριστικά *srcip* (Source IP) και *dstip* (Destination IP), τα οποία θα δημιουργούσαν παραπλανητικά και μη έγκυρα τελικά αποτελέσματα, όπου για παράδειγμα κατά την εκπαίδευση το μοντέλο απομνημονεύει συγκεκριμένες διευθύνσεις από εγγραφές που αντιστοιχούν σε επιθέσεις. Αυτό θα είχε ως συνέπεια, το τελικό μοντέλο να ταξινομεί κανονικές δικτυακές συνδέσεις λανθασμένα ως επιθέσεις λόγω ταύτισης της διεύθυνσης IP. Ακολούθησε η αφαίρεση των θυρών πηγής και προορισμού (*sport* - Source port, *dsport* - Destination Port) ώστε το μοντέλο να μην εξαρτάται υπερβολικά από συγκεκριμένους αριθμούς θυρών ή υπηρεσίες του συγκεκριμένου συνόλου δεδομένων, γεγονός που θα μπορούσε να περιορίσει τη γενίκευσή του. Επίσης, αφαιρέθηκαν τα χαρακτηριστικά *Stime* (Timestamp of record start time) και *Ltime* (Timestamp of record last time). Τα παραπάνω χαρακτηριστικά δεν είναι χρήσιμα για την εκπαίδευση των μοντέλων, καθώς εξυπηρετούν καθαρά λειτουργικούς και καταγραφικούς σκοπούς, ενώ η λανθασμένη χρήση τους θα οδηγήσει επίσης σε

ελλιπή τελική απόδοση [44]. Επειδή πρόκειται για επιβλεπόμενη μάθηση (Supervised Learning), τα δεδομένα χωρίζονται σε χαρακτηριστικά (Features) και στόχο (Target), οπότε αφαιρέθηκαν και η κατηγορία επίθεσης (*attack_cat*) και η δυαδική ετικέτα κλάσης (Label), αφού όμως έγινε πρώτα η αποθήκευση των συγκεκριμένων δεδομένων σε εξωτερικό αρχείο για τη χρήση τους αργότερα. Εν κατακλείδι, με την αφαίρεση των 6 παραπάνω χαρακτηριστικών, το σύνολο δεδομένων έχει πλέον συνολικά 43 χαρακτηριστικά εισόδου.

4.3.2 One-Hot Encoding

Τα Ασαφή Συστήματα και κατ' επέκταση το ANFIS παρουσιάζουν ένα πολύ σημαντικό μειονέκτημα. Αυτό είναι το γεγονός ότι μπορούν να επεξεργαστούν αποκλειστικά αριθμητικά δεδομένα. Το σύνολο δεδομένων UNSW-NB15 περιλαμβάνει 3 χαρακτηριστικά τα οποία χρησιμοποιούν ονομαστικά δεδομένα (nominal data). Φαινομενικά η τεχνική Label Encoding, δηλαδή η μετατροπή των κατηγορικών/ονομαστικών δεδομένων σε αριθμητικά δεδομένα με σειριακή ταξινόμηση (πχ. TCP=1, UDP=2, ICMP=3), μοιάζει με μια λογική προσέγγιση. Ωστόσο, για την παρούσα εργασία αποτελεί μια λανθασμένη λύση. Η χρήση Label Encoding δημιουργεί για το ANFIS μια λανθασμένη κατάσταση όπου υπάρχει σχέση διάταξης (ordinal relationship) ή κάποιου είδους βαθμωτή αξία. Ουσιαστικά, το ANFIS θα εκπαιδευόταν με την εσφαλμένη παραδοχή πως για παράδειγμα το TCP με τιμή 1 είναι μικρότερο από το UDP με τιμή 2 ενώ παράλληλα η μέση τιμή τους είναι το 1.5, όμως αυτή η συνθήκη δεν έχει κανένα απολύτως πραγματικό νόημα στα πρωτόκολλα επικοινωνίας.

Για να ξεπεραστεί αυτό το πρόβλημα, χρησιμοποιήθηκε η τεχνική One-Hot Encoding. Η συγκεκριμένη τεχνική μετατρέπει τα συγκεκριμένα κατηγορικά δεδομένα σε αριθμητικά, με μορφή που μπορούν τα μοντέλα μηχανικής μάθησης να κατανοήσουν. Για κάθε μία ξεχωριστή τιμή ενός χαρακτηριστικού με ονομαστικά δεδομένα, δημιουργείται ένα νέο χαρακτηριστικό, με κατάλληλη ονομασία, όπου αποδεκτές τιμές είναι το 0 και το 1. Όταν το νέο χαρακτηριστικό έχει τιμή 1 σημαίνει πως η συγκεκριμένη κατηγορία τιμής υπάρχει στη συγκεκριμένη εγγραφή του συνόλου, ενώ για 0 το αντίθετο. Με αυτόν τον τρόπο διαχωρισμού της πληροφορίας των ονομαστικών δεδομένων, εξαλείφεται το πρόβλημα της σχέσης διάταξης και κάθε αρχική τιμή που μετατράπηκε σε ξεχωριστό χαρακτηριστικό πλέον αντιμετωπίζεται ανεξάρτητα. Συγκεκριμένα, τα χαρακτηριστικά τα οποία περιέχουν κατηγορική μορφή ονομαστικών δεδομένων είναι τα *proto* (πρωτόκολλο επικοινωνίας), *service* (υπηρεσία) και *state* (κατάσταση). Μετά την εφαρμογή του One-Hot Encoding, το αποτέλεσμα είναι η αύξηση του αρχικού συνόλου των 43 χαρακτηριστικών εισόδου σε 202.

4.3.3 Περιορισμός ακραίων τιμών και μετασχηματισμός

Τα δεδομένα κίνησης δικτύου γενικά χαρακτηρίζονται από μεγάλη ασυμμετρία (skewness) και μεγάλες διαφορές των ακραίων τιμών (outliers). Για παράδειγμα, μία εγγραφή μπορεί να περιέχει αποστολή μηδενικού όγκου πακέτων, ενώ μια άλλη εγγραφή να φτάνει τα μερικά εκατομμύρια bytes. Για να μειωθεί η επίδραση των ακραίων τιμών στα δεδομένα, εφαρμόστηκε περιορισμός ακραίων τιμών (clipping). Τα όρια ορίστηκαν στο 1ο και 99ο εκατοστημόριο των δεδομένων εκπαίδευσης. Αυτό σημαίνει ότι οποιαδήποτε τιμή βρίσκεται χαμηλότερα από το 1% των υπόλοιπων τιμών, εξισώνεται και αντικαθίσταται με βάση την αμέσως επόμενη τιμή εντός των ορίων. Αντίθετα, όποια τιμή βρίσκεται πάνω από το 99% των υπόλοιπων τιμών, μειώνεται. Έτσι, απομονώνεται ο θόρυβος που μπορεί να προκαλέσουν οι ακραίες τιμές κάθε χαρακτηριστικού. Είναι σημαντικό να αναφερθεί πως τα όρια ρυθμίστηκαν ώστε να μην αλλοιώσουν τα δεδομένα των One-Hot Encoded χαρακτηριστικών. Επιπλέον, για την αποφυγή οποιασδήποτε διαρροής πληροφορίας (data leakage), η μέτρηση του χαμηλότερου 1% και του υψηλότερου 99% έγινε αποκλειστικά στο σύνολο εκπαίδευσης (training set) και στην πορεία

τα ίδια όρια εφαρμόστηκαν αυτούσια στο εσωτερικό σύνολο επικύρωσης, στο σύνολο αξιολόγησης καταλληλότητας και στο τελικό σύνολο ελέγχου.

Μετά τον περιορισμό των ακραίων τιμών, εφαρμόστηκε λογαριθμικός μετασχηματισμός ($\log_{10}$). Ο συγκεκριμένος μετασχηματισμός βοηθά σε δύο σημεία. Το πρώτο σημείο είναι η λύση που παρέχει για το ANFIS, το οποίο δυσκολεύεται με μεγάλες διαφορές τιμών δίνοντας περισσότερη σημασία στις μεγαλύτερες τιμές. Το πρόβλημα αντιμετωπίζεται συμπιέζοντας τις πολύ μεγάλες τιμές και μετατρέποντας τις μεγάλες διαφορές τιμών σε πιο ελεγχόμενες. Για παράδειγμα όταν σε μια εγγραφή του συνόλου δεδομένων υπάρχει μεταφορά 10 byte με τον μετασχηματισμό $\log_{10}(10)$ το αποτέλεσμα θα είναι περίπου 2.4, ενώ όταν υπάρχει μεταφορά 100.000 byte με $\log_{10}(100.000)$ θα έχει αποτέλεσμα ίσο με περίπου 11.5 μονάδες. Αυτή η διαφορά που υπάρχει μεταξύ τους είναι σημαντικά περισσότερο διαχειρίσιμη από το ANFIS. Το δεύτερο σημείο είναι η διατήρηση των μηδενικών του συνόλου δεδομένων ως έχει. Λόγω των μηδενικών τιμών που υπάρχουν στα δεδομένα, η εφαρμογή απλού λογαρίθμου στο μηδέν, δηλαδή $\log(0)$, δίνει αποτέλεσμα που τείνει προς το μείον άπειρο που σημαίνει πως ο κώδικας δεν θα τρέξει λόγω απροσδιοριστίας. Το $\log_{10}$ προσθέτει εσωτερικά το 1, οπότε ακόμα και αν η είσοδος είναι το 0, δηλαδή $\log_{10}(0)$ και άρα $\log_{10}(1)$, το αποτέλεσμα παραμένει 0. Έτσι υπάρχει μία επιπλέον δικλείδα ασφαλείας για την ομαλή και χωρίς σφάλματα προεπεξεργασία των δεδομένων του UNSW-NB15.

4.3.4 Αντιμετώπιση ανισορροπίας κλάσεων με SMOTE

Στο σύνολο δεδομένων UNSW-NB15 υπάρχει μεγάλη ανισορροπία κλάσεων, όπως δείχνει και ο Πίνακας 4.1, όπου η φυσιολογική κίνηση ανέρχεται σε περίπου 87.36% του συνόλου εγγραφών σε σχέση με το σύνολο των κλάσεων επιθέσεων που ανέρχεται στο 12.64%. Επιπλέον, ανισορροπία παρουσιάζεται και από τις κλάσεις επιθέσεων στη μεταξύ τους σύγκριση, καθώς ορισμένες κλάσεις εμφανίζονται δεκάδες χιλιάδες φορές ενώ άλλες εμφανίζονται πολύ σπανιότερα. Η ελάχιστη εκπροσώπηση αυτών των κλάσεων κατά την εκπαίδευση του μοντέλου μπορεί να οδηγήσει σε μη αντιπροσωπευτικά αποτελέσματα. Επίσης, όπως αναφέρει το [45], η παρουσία ανισορροπίας κλάσεων έχει σοβαρές επιπτώσεις στη διαδικασία εκπαίδευσης. Συνήθως οδηγεί στη δημιουργία μοντέλων με μικρή ακρίβεια πρόβλεψης για τις κλάσεις όπου υπάρχει μικρός αριθμός δειγμάτων και τα μοντέλα τείνουν να ταξινομούν τα περισσότερα δείγματα κατά τον τελικό έλεγχο, προς την πλειοψηφική κλάση.

Για την αντιμετώπιση αυτού του προβλήματος, χρησιμοποιήθηκε η τεχνική SMOTE (Synthetic Minority Over-sampling Technique). Η συγκεκριμένη τεχνική προτάθηκε το 2002, ως μια λύση της ανισορροπίας κλάσεων στην εκπαίδευση μοντέλων Μηχανικής Μάθησης [46]. Με την τεχνική SMOTE δημιουργούνται νέες συνθετικές εγγραφές (Synthetic Samples), οι οποίες ακολουθούν το εύρος τιμών που έχουν και τα δεδομένα των αρχικών εγγραφών, χωρίς να γίνεται απλή αντιγραφή. Με αυτόν τον τρόπο αυξάνονται οι εγγραφές των σπάνιων κλάσεων παρέχοντας τη δυνατότητα στα μοντέλα να έχουν αρκετό χώρο εκπαίδευσης για πιο ακριβή τελικά αποτελέσματα. Είναι υψίστης σημασίας να αναφερθεί πως η συγκεκριμένη τεχνική εφαρμόζεται μόνο στο υποσύνολο εκπαίδευσης (Training set). Η παραγωγή συνθετικών δεδομένων στα υποσύνολα επικύρωσης και ελέγχου θα προκαλούσε διαρροή πληροφορίας (data leakage). Έτσι, η χρήση SMOTE μόνο στα δεδομένα εκπαίδευσης, εξασφαλίζει την αντικειμενική αξιολόγηση του μοντέλου με ανέπαφα υποσύνολα δεδομένων που αντικατοπτρίζουν πραγματικές συνθήκες δικτύου.

4.3.5 Κανονικοποίηση δεδομένων

Ο περιορισμός τιμών και ο λογαριθμικός μετασχηματισμός εφαρμόζονται σε κάθε ένα ξεχωριστό χαρακτηριστικό του συνόλου δεδομένων. Όμως το κάθε χαρακτηριστικό έχει διαφορετικό εύρος τιμών, για παράδειγμα ακόμα και μετά την εφαρμογή των παραπάνω τεχνικών το χαρακτηριστικό *sbytes* (Source to Destination Transaction Bytes) μπορεί να έχει εύρος τιμών από 0 έως 20 και ένα άλλο χαρακτηριστικό όπως το *dur* (Record total duration) να έχει εύρος τιμών 0 έως 2. Για το ANFIS αυτή η διαφορά στα εύρη τιμών δεν μπορεί να τεκμηριωθεί λογικά, καθώς το ANFIS είναι ένα καθαρά υπολογιστικό μοντέλο και δεν αντιλαμβάνεται τη φυσική σημασία της διαφοράς των τιμών, επομένως το μοντέλο θα δώσει μεγαλύτερη έμφαση στο χαρακτηριστικό με το μεγαλύτερο εύρος τιμών.

Λύση σε αυτό το πρόβλημα δίνει η τεχνική κανονικοποίησης δεδομένων Z-Score (Standard Score). Ο μαθηματικός τύπος που περιγράφει την τεχνική παρουσιάζεται παρακάτω στη σχέση (4.1):

$$Z = \frac{x - \mu}{\sigma} \quad (4.1)$$

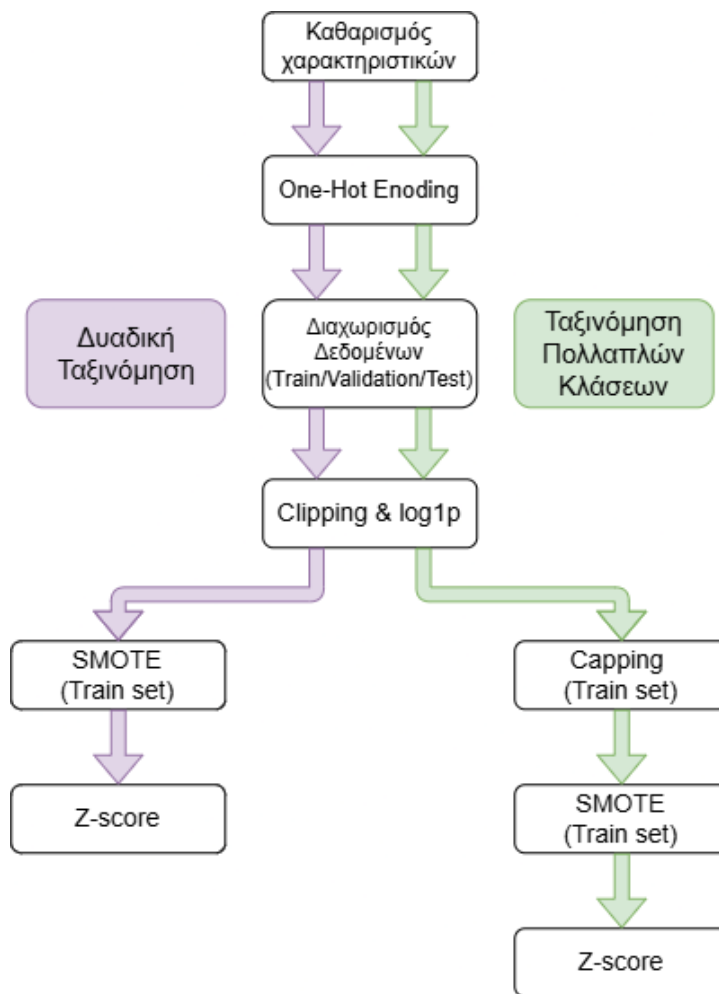
Όπου Z είναι η νέα τυποποιημένη τιμή του χαρακτηριστικού, x η αρχική τιμή πριν τη μετατροπή, μ η μέση τιμή (mean) όλων των δειγμάτων του χαρακτηριστικού και σ η τυπική απόκλιση (Standard Deviation) του χαρακτηριστικού. Με αυτήν την τεχνική πραγματοποιούνται δύο αλλαγές στα δεδομένα. Η πρώτη αλλαγή είναι η κεντροποίηση (centering) των δεδομένων στο 0, δηλαδή ο νέος μέσος όρος της κατανομής μετατρέπεται σε 0. Η δεύτερη αλλαγή είναι η κλιμάκωση της διακύμανσης των τιμών. Έτσι, τα χαρακτηριστικά αποκτούν μια κοινή κλίμακα τιμών που λειτουργεί σωστά για την εκπαίδευση του ANFIS. Για παράδειγμα, μία τιμή του Z-score ίση με το 0 υποδηλώνει ταύτιση με τον μέσο όρο, ενώ μια θετική τιμή ίση με 2 υποδηλώνει πως η τιμή είναι αυξημένη κατά 2 φορές της τυπικής απόκλισης πάνω από τον μέσο όρο. Παράλληλα έγινε περαιτέρω περικοπή των ακραίων τιμών, ως ένα επιπλέον μέτρο προστασίας για το πρόβλημα ανισορροπίας τιμών, στο διάστημα $[-4.0, 4.0]$. Όπως και στην περίπτωση του περιορισμού των ακραίων τιμών, είναι εξαιρετικά σημαντικό να διευκρινιστεί ότι οι παράμετροι της κανονικοποίησης, δηλαδή η μέση τιμή (μ) και η τυπική απόκλιση (σ), υπολογίστηκαν αποκλειστικά από το σύνολο εκπαίδευσης (training set). Στη συνέχεια, αυτές οι ίδιες παράμετροι εφαρμόστηκαν αυτούσιες για την κανονικοποίηση του εσωτερικού συνόλου επικύρωσης, του συνόλου αξιολόγησης καταλληλότητας και του τελικού συνόλου ελέγχου, αποτρέποντας με αυτόν τον τρόπο οποιαδήποτε μελλοντική διαρροή πληροφορίας (data leakage).

4.3.6 Πλαίσιο προεπεξεργασίας δεδομένων για δυαδική ταξινόμηση

Η διαχείριση για την προεπεξεργασία των δεδομένων μετά από ένα σημείο διαχωρίστηκε σε προεπεξεργασία δεδομένων για δυαδική ταξινόμηση (binary classification) και ταξινόμηση πολλαπλών κλάσεων (multiclass classification). Αυτός ο διαχωρισμός έγινε λόγω των διαφορετικών προσεγγίσεων για κάθε μία από δύο ταξινομήσεις που υλοποιήθηκαν.

Αρχικά, όπως αναφέρθηκε στις προηγούμενες υποενότητες, έγινε καθαρισμός των χαρακτηριστικών του συνόλου δεδομένων που δεν είναι χρήσιμα. Ακολούθησε η εφαρμογή της τεχνικής One-Hot Encoding που μετέτρεψε τα κατηγορικά δεδομένα σε νέα χαρακτηριστικά αριθμητικών δεδομένων. Έπειτα, έγινε ο διαχωρισμός του συνόλου δεδομένων σε υποσύνολα για την εκπαίδευση, επικύρωση και τελικό έλεγχο του μοντέλου. Αμέσως επόμενο βήμα αποτέλεσε ο περιορισμός ακραίων τιμών και ο λογαριθμικός μετασχηματισμός των δεδομένων για κάθε ένα ξεχωριστό χαρακτηριστικό. Για την αυστηρή αποφυγή διαρροής δεδομένων (data leakage), τα όρια περικοπής (clipping thresholds) υπολογίστηκαν αποκλειστικά από το υποσύνολο εκπαίδευσης και στη συνέχεια εφαρμόστηκαν στα υποσύνολα επικύρωσης και ελέγχου. Από αυτό το σημείο και έπειτα διαχωρίζεται η προσέγγιση

προεπεξεργασίας για τα δεδομένα των δύο ταξινόμησεων. Ακολουθώντας λοιπόν την προεπεξεργασία για τη δυαδική ταξινόμηση, το επόμενο βήμα είναι η εφαρμογή της τεχνικής SMOTE, η οποία εφαρμόζεται αποκλειστικά στα δεδομένα που προορίζονται για την εκπαίδευση του μοντέλου. Έγινε χρήση του SMOTE με τέτοιο τρόπο ώστε να αυξηθούν οι εγγραφές επιθέσεων και ορίστηκε όριο 800 χιλιάδων συνολικών εγγραφών, το οποίο έχει αναλογία 50%-50% για εγγραφές φυσιολογικής κίνησης και για επιθέσεις αντίστοιχα. Επίσης, μετά την εφαρμογή της συγκεκριμένης τεχνικής πραγματοποιήθηκε ανακάτεμα (shuffling) των εγγραφών για την αποφυγή διαδοχικών συνθετικών εγγραφών που προκαλούνται από το SMOTE. Τελευταίο βήμα είναι η εφαρμογή της τεχνικής Z-score για την τυποποίηση των δεδομένων, η οποία μετατρέπει τα δεδομένα σε μια κοινή κλίμακα για όλα τα 202 χαρακτηριστικά. Η πορεία προεπεξεργασίας που περιεγράφηκε παραπάνω, απεικονίζεται και στο Σχήμα 4.1.



Σχήμα 4.1: Διάγραμμα ροής προεπεξεργασίας δεδομένων

4.3.7 Πλαίσιο προεπεξεργασίας δεδομένων για ταξινόμηση πολλαπλών κλάσεων

Για την προεπεξεργασία των δεδομένων που προορίζονται για ταξινόμηση πολλαπλών κλάσεων, ακολούθησαν τα τέσσερα πρώτα βήματα ακριβώς με την ίδια εφαρμογή που έγινε στην προεπεξεργασία δεδομένων για δυαδική ταξινόμηση. Συγκεκριμένα, πραγματοποιήθηκαν καθαρισμός χαρακτηριστικών, One-Hot Encoding, διαχωρισμός δεδομένων σε υποσύνολα, περιορισμός ακραίων τιμών και χρήση λογαριθμικού μετασχηματισμού. Έπειτα, λόγω της συγκεκριμένης ταξινόμησης,

κρίθηκε αναγκαία η ύπαρξη επαρκούς πλήθους εγγραφών από κάθε κλάση επίθεσης, ειδικά για τις κλάσεις επίθεσης με μικρό αρχικό πλήθος.

Έτσι, έγινε αρχικά η εφαρμογή ενός ανώτερου ορίου (Capping - υποδειγματοληψία) σε κάθε κλάση, από τις 9 των επιθέσεων και τη 1 της φυσιολογικής κίνησης, με τιμή 80.000 εγγραφές. Το όριο αυτό εφαρμόστηκε για να δημιουργηθεί χώρος για τις νέες συνθετικές εγγραφές που θα δημιουργηθούν με SMOTE, έχοντας παράλληλα στόχο το σύνολο δεδομένων εκπαίδευσης να μην ξεπεράσει τις 800.000 εγγραφές (ακολουθώντας δηλαδή το όριο που υπάρχει και στα δεδομένα της δυαδικής ταξινόμησης). Αυτή η προσέγγιση είναι διαφορετική από αυτή της δυαδικής ταξινόμησης, καθώς σε εκείνη εφαρμόζεται SMOTE για τη δημιουργία συνθετικών εγγραφών επιθέσεων ασχέτως της κατηγορίας.

Στη συνέχεια, πραγματοποιείται η διαδικασία SMOTE για το σύνολο εκπαίδευσης, όπου μόνο για τις κλάσεις που έχουν κάτω από 80.000 εγγραφές παράγονται συνθετικές εγγραφές μέχρι το πλήθος τους να φτάσει το όριο αυτό. Με αυτόν τον τρόπο, μετά την ολοκλήρωση του SMOTE το σύνολο εκπαίδευσης αποτελείται από 10 κλάσεις (9 επιθέσεων και 1 φυσιολογικής κίνησης), οι οποίες έχουν από 80.000 εγγραφές η καθεμία, αθροίζοντας τελικά τις 800.000 εγγραφές. Για να γίνει περισσότερο κατανοητός ο τρόπος διαμόρφωσης του συνόλου εκπαίδευσης, ο Πίνακας 4.3 παρουσιάζει αναλυτικά αυτή τη διαδικασία.

Πίνακας 4.3: Εξισορρόπηση δεδομένων εκπαίδευσης για τη ταξινόμηση πολλαπλών κλάσεων

Κλάση	Αρχικό Πλήθος εγγραφών	Πλήθος εγγραφών μετά το Capping	Συνθετικές εγγραφές SMOTE	Τελικό Πλήθος εγγραφών
Normal	1.331.257	80.000	0	80.000
Fuzzers	14.548	14.548	65.452	80.000
Analysis	1.606	1.606	78.394	80.000
Backdoors	1.397	1.397	78.603	80.000
Dos	9.812	9.812	70.188	80.000
Exploits	26.715	26.715	53.285	80.000
Generic	129.289	80.000	0	80.000
Reconnaissance	8.392	8.392	71.608	80.000
Shellcode	907	907	79.093	80.000
Worms	104	104	79.896	80.000

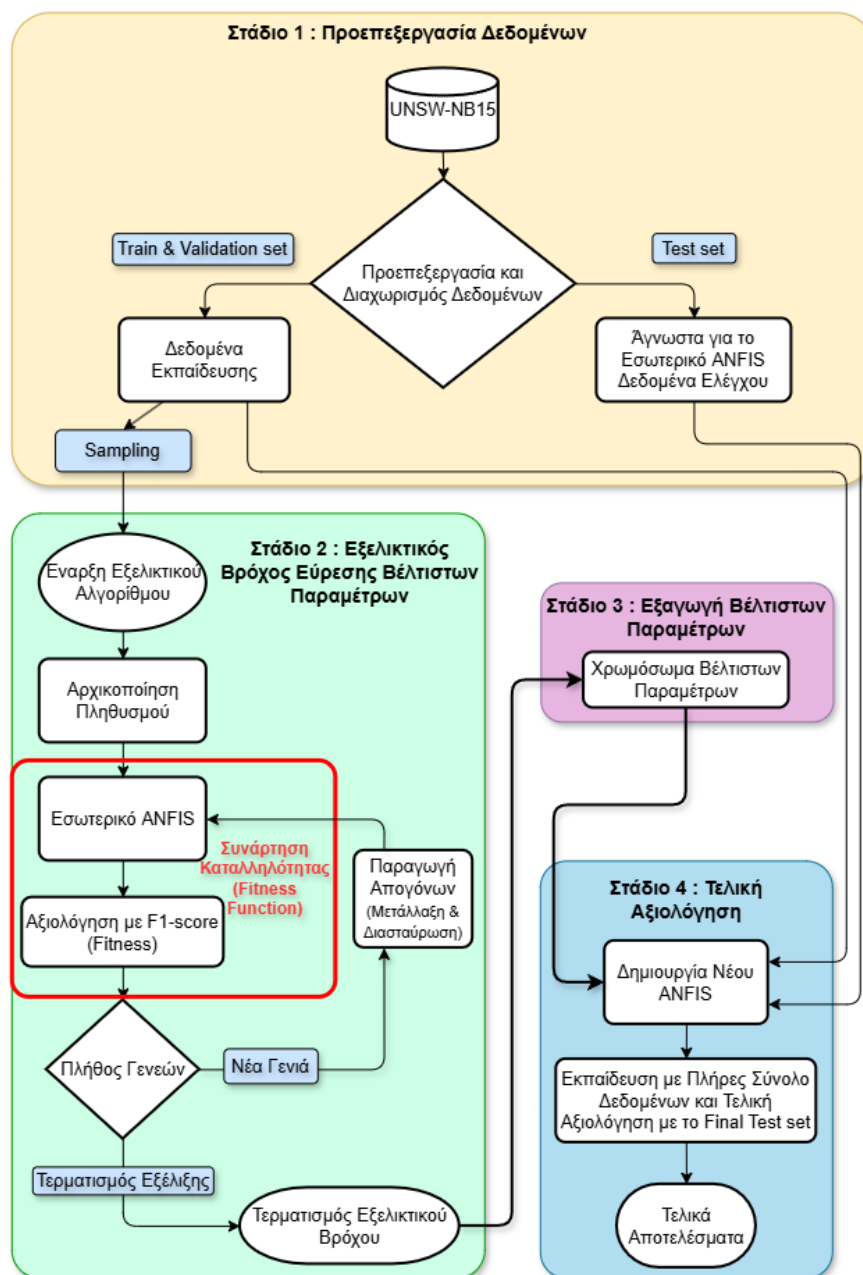
Η παρουσίαση αυτών των αριθμητικών στοιχείων αναδεικνύει μια κρίσιμη πτυχή της διαδικασίας εξισορρόπησης δεδομένων. Πολύ σπάνιες κλάσεις (όπως τα Worms και τα Shellcode) ενισχύθηκαν με έναν εξαιρετικά μεγάλο αριθμό συνθετικών εγγραφών, αυξάνοντας το πλήθος τους από ελάχιστες πραγματικές εγγραφές σε 80.000. Αν και αυτή η δραστική αύξηση προσφέρει στο ANFIS αρκετά δεδομένα για την αναγνώριση μοτίβων, η συντριπτική υπεροχή των συνθετικών δεδομένων έναντι των πραγματικών δημιουργεί τον κίνδυνο της υπερπροσαρμογής (overfitting) σε συνθετικά μοτίβα, γεγονός που μπορεί να οδηγήσει σε χαμηλή τελική ακρίβεια (precision).

Τέλος, εφαρμόστηκε και η τεχνική Z-score όπως και στην προεπεξεργασία για δυαδική ταξινόμηση. Στο Σχήμα 4.1 απεικονίζεται η συνολική ροή προεπεξεργασίας δεδομένων για την ταξινόμηση πολλαπλών κλάσεων.

Κεφάλαιο 5ο: Μεθοδολογία και υλοποίηση

5.1 Γενική αρχιτεκτονική συστήματος

Η αρχιτεκτονική του προτεινόμενου συστήματος ανίχνευσης εισβολών σε δίκτυα βασίζεται στην υβριδική συνεργασία μεταξύ ANFIS και Γενετικού Αλγορίθμου. Ο κύριος στόχος αυτής της συνεργασίας είναι η αυτοματοποιημένη βελτιστοποίηση των παραμέτρων του μοντέλου. Απώτερος σκοπός είναι η μεγιστοποίηση απόδοσης ταξινόμησης από το μοντέλο. Η αρχιτεκτονική για το κομμάτι του γενετικού αλγορίθμου είναι τύπου Wrapper, όπου μέσα στον γενετικό αλγόριθμο λειτουργεί ένα εσωτερικό ANFIS με τα αποτελέσματα του οποίου επηρεάζεται η πορεία της εξέλιξης. Η λειτουργία του μοντέλου απαρτίζεται από τέσσερα βασικά στάδια. Στο Σχήμα 5.1 απεικονίζεται το διάγραμμα ροής της γενικής λειτουργίας του προτεινόμενου μοντέλου.



Σχήμα 5.1: Διάγραμμα ροής της αρχιτεκτονικής του προτεινόμενου συστήματος

Το πρώτο στάδιο, αποτελείται από την προεπεξεργασία των δεδομένων. Όπως αναλύθηκε εκτενώς το κομμάτι της προεπεξεργασίας των δεδομένων στο Κεφάλαιο 4, τα πρωτογενή δεδομένα του συνόλου UNSW-NB15 υποβάλλονται σε μια αυστηρή διαδικασία επεξεργασίας ώστε να είναι έτοιμα για να χρησιμοποιηθούν στην εκπαίδευση, επικύρωση και τελικό έλεγχο του ANFIS.

Στο δεύτερο στάδιο, ο εξελικτικός βρόχος με την εσωτερική εφαρμογή του ANFIS λειτουργεί για την εύρεση των βέλτιστων παραμέτρων (επιλογή χαρακτηριστικών και υπερπαραμέτροι ANFIS) που θα εφαρμοστούν στον τελικό έλεγχο. Ένα χρωμόσωμα αντιπροσωπεύει τις τιμές των παραμέτρων που εφαρμόζονται στο μοντέλο. Για κάθε υποψήφιο χρωμόσωμα κάθε γενιάς, δημιουργείται ένα νέο προσωρινό μοντέλο ANFIS. Χρησιμοποιήθηκε η τεχνική FGA (Fast Genetic Algorithm), με την οποία το εσωτερικό μοντέλο αξιολόγησης λειτουργεί ως μια ποιοτική προσέγγιση (qualitative approximation) του αρχικού πλαισίου. Για να επιτευχθεί αυτό, το μοντέλο χρησιμοποιεί στοχευμένα, αντιπροσωπευτικά δείγματα εγγραφών για κάθε υποσύνολο δεδομένων που προορίζεται για την εκπαίδευση, επικύρωση και έλεγχο, με στόχο την επιτάχυνση της εξελικτικής διαδικασίας. Το εσωτερικό ANFIS, που υλοποιείται μέσα στη συνάρτηση καταλληλότητας (Fitness Function) του γενετικού αλγορίθμου, για την επικύρωση και έλεγχό του, χρησιμοποιείται μόνο το υποσύνολο επικύρωσης (Validation set) το οποίο χωρίζεται σε δύο νέα υποσύνολα, το εσωτερικής επικύρωσης (Internal validation set) και το αξιολόγησης καταλληλότητας (Fitness evaluation set) αντίστοιχα, τα οποία παρουσιάζονται και στον Πίνακα 4.2. Έτσι, αφήνοντας το υποσύνολο τελικού ελέγχου (Final Test set) εκτός του εξελικτικού βρόχου δεν θα υπάρξει διαρροή δεδομένων στην τελική αξιολόγηση του τέταρτου σταδίου. Το εσωτερικό ANFIS αξιολογείται με βάση την τιμή του F1-score, αποτελώντας την τιμή της μεταβλητής καταλληλότητας (Fitness) του συγκεκριμένου χρωμοσώματος. Στη συνέχεια, ανάλογα με την καταλληλότητα του χρωμοσώματος, ο γενετικός αλγόριθμος διαλέγει τα χρωμοσώματα που θα ακολουθήσουν στην επόμενη γενιά αναπροσαρμόζοντάς τα μέσω μετάλλαξης (Mutation) και ανασυνδυασμού (Crossover). Έτσι κάθε επόμενη γενιά αποτελείται από όλο και περισσότερα χρωμοσώματα που αποδίδουν καλύτερα αποτελέσματα. Ουσιαστικά, σε αυτό το στάδιο ο γενετικός αλγόριθμος λειτουργεί ως μια αυτοματοποιημένη εύρεση των βέλτιστων παραμέτρων, χρησιμοποιώντας ένα ελαφρύ εσωτερικό μοντέλο και δεν αποτελεί την τελική αξιολόγηση του συστήματος.

Το τρίτο στάδιο αποτελείται από την εξαγωγή των βέλτιστων παραμέτρων. Με την ολοκλήρωση των προγραμματισμένων γενεών, η εξέλιξη που περιγράφεται στο δεύτερο στάδιο τερματίζεται ολοκληρώνοντας τον ρόλο του γενετικού αλγορίθμου στη λειτουργία του προτεινόμενου μοντέλου. Έπειτα, εξάγεται το χρωμόσωμα το οποίο είχε τη μεγαλύτερη τιμή της συνάρτησης καταλληλότητας.

Στο τέταρτο και τελικό στάδιο, υλοποιείται η φάση της τελικής αξιολόγησης. Οι τιμές των γονιδίων του βέλτιστου χρωμοσώματος χρησιμοποιούνται ως παράμετροι σε ένα μοντέλο ANFIS το οποίο χρησιμοποιεί όλο το σύνολο των προεπεξεργασμένων δεδομένων, σε αντίθεση με την προσεγγιστική υλοποίηση του εσωτερικού ANFIS του γενετικού αλγορίθμου. Δημιουργείται εξαρχής ένα νέο ανεξάρτητο σύστημα ANFIS στο οποίο εισάγονται οι παράμετροι από τα γονίδια του βέλτιστου χρωμοσώματος του 3^{ου} σταδίου. Έπειτα ακολουθεί η εκπαίδευση του μοντέλου και η τελική αξιολόγηση ταξινόμησης με το υποσύνολο δεδομένων τελικού ελέγχου που εισάγεται σε ANFIS για πρώτη φορά, καθώς δεν χρησιμοποιήθηκε στο εσωτερικό ANFIS του δεύτερου σταδίου όπως αναφέρθηκε για την αποφυγή διαρροής δεδομένων.

5.2 Υλοποίηση του ANFIS

Ο κύριος πυρήνας του προτεινόμενου υβριδικού μοντέλου βασίζεται σε ένα Προσαρμοστικό Νευρο-Ασφές Σύστημα Συμπερασμού (ANFIS). Για την υλοποίηση, χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python στην 3.12.10 έκδοση μαζί με τη βιβλιοθήκη X-ANFIS που παρέχει όλη την υποδομή για τη λειτουργία του ANFIS [47][48]. Επιλέχθηκε να γίνει η χρήση του κλασικού μοντέλου ANFIS (Classic ANFIS), και πιο συγκεκριμένα της κλάσης `AnfisClassifier`, καθώς η βιβλιοθήκη παρέχει διάφορες προσεγγίσεις υλοποίησης του ANFIS. Για την εκπαίδευση χρησιμοποιείται η λογική του υβριδικού αλγορίθμου που πρότεινε ο Jang το 1993 (συνδυασμός Gradient Descent και Least Squares Estimation). Ωστόσο, διευκρινίζεται ότι στη συγκεκριμένη υλοποίηση μέσω της βιβλιοθήκης X-ANFIS, ο βελτιστοποιητής Adam (Adaptive Moment Estimation) αντικαθιστά την παραδοσιακή μέθοδο Gradient Descent για την ενημέρωση παραμέτρων των συναρτήσεων συμμετοχής του μοντέλου. Ο Adam συνδυάζει τα πλεονεκτήματα των μεθόδων Momentum και RMSProp, προσαρμόζοντας δυναμικά τον ρυθμό εκμάθησης. Η χρήση του Adam, σε συνδυασμό με τη μέθοδο Ελαχίστων Τετραγώνων (Least Squares), μειώνει τον κίνδυνο εγκλωβισμού της εκπαίδευσης σε τοπικά βέλτιστα και εξασφαλίζει ταχύτερη και πιο ομαλή σύγκλιση. Λόγω του μεγάλου όγκου δεδομένων, η διαδικασία εκπαίδευσης υλοποιείται σε παρτίδες (batches) μικρότερων συνόλων ώστε να διασφαλιστεί η καλύτερη διαχείριση της διαθέσιμης μνήμης RAM, διατηρώντας όμως την ίδια κατεύθυνση εκπαίδευσης.

Είναι σημαντικό να αναφερθεί πως έγινε τροποποίηση του πηγαίου κώδικα της βιβλιοθήκης αυτής καθώς δεν εξυπηρετούσε ορισμένες καθοριστικές απαιτήσεις για τη χρήση που προορίζεται. Συγκεκριμένα, η πρώτη τροποποίηση του κώδικα αφορά την ικανότητα του ANFIS να δέχεται τα κέντρα που ορίζονται από την αρχικοποίηση με Fuzzy C-Means. Η δεύτερη τροποποίηση αφορά τη δυνατότητα χρήσης ξεχωριστού υποσυνόλου επικύρωσης, καθώς ο κώδικας αρχικά δεχόταν μόνο υποσύνολο εκπαίδευσης και ελέγχου και για επικύρωση χρησιμοποιούσε ένα ποσοστό του υποσυνόλου εκπαίδευσης.

Για την αρχικοποίηση εφαρμόστηκε η τεχνική Fuzzy C-Means, η οποία υλοποιήθηκε με τη συνάρτηση `fuzzy.cluster.cmeans` της βιβλιοθήκης `scikit-learn` [49]. Η συνάρτηση δέχεται ένα δείγμα (ίσο με το 10% του συνόλου εκπαίδευσης) με σκοπό τη δημιουργία κέντρων-ομάδων (clusters), τα οποία χρησιμοποιούνται για την αρχικοποίηση της βάσης κανόνων του ANFIS. Το πλήθος των κέντρων αντικατοπτρίζει τους ασαφείς κανόνες του συστήματος, δηλαδή το πλήθος των clusters είναι ίσο με το πλήθος των κανόνων. Τα κέντρα αυτά αποτελούν ένα αρχικό σημείο αναφοράς, ενώ στη συνέχεια οι παράμετροι των κανόνων προσαρμόζονται και βελτιστοποιούνται περαιτέρω κατά τη διάρκεια της εκπαίδευσης του μοντέλου. Για τη διαδικασία της ασαφοποίησης (fuzzification) των δεδομένων εισόδου, δύο διαφορετικές συναρτήσεις συμμετοχής είναι διαθέσιμες για χρήση, οι Τριγωνικές (Triangular) και οι Γκαουσιανές (Gaussian), παρέχοντας ευελιξία στο μοντέλο για καλύτερη απόδοση. Για την αποφυγή του φαινομένου της υπερπροσαρμογής (overfitting), χρησιμοποιείται ο μηχανισμός Πρόωρου Τερματισμού (Early Stopping). Ο συγκεκριμένος μηχανισμός παρακολουθεί το σφάλμα του εσωτερικού συνόλου επικύρωσης (Internal Validation set) και τερματίζει αυτόματα την εκπαίδευση του μοντέλου όταν δεν υπάρχει βελτίωση του σφάλματος για ένα προκαθορισμένο πλήθος εποχών (patience). Για τη συγκεκριμένη υλοποίηση, το patience ορίστηκε στη μισή τιμή των εκάστοτε ορισμένων εποχών εκπαίδευσης.

Η έξοδος του συστήματος είναι γραμμική (Linear) όπως και η βασική αρχιτεκτονική του ANFIS το οποίο βασίζεται στο Takagi-Sugeno-Kang. Για τον λόγο αυτό, έγινε ορισμός ώστε να μην χρησιμοποιείται κάποια συνάρτηση ενεργοποίησης στο τελικό επίπεδο (`act_output=None`). Η γραμμική αυτή έξοδος επεξεργάζεται περαιτέρω με τη βοήθεια της βιβλιοθήκης `scikit-learn` για τον υπολογισμό

διάφορων μετρικών, όπως το F1-score, οι οποίες επιτρέπουν τη μετάφραση της τελικής απόδοσης ταξινόμησης του μοντέλου. Τέλος, πέρα από την εμφάνιση των μετρικών, αξιοποιείται ένα από τα σημαντικότερα χαρακτηριστικά του μοντέλου ANFIS, το οποίο είναι η εξαγωγή των κανόνων του μοντέλου σε ερμηνεύσιμη μορφή. Οι κανόνες τύπου AN-TOTE (IF-THEN) δίνουν τη δυνατότητα κατανόησης της λογικής πορείας που ακολουθεί το μοντέλο ANFIS για τη λήψη αποφάσεων. Το σημαντικό αυτό πλεονέκτημα μειώνει το φαινόμενο «μαύρου κουτιού» που χαρακτηρίζει άλλα μοντέλα Μηχανικής Μάθησης, προσφέροντας διαφάνεια στη λειτουργία του συστήματος.

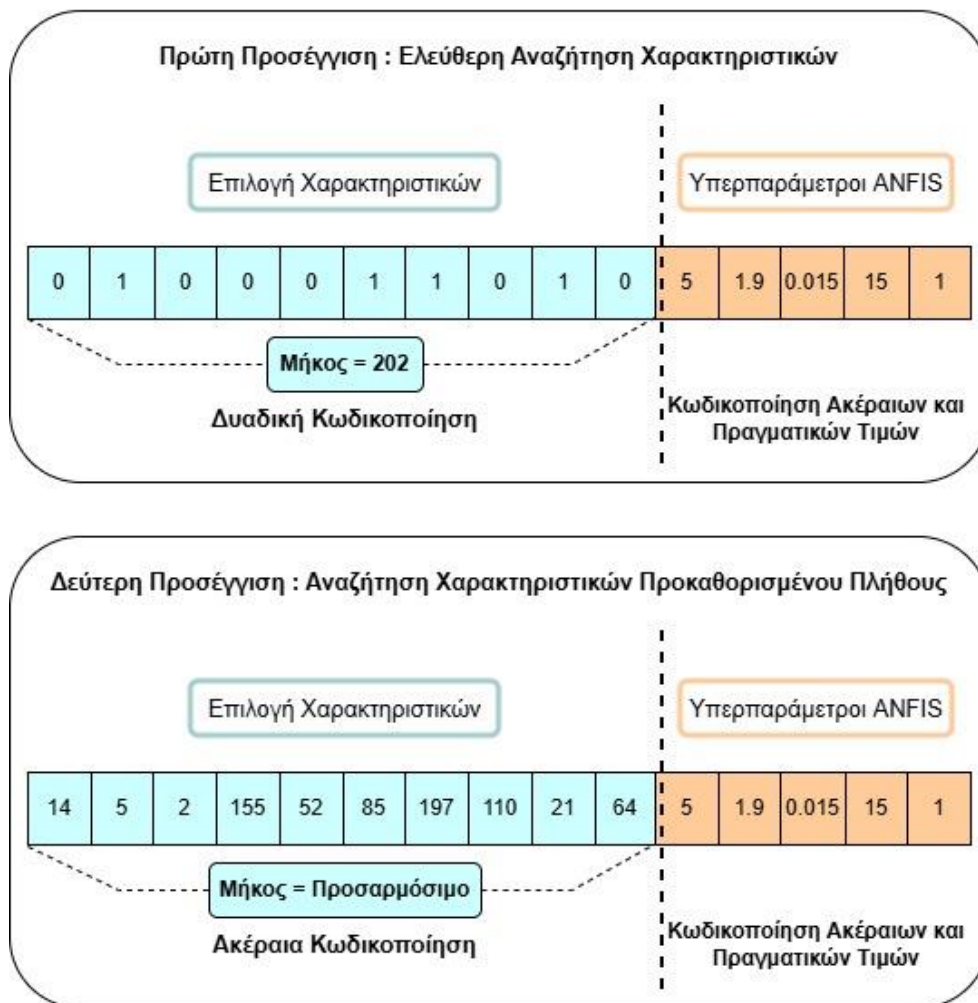
5.3 Υλοποίηση του Γενετικού Αλγορίθμου

Η υλοποίηση του εξελικτικού και πιο συγκεκριμένα του γενετικού αλγορίθμου έγινε με τη βοήθεια της βιβλιοθήκης DEAP [50]. Η συγκεκριμένη βιβλιοθήκη παρέχει τη δυνατότητα στον γενετικό αλγόριθμο να λειτουργεί σύμφωνα με την προτεινόμενη αρχιτεκτονική που αναλύθηκε σε προηγούμενη ενότητα. Η λειτουργία του γενετικού αλγορίθμου στο προτεινόμενο σύστημα στοχεύει στην εύρεση δύο διαφορετικών κατηγοριών βέλτιστων παραμέτρων. Η πρώτη κατηγορία αφορά την επιλογή χαρακτηριστικών (Feature Selection) από το σύνολο δεδομένων. Με την επιλογή χαρακτηριστικών μειώνεται ο χώρος αναζήτησης, το οποίο είναι αρκετά σημαντικό για το ANFIS το οποίο αντιμετωπίζει δυσκολίες σε δεδομένα μεγάλης διάστασης (High-Dimensional Data) και της «κατάρας της διαστατικότητας» (Curse of Dimensionality). Συγκεκριμένα, καθώς αυξάνεται ο αριθμός των εισόδων του συστήματος, δημιουργείται πολλαπλάσιο υπολογιστικό κόστος κατά την εκπαίδευση. Η δεύτερη κατηγορία αφορά την υπερπαραμετροποίηση του ANFIS για παραμέτρους όπως το πλήθος εποχών, το πλήθος ασαφών κανόνων του συστήματος και άλλα.

Η αναπαράσταση των χρωμοσωμάτων υλοποιήθηκε με δύο προσεγγίσεις, η πρώτη αφορά την ελεύθερου πλήθους επιλογή χαρακτηριστικών με παράλληλη υπερπαραμετροποίηση του ANFIS και η δεύτερη την προκαθορισμένου πλήθους επιλογή χαρακτηριστικών με επίσης παράλληλη υπερπαραμετροποίηση ANFIS. Για την πρώτη προσέγγιση, το χρωμόσωμα δεν μπορεί να αναπαρασταθεί αποκλειστικά από έναν τύπο κωδικοποίησης (Encoding) για να εξυπηρετήσει τις δύο κατηγορίες παραμέτρων. Για τον λόγο αυτό, χρησιμοποιήθηκε υβριδική αναπαράσταση χρωμοσώματος (Hybrid Chromosome Representation) με την οποία μπορούν να συνδυαστούν δύο ή περισσότεροι τύποι κωδικοποίησης μέσα σε ένα χρωμόσωμα. Το πρώτο τμήμα του χρωμοσώματος της πρώτης προσέγγισης λειτουργεί ως ένα δυαδικό χρωμόσωμα (Binary Chromosome) το οποίο αφορά την επιλογή χαρακτηριστικών, όπου τα δυαδικά κωδικοποιημένα γονίδια (Genes) λαμβάνουν τιμή 1 για την επιλογή του χαρακτηριστικού και αντίθετα τιμή 0 για τη μη επιλογή του χαρακτηριστικού. Το δεύτερο τμήμα του χρωμοσώματος λειτουργεί ως ένα χρωμόσωμα πραγματικών και ακέραιων τιμών (Real/Integer-Valued Chromosome) στο οποίο κάθε γονίδιο λαμβάνει τιμές που αντιστοιχούν σε ξεχωριστές υπερπαραμέτρους για το ANFIS. Έτσι, η πρώτη προσέγγιση συνδυάζει ένα δυαδικό χρωμόσωμα 202 γονιδίων με ένα χρωμόσωμα πραγματικών και ακέραιων τιμών 5 γονιδίων, δημιουργώντας ένα ενιαίο υβριδικό χρωμόσωμα 207 γονιδίων. Αντίθετα η δεύτερη προσέγγιση, όπου το πλήθος επιλεγμένων χαρακτηριστικών είναι προκαθορισμένο, χρησιμοποιεί ένα ενιαίο χρωμόσωμα πραγματικών και ακέραιων τιμών το οποίο χωρίζεται επίσης σε δύο θεωρητικά τμήματα. Το πρώτο τμήμα αυτής της προσέγγισης αποτελείται από ένα προκαθορισμένο πλήθος γονιδίων, των οποίων οι τιμές αντιστοιχίζονται με τον αριθμό της σειριακής αρίθμησης των χαρακτηριστικών από το σύνολο δεδομένων. Το δεύτερο τμήμα ακολουθεί την ίδια υλοποίηση όπως και στην πρώτη προσέγγιση, αναλαμβάνοντας την αντιστοίχιση των τελευταίων 5 γονιδίων με την υπερπαραμετροποίηση του ANFIS. Ουσιαστικά, η διαφορά των δύο προσεγγίσεων βρίσκεται στην υλοποίηση του πρώτου τμήματος του χρωμοσώματος, με την πρώτη προσέγγιση να ακολουθεί επιλογή χαρακτηριστικών χωρίς

όριο πλήθους και στη δεύτερη να χρησιμοποιεί ένα αρχικό προκαθορισμένο πλήθος χαρακτηριστικών με βάση τις απαιτήσεις του πειράματος, δημιουργώντας τη δυνατότητα προσαρμογής του επιθυμητού πλήθους και πολυπλοκότητας. Η διαφοροποίηση των δύο προσεγγίσεων έγινε καθαρά για ερευνητικούς λόγους, ώστε να πραγματοποιηθεί μια μελέτη ανάμεσα στην απόδοση του γενετικού αλγορίθμου σε έναν ελεύθερο και σε έναν προκαθορισμένο χώρο αναζήτησης, καθώς η επίδραση του πλήθους χαρακτηριστικών τόσο στην τελική απόδοση ταξινόμησης όσο και στο υπολογιστικό κόστος του συστήματος είναι σημαντική. Στόχος είναι να διαπιστωθεί ποια από τις δύο προσεγγίσεις αποδίδει καλύτερα αποτελέσματα. Η αναπαράσταση των χρωμοσωμάτων απεικονίζεται στο Σχήμα 5.2 για την καλύτερη κατανόησή τους.

Υβριδικό Χρωμόσωμα



Σχήμα 5.2: Αναπαράσταση των δύο προσεγγίσεων υβριδικού χρωμοσώματος για το Σενάριο 2

Η συνάρτηση καταλληλότητας αποτελεί τον πυρήνα της εξελικτικής διαδικασίας, καθώς αναλαμβάνει την αξιολόγηση απόδοσης κάθε χρωμοσώματος. Η υλοποίηση του ANFIS μπορεί να δημιουργήσει αρκετά μεγάλο υπολογιστικό κόστος το οποίο έχει ως αποτέλεσμα μεγάλες χρονικές περιόδους για την ολοκλήρωση της εκπαίδευσής του. Για τη συγκεκριμένη λειτουργία του ANFIS εσωτερικά της συνάρτησης καταλληλότητας, αυτό σημαίνει πολλαπλασιασμός του υπολογιστικού κόστους και της χρονικής διάρκειας που απαιτείται για την ολοκλήρωση της εξελικτικής διαδικασίας. Για την αποφυγή αυτού του φαινομένου, έγινε χρήση της αρχιτεκτονικής Fast Genetic Algorithm, όπως είχε αναφερθεί

σε προηγούμενη ενότητα. Με τη συγκεκριμένη αρχιτεκτονική, το εσωτερικό μοντέλο αξιολόγησης, που στη συγκεκριμένη περίπτωση είναι το ANFIS που βρίσκεται μέσα στη συνάρτηση αξιολόγησης, χρησιμοποιεί ένα μικρότερο υποσύνολο δεδομένων για τη γρηγορότερη εκπαίδευση και αξιολόγησή του. Η ίδια αρχιτεκτονική χρησιμοποιείται και στο [17], όπου ο γενετικός αλγόριθμος έχει ρόλο Wrapper για την εύρεση των βέλτιστων χαρακτηριστικών με ένα «ελαφρύ» εσωτερικό μοντέλο, χρησιμοποιώντας ποιοτική προσέγγιση με τη δημιουργία ενός μικρότερου υποσυνόλου δεδομένων εκπαίδευσης. Στο προτεινόμενο μοντέλο, αξιοποιούνται αρχικά τα δύο από τα τρία κύρια σύνολα δεδομένων: το σύνολο εκπαίδευσης και το σύνολο επικύρωσης. Το σύνολο εκπαίδευσης χρησιμοποιήθηκε αποκλειστικά για τη δημιουργία νέου μικρότερου δείγματος εκπαίδευσης. Αντίθετα, το αρχικό σύνολο επικύρωσης (20%) διαιρέθηκε περαιτέρω προκειμένου να δημιουργηθούν το εσωτερικό σύνολο επικύρωσης (το οποίο χρησιμοποιείται από το ANFIS για τον πρόωρο τερματισμό) και το σύνολο αξιολόγησης καταλληλότητας (το οποίο χρησιμοποιείται από τον εξελικτικό αλγόριθμο για τον υπολογισμό του F1-score των υποψηφίων λύσεων). Το υποσύνολο τελικού ελέγχου (Final Test set) δεν χρησιμοποιήθηκε ώστε να γίνει αποφυγή δημιουργίας του φαινομένου διαρροής δεδομένων (Data Leakage / Optimization Bias), καθώς ο Γενετικός Αλγόριθμος θα χρησιμοποιούσε τα δεδομένα για να επιλέξει τις βέλτιστες παραμέτρους, χάνοντας έμμεσα την αντικειμενικότητά του. Παράλληλα, η αποφυγή της υπερεκπαίδευσης, πραγματοποιείται με την εφαρμογή πρόωρου τερματισμού. Με το φαινόμενο της υπερεκπαίδευσης το μοντέλο απομνημονεύει τα δεδομένα εκπαίδευσης αντί να αναγνωρίσει τη γενική λογική για τη λήψη αποφάσεων. Αυτή η κατάσταση είναι ανεπιθύμητη καθώς αν το σύστημα δεχόταν νέα άγνωστα δεδομένα είναι σίγουρο πως οι αποφάσεις του θα ήταν λανθασμένες λόγω της ανικανότητας του μοντέλου για γενίκευση. Στο τέλος της εκπαίδευσης του εσωτερικού ANFIS το μοντέλο αξιολογείται με βάση το υποσύνολο αξιολόγησης καταλληλότητας (Fitness evaluation set) και η τιμή καταλληλότητας (Fitness) του χρωμοσώματος αντιστοιχεί στην τιμή της μετρικής F1-score.

Η τιμή καταλληλότητας στη συνέχεια χρησιμοποιείται στον εξελικτικό βρόχο για την επιλογή γονέων που θα καθορίσουν τον πληθυσμό χρωμοσωμάτων της επόμενης γενιάς. Η επιλογή γονέων καθορίζει ποια χρωμοσώματα θα «αναπαραχθούν» για τη δημιουργία απογόνων στην επόμενη γενιά του εξελικτικού βρόχου. Για την επιλογή χρησιμοποιήθηκε η Tournament Selection, με την οποία ένα προκαθορισμένο πλήθος X τυχαίων χρωμοσωμάτων του πληθυσμού συμμετέχει σε ένα διαγωνισμό-τουρνουά στον οποίο κερδίζει μόνο το χρωμόσωμα που έχει την υψηλότερη βαθμολογία καταλληλότητας μεταξύ των συμμετεχόντων. Το χρωμόσωμα στη συνέχεια τοποθετείται σε μια προσωρινή λίστα γονέων που θα δημιουργήσουν τον πληθυσμό της επόμενης γενιάς. Η διαδικασία του τουρνουά επαναλαμβάνεται για τόσες φορές όσες είναι η προκαθορισμένη τιμή πληθυσμού. Έτσι, κατά την ολοκλήρωση της επιλογής γονέων τα χρωμοσώματα με μεγαλύτερη τιμή καταλληλότητας έχουν περισσότερες πιθανότητες να «επιβιώσουν» με την ύπαρξή τους μέσα στην προσωρινή λίστα γονέων πολλαπλές φορές ως κλώνοι. Αντίθετα ένα χρωμόσωμα με κακή τιμή καταλληλότητας πιθανότατα δεν θα εισαχθεί σε αυτή τη λίστα και δεν θα συνεχίσει στις επόμενες γενιές.

Έπειτα, αφού έχει γίνει η επιλογή γονέων-χρωμοσωμάτων, ακολουθεί η διαδικασία διασταύρωσης και μετάλλαξης. Λόγω των δύο διαφορετικών προσεγγίσεων αναπαραστάσης χρωμοσωμάτων, χρειάστηκε και διαφορετική υλοποίηση για τις διαδικασίες αυτές. Αρχικά για την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών με το υβριδικό χρωμόσωμα, η διασταύρωση γίνεται διαχωρίζοντας το χρωμόσωμα σε δύο τμήματα (χαρακτηριστικά και υπερπαραμέτροι) και σε δύο ομάδες για κάθε τμήμα (χαρακτηριστικά(1) και χαρακτηριστικά(2), υπερπαραμέτροι(1) και υπερπαραμέτροι(2)). Ο διαχωρισμός σε δύο ομάδες γίνεται ώστε τα χρωμοσώματα-γονείς να ανταλλάξουν σωστά το γενετικό τους υλικό χωρίς να υπάρχει πιθανότητα μίξης τιμών που προορίζονται για υπερπαραμέτρους με τιμές

για χαρακτηριστικά, χωρίζοντας το χρωμόσωμα του καθενός με τέτοιο τρόπο ώστε να πραγματοποιείται η σωστή δημιουργία νέων απογόνων. Έπειτα αφού έχει πραγματοποιηθεί ο διαχωρισμός, εφαρμόζεται διασταύρωση τύπου Δύο Σημείων (Two-Point Crossover) για τις ομάδες του δυαδικού τμήματος που αφορά τα χαρακτηριστικά. Η διασταύρωση τύπου δύο σημείων χωρίζει και τα δύο χρωμοσώματα-γονείς σε δύο σημεία δημιουργώντας τρία τμήματα στο καθένα, τα οποία στη συνέχεια επιλέγονται τυχαία για την παραγωγή των απογόνων. Η διασταύρωση για τις ομάδες του δεύτερου τμήματος των υπερπαραμέτρων του ANFIS πραγματοποιήθηκε με διασταύρωση τύπου Μίξης (Blend Crossover), η οποία είναι αρκετά δημοφιλής επιλογή διασταύρωσης χρωμοσωμάτων πραγματικών τιμών. Με τον συγκεκριμένο τύπο διασταύρωσης, παράγονται νέα γονίδια με βάση τη διαφορά τιμών των γονέων χρησιμοποιώντας ανώτατο και κατώτατο όριο, ενώ παράλληλα δίνεται η δυνατότητα με τον ορισμό της μεταβλητής α , να γίνει εξερεύνηση εκτός των ορίων με στόχο την ανακάλυψη νέων καλύτερων απογόνων. Η διασταύρωση τύπου Μίξης εφαρμόζεται αποκλειστικά σε ένα γονίδιο των δύο χρωμοσωμάτων-γονέων τη φορά, το οποίο είναι ιδιαίτερα σημαντικό λόγω της διαφοράς στα εύρη τιμών που έχει κάθε ένα γονίδιο του τμήματος υπερπαραμέτρων, καθώς η διασταύρωση μεταξύ αυτών θα οδηγούσε σε καταστροφικά αποτελέσματα. Τέλος, γίνεται η επανένωση των τμημάτων σε ενιαία υβριδικά χρωμοσώματα.

Η μετάλλαξη, της ελεύθερης προσέγγισης, γίνεται επίσης διαχωρίζοντας αρχικά το χρωμόσωμα σε δύο τμήματα που αντιστοιχούν στο δυαδικό χρωμόσωμα των χαρακτηριστικών και το χρωμόσωμα πραγματικών/ακέραιων τιμών για τις υπερπαραμέτρους του ANFIS. Για την εφαρμογή μετάλλαξης στο πρώτο δυαδικό τμήμα των χαρακτηριστικών, εφαρμόστηκε μετάλλαξη τύπου Αναστροφής Ψηφίου (Bit-flip Mutation), με την οποία κάθε γονίδιο έχει μια ποσοστιαία πιθανότητα να αλλάξει η δυαδική του τιμή. Στο δεύτερο τμήμα των υπερπαραμέτρων εφαρμόστηκε Γκαουσιανή Μετάλλαξη (Gaussian Mutation), με την οποία για το κάθε γονίδιο ορίζεται η τυπική απόκλιση που μπορεί να εφαρμοστεί στην τιμή κατά τη μετάλλαξη μέσα από τη μεταβλητή σ . Αυτή η δυνατότητα ορισμού διαφορετικής κλίμακας μετάλλαξη για κάθε γονίδιο είναι σημαντική λόγω της γνωστής διαφοράς κλίμακας τιμών των υπερπαραμέτρων. Έπειτα, τα δύο τμήματα επανενώνονται ξανά.

Στην προσέγγιση προκαθορισμένου πλήθους επιλεγμένων χαρακτηριστικών, η διασταύρωση ακολουθεί ουσιαστικά την ίδια πορεία με αυτή της προσέγγισης ελεύθερης επιλογής, ξεκινώντας από τον διαχωρισμό του χρωμοσώματος. Έπειτα, στο πρώτο τμήμα γίνεται διασταύρωση Δύο Σημείων, ενώ για το δεύτερο τμήμα γίνεται με διασταύρωση τύπου Μίξης. Τελευταίο βήμα είναι η σύνδεση των δύο τμημάτων σε ένα χρωμόσωμα. Ουσιαστικά η διασταύρωση και στις δύο προσεγγίσεις δεν έχει σημαντικές αλλαγές, με τη μοναδική αλλαγή να είναι στο μήκος του πρώτου τμήματος του χρωμοσώματος. Ωστόσο, είναι σημαντικό να αναφερθεί πως η χρήση διασταύρωσης σε αυτή την προσέγγιση χρωμοσώματος μπορεί να δημιουργήσει διπλότυπα στο πρώτο τμήμα του χρωμοσώματος, λόγω της τυχαίας ανταλλαγής τμημάτων. Για παράδειγμα η διασταύρωση δύο χρωμοσωμάτων που έχουν κάποιες κοινές τιμές μπορεί να αποδώσει ένα νέο χρωμόσωμα στο οποίο θα επαναλαμβάνονται συγκεκριμένες τιμές. Για την αποφυγή αυτού του φαινομένου, εφαρμόζεται ένας μηχανισμός επιδιόρθωσης. Ο μηχανισμός αυτός εκτελείται στην αρχή της συνάρτησης καταλληλότητας, όπου εντοπίζονται και διαγράφονται διπλότυπες τιμές για το πρώτο τμήμα που αφορά τα χαρακτηριστικά, αναπληρώνοντάς αυτές με νέες τιμές. Έτσι, εξασφαλίζεται η σωστή λειτουργία της διασταύρωσης και κατ' επέκταση της εξελικτικής διαδικασίας.

Για την εφαρμογή μετάλλαξης στη προσέγγιση προκαθορισμένου πλήθους επιλεγμένων χαρακτηριστικών, αρχικά γίνεται ο διαχωρισμός του ενιαίου σε αυτή την προσέγγιση, χρωμοσώματος πραγματικών/ακέραιων τιμών σε δύο τμήματα που αφορούν την επιλογή χαρακτηριστικών και υπερπαραμέτρων του ANFIS. Το πρώτο τμήμα χαρακτηριστικών, λόγω των ακέραιων δεκαδικών τιμών

που μπορεί να λάβει, δεν είναι δυνατό να χρησιμοποιηθεί μετάλλαξη αναστροφής ψηφίου. Οπότε, χρησιμοποιήθηκε μετάλλαξη Τυχαίας Αντικατάστασης (Random Resetting Mutation). Με αυτόν τον τύπου μετάλλαξης, για κάθε ένα από τα γονίδια-χαρακτηριστικά παράγεται ένας τυχαίος αριθμός με εύρος τιμών 0,0-1,0. Αν η τιμή αυτή είναι κάτω από το 0,1, η δεκαδική τιμή του γονιδίου αντικαθίσταται με μία άλλη τιμή-χαρακτηριστικό που δεν υπάρχει ήδη στο συγκεκριμένο τμήμα του χρωμοσώματος. Η μετάλλαξη στο δεύτερο τμήμα γίνεται με Γκαουσιανή μετάλλαξη, όπου με τη χρήση μεταβλητών σ ορίζονται οι τυπικές αποκλίσεις κάθε γονιδίου κατά τη μετάλλαξη. Στη συνέχεια, πραγματοποιείται επανένωση των δύο τμημάτων ξανά σε ένα ενιαίο.

5.4 Πειραματικό πρωτόκολλο για δυαδική ταξινόμηση

Για την αντικειμενική αξιολόγηση της απόδοσης του προτεινόμενου μοντέλου στη δυαδική ταξινόμηση, υιοθετήθηκε ένα συγκεκριμένο πειραματικό πρωτόκολλο. Κύριος στόχος των πειραμάτων είναι η αξιολόγηση της ικανότητας του μοντέλου να διαχωρίζει με ακρίβεια τα μοτίβα της δικτυακής κίνησης που προέρχονται από το ειδικά επεξεργασμένο για δυαδική ταξινόμηση σύνολο δεδομένων, όπως αναλύεται στο Κεφάλαιο 4. Στόχος είναι η ταξινόμηση εγγραφών σε δύο διακριτές κατηγορίες, τη φυσιολογική κίνηση (Normal) και την κακόβουλη κίνηση/εισβολή (Attack). Αρχικά καθορίζονται δύο μοντέλα αναφοράς (Baseline Models) που θα αποτελέσουν ένα αρχικό μέτρο σύγκρισης απόδοσης. Έπειτα, ακολουθεί το Σενάριο 1, στο οποίο ο γενετικός αλγόριθμος χρησιμοποιείται μόνο για την εύρεση των βέλτιστων υπερπαραμέτρων του ANFIS, οι οποίες έπειτα εφαρμόζονται στο τελικό μοντέλο ANFIS για την τελική αξιολόγηση. Στο Σενάριο 2, υλοποιείται η καθολική αρχιτεκτονική του προτεινόμενου μοντέλου, όπου με τη βοήθεια του γενετικού αλγορίθμου γίνεται αναζήτηση του βέλτιστου συνόλου χαρακτηριστικών και των υπερπαραμέτρων του ANFIS, τα οποία χρησιμοποιούνται σε μια τελική αξιολόγηση με το ANFIS.

5.4.1 Μοντέλα αναφοράς

Επιλέχθηκαν δύο μοντέλα αναφοράς, με αυτά να είναι ένα μοντέλο ANFIS το οποίο χρησιμοποιεί τις προεπιλεγμένες τιμές υπερπαραμέτρων που προτείνει η βιβλιοθήκη για γενική χρήση, ενώ το δεύτερο μοντέλο αναφοράς είναι ο κλασικός αλγόριθμος Μηχανικής Μάθησης Random Forest, για τον οποίο είναι γνωστή η εξαιρετική απόδοσή του σε δεδομένα μεγάλης διαστατικότητας.

Συγκεκριμένα για το ANFIS βασικών υπερπαραμέτρων, έγινε η αρχικοποίηση με FCM και 2 κανόνες/ομάδες (clusters), με τιμή 2,0 για τον εκθέτη ασαφοποίησης (Fuzziness Parameter – m) ο οποίος ουσιαστικά ελέγχει την ασάφεια των ομάδων, ενώ ως δείγμα για τον ορισμό της αρχικοποίησης αυτής χρησιμοποιήθηκε το 10% του αρχικού υποσυνόλου εκπαίδευσης. Για το εσωτερικό κομμάτι του ANFIS, χρησιμοποιήθηκε Γκαουσιανή συνάρτηση συμμετοχής, τιμή 0,01 για τον ρυθμό μάθησης (Learning Rate - lr), ενώ η εκπαίδευση ορίστηκε σε 20 εποχές με ενεργοποιημένη τη δυνατότητα πρόωρου τερματισμού.

Το δεύτερο μοντέλο αναφοράς που επιλέχθηκε, όπως αναφέρθηκε παραπάνω, είναι με τον αλγόριθμο Random Forest. Η επιλογή του συγκεκριμένου αλγορίθμου βασίστηκε στην ικανότητά του να διαχειρίζεται δεδομένα σε μορφή πίνακα (tabular data), να ανακαλύπτει πολύπλοκες μη γραμμικές σχέσεις μεταξύ των χαρακτηριστικών και να παρουσιάζει ανθεκτικότητα στο φαινόμενο της υπερεκπαίδευσης. Η υλοποίηση του αλγορίθμου έγινε με τη βοήθεια της βιβλιοθήκης scikit-learn, ενώ ορίστηκαν 100 δέντρα απόφασης με το μέγιστο βάθος κάθε δέντρου να περιορίζεται σε 20 επίπεδα. Οι συγκεκριμένες παράμετροι ορίστηκαν σε αυτές τις τιμές ώστε να μπορεί ο αλγόριθμος να ανταπεξέλθει

στην πολυπλοκότητα του συνόλου δεδομένων. Η χρήση Random Forest αποτελεί ένα αρκετά σημαντικό μοντέλο αναφοράς για τη σύγκριση με το προτεινόμενο μοντέλο.

Η υλοποίηση και αξιολόγηση αυτών των μοντέλων αναφοράς γίνεται σε ένα στάδιο, σε αντίθεση με τα επόμενα πειράματα που χρησιμοποιείται γενετικός αλγόριθμος αρχικά για την εύρεση παραμέτρων και στη συνέχεια υλοποιείται ένα τελικό ANFIS ως συνολική αξιολόγηση μοντέλου.

5.4.2 Υλοποίηση γενετικού αλγορίθμου για δυαδική ταξινόμηση

Στον γενετικό αλγόριθμο για τη γρήγορη λειτουργία του εσωτερικού μοντέλου ANFIS, γίνεται χρήση αντιπροσωπευτικού υποσυνόλου δεδομένων από τα αρχικά υποσύνολα εκπαίδευσης και επικύρωσης όπως αναλύθηκε σε προηγούμενη ενότητα, ώστε να γίνεται αξιολόγηση κάθε χρωμοσώματος μέσω της συνάρτησης καταλληλότητας. Λόγω της διαφοράς στην πολυπλοκότητα των πειραμάτων, ορίστηκαν διαφορετικές παράμετροι πληθυσμού. Ειδικότερα, στο Σενάριο 1 ο γενετικός αλγόριθμος ορίστηκε να παράγει 10 χρωμοσώματα (πληθυσμός) για καθεμία από τις 20 γενιές (συνολικά 200 χρωμοσώματα). Αντίθετα, στο Σενάριο 2, λόγω του σημαντικά μεγαλύτερου χώρου αναζήτησης, ο πληθυσμός αυξήθηκε στα 20 χρωμοσώματα για 40 γενιές. Η επιλογή γονέων και στα δύο σενάρια γίνεται με Tournament Selection, στην οποία γίνεται επιλογή 3 τυχαίων χρωμοσωμάτων για τη λειτουργία του που έχει αναλυθεί παραπάνω. Με την ολοκλήρωση της διαδικασίας επιλογής των γονέων, τα χρωμοσώματα τοποθετούνται σε ζευγάρια και με πιθανότητα 50% το κάθε ζευγάρι, θα διασταυρωθεί για να παραγάγει δύο νέους απογόνους που θα αντικαταστήσουν τα χρωμοσώματα-γονείς στην επόμενη γενιά. Έπειτα αφού τελειώσει η διαδικασία διασταύρωσης, ακολουθεί γκαουσιανή μετάλλαξη με την οποία κάθε ένα ξεχωριστό χρωμόσωμα έχει 30% πιθανότητα να υποστεί μετάλλαξη. Σε περίπτωση που δεν ικανοποιούνται οι πιθανότητες για την υλοποίηση διασταύρωσης και μετάλλαξης, τα χρωμοσώματα-γονείς περνούν αυτούσια στην επόμενη γενιά χωρίς αλλαγές. Με στόχο την ομαλή έναρξη της εξελικτικής διαδικασίας και έναν ελεγχόμενο χώρο λύσεων, ορίστηκαν για την αρχικοποίηση του συγκεκριμένα όρια αναζήτησης για τις υπερπαραμέτρους του ANFIS. Ωστόσο, για τη φάση της εξέλιξης τα άνω όρια αυτών των υπερπαραμέτρων αυξήθηκαν, επιτρέποντας στους γενετικούς τελεστές να εξερευνήσουν ευρύτερες περιοχές λύσεων. Οι ακριβείς τιμές των ορίων προσαρμόζονται ανάλογα με την πολυπλοκότητα του κάθε πειράματος και παρουσιάζονται στις επόμενες υποενότητες των αντίστοιχων σεναρίων. Τα όρια αναζήτησης καθορίστηκαν έπειτα από προκαταρκτικές εκτελέσεις και με βάση τους λειτουργικούς περιορισμούς της βιβλιοθήκης, ώστε να αποκλείονται μη λειτουργικές τιμές και να διατηρείται ο χρόνος εκτέλεσης σε εφικτά επίπεδα. Ωστόσο, πρέπει να αναγνωριστεί ότι οι επιλεγμένες τιμές περιορίζουν τον χώρο αναζήτησης και ενδέχεται να επηρεάζουν το τελικό αποτέλεσμα της εξελικτικής αναζήτησης. Είναι σημαντικό να αναφερθεί πως για τον τύπο της συνάρτησης συμμετοχής έχει πραγματοποιηθεί κατηγορική κωδικοποίηση, ώστε η τιμή 0 να αντιστοιχεί σε τριγωνική συνάρτηση συμμετοχής και η τιμή 1 σε γκαουσιανή συνάρτηση συμμετοχής.

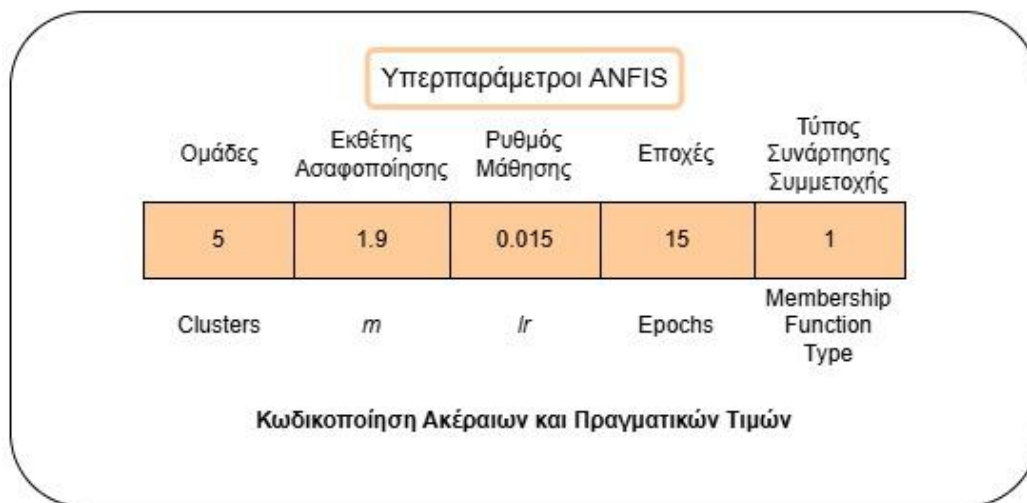
Επιπλέον, η αναπαράσταση του χρωμοσώματος σχεδιάστηκε με τέτοιο τρόπο ώστε να υποστηρίζει τόσο την απλή μορφή ενιαίου χρωμοσώματος (για το Σενάριο 1), όσο και την υβριδική μορφή για την ταυτόχρονη επιλογή χαρακτηριστικών, υποστηρίζοντας προσεγγίσεις ελεύθερου και προκαθορισμένου πλήθους χαρακτηριστικών, οι οποίες εφαρμόζονται στο Σενάριο 2. Για το Σενάριο 1, η παραμετροποίηση του γενετικού αλγορίθμου έγινε με τη χρήση των κλασικών λειτουργιών διασταύρωσης και μετάλλαξης, καθώς σε αυτό το πείραμα δεν χρησιμοποιείται υβριδική ή διαφοροποιημένη τμηματικά αναπαράσταση χρωμοσωμάτων. Αντίθετα, στο Σενάριο 2, οι διαδικασίες διασταύρωσης και μετάλλαξης υλοποιούνται με τις εξατομικευμένες διαδικασίες που αναλύθηκαν στην υποενότητα 5.3.

5.4.3 Σενάριο 1 – Βελτιστοποίηση υπερπαραμέτρων του ANFIS

Κρίθηκε αναγκαία η σταδιακή υλοποίηση του τελικού προτεινόμενου μοντέλου για την παρακολούθηση και επιβεβαίωση της σταδιακής σωστής λειτουργίας του, ενώ παράλληλα το συγκεκριμένο πείραμα αποτελεί ένα επιπλέον συγκριτικό μοντέλο. Για τον λόγο αυτό, δημιουργήθηκε το πειραματικό πλαίσιο Σενάριο 1. Στόχος του συγκεκριμένου πειράματος είναι η αξιολόγηση της ικανότητας του γενετικού αλγορίθμου να εντοπίσει τις βέλτιστες υπερπαραμέτρους για το ANFIS χωρίς να πραγματοποιεί επιλογή χαρακτηριστικών, χρησιμοποιώντας δηλαδή και τα 202 χαρακτηριστικά του συνόλου δεδομένων.

Η αναπαράσταση του χρωμοσώματος αυτού του πειράματος γίνεται με ένα ενιαίο χρωμόσωμα πραγματικών και ακέραιων τιμών το οποίο αποτελείται από 5 θέσεις-γονίδια. Τα γονίδια, τα οποία απεικονίζονται στο Σχήμα 5.3, αντιστοιχούν στο πλήθος των clusters, στον εκθέτη ασαφοποίησης (m), στον ρυθμό μάθησης (lr), στις εποχές και στον τύπο συνάρτησης συμμετοχής. Τα συγκεκριμένα χρησιμοποιούνται για την αρχικοποίηση με τον αλγόριθμο FCM αλλά και για την άμεση υπερπαραμετροποίηση του ANFIS.

Χρωμόσωμα του Σενάριο 1



Σχήμα 5.3: Αναπαράσταση χρωμοσώματος για το Σενάριο 1

Η τιμή καταλληλότητας αποτελείται από το F1-score της κλάσης επιθέσεων. Τα όρια των γονιδίων ορίστηκαν στο εύρος [2, 10] (με επέκταση έως 14 κατά την εξέλιξη) για τα clusters, στο [1, 2, 2, 5] (έως 2,8 κατά την εξέλιξη) για τον εκθέτη ασαφοποίησης, στο [0,0001, 0,03] (έως 0,04 κατά την εξέλιξη) για τον ρυθμό μάθησης, στο [5, 20] (έως 30 κατά την εξέλιξη) για το πλήθος εποχών και τέλος για τον τύπο συνάρτησης συμμετοχής το εύρος είναι [0, 1] και δεν γίνεται κάποια μεταβολή του.

5.4.4 Σενάριο 2 - Βελτιστοποίηση υπερπαραμέτρων και επιλογή χαρακτηριστικών

Στο Σενάριο 2, υλοποιείται η πλήρης αρχιτεκτονική του προτεινόμενου μοντέλου. Ο γενετικός αλγόριθμος έχει διττό ρόλο. Ο στόχος είναι να λυθεί το υψηλό υπολογιστικό κόστος λόγω της διαστατικότητας των δεδομένων, το οποίο είναι σημαντικό πρόβλημα για το ANFIS. Ο γενετικός αλγόριθμος επιλέγει μόνο τα πιο χρήσιμα από τα 202 χαρακτηριστικά, ενώ παράλληλα γίνεται και η αναζήτηση των βέλτιστων υπερπαραμέτρων για την εκπαίδευση του ANFIS.

Η αναπαράσταση των χρωμοσωμάτων σε αυτό το πείραμα εκπροσωπεί την επιλογή χαρακτηριστικών αλλά και την επιλογή υπερπαραμέτρων ANFIS, χωρίζοντας το χρωμόσωμα σε δύο τμήματα. Η υλοποίηση της αναπαράστασης των χρωμοσωμάτων έγινε με δύο προσεγγίσεις, η πρώτη αφορά την ελεύθερη επιλογή χαρακτηριστικών, η οποία αποτελείται από 207 γονίδια, ενώ η δεύτερη αφορά την επιλογή χαρακτηριστικών προκαθορισμένου πλήθους παράλληλα με την αναζήτηση υπερπαραμέτρων. Η διαφοροποίηση έγινε για να εξεταστεί η διαφορά στη μεταξύ τους απόδοση. Στην προσέγγιση προκαθορισμένου πλήθους χαρακτηριστικών, τιμή καταλληλότητας ενός χρωμοσώματος αποτελείται από την τιμή του F1-score της πρόβλεψης του εσωτερικού ANFIS για την κλάση επίθεσης. Ωστόσο, για την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών η τιμή καταλληλότητας αποτελείται από την πρόσθεση του F1-score με τη μεταβλητή *sparsity_bonus*, με την οποία όσο λιγότερα χαρακτηριστικά επιλέγονται τόσο αυξάνεται η τιμή της. Είναι πολύ σημαντικό να αναφερθεί πως η τιμή του *sparsity_bonus* ορίστηκε στο 0,05. Η συγκεκριμένη τιμή επιλέχθηκε εμπειρικά έπειτα από προκαταρκτικές δοκιμές, ώστε να επιτυγχάνεται ισορροπία μεταξύ της απόδοσης και του περιορισμού του αριθμού των χαρακτηριστικών, και αποτελεί σχεδιαστική επιλογή της παρούσας υλοποίησης. Ο σκοπός συμπερίληψής της στην τιμή καταλληλότητας είναι για τη διαφοροποίηση δύο χρωμοσωμάτων τα οποία έχουν ισοδυναμία στο F1-score, ευνοώντας το χρωμόσωμα με τα λιγότερα επιλεγμένα χαρακτηριστικά. Αυτό πραγματοποιήθηκε για την αποφυγή επιπλέον αχρείαστου υπολογιστικού κόστους που θα προκαλούσαν τα περισσότερα χαρακτηριστικά. Ο υπολογισμός της τιμής καταλληλότητας της προσέγγισης ελεύθερης επιλογής χαρακτηριστικών παρουσιάζεται παρακάτω στη σχέση (5.1):

$$Fitness\ score = (F1\ score) + sparsity_bonus \quad (5.1)$$

Τα όρια τιμών των γονιδίων για την αρχικοποίηση του πληθυσμού αλλά και στη συνέχεια για τη διαδικασία εξέλιξης, διατηρήθηκαν ίδια με το Σενάριο 1, με εξαίρεση τα όρια αρχικοποίησης για τα clusters, όπου το εύρος αυξήθηκε στο [2, 12] με το άνω όριο να αυξάνεται κατά την εξελικτική διαδικασία σε 14 και το όριο εκθέτη ασαφопоποίησης να μπορεί να αυξηθεί μέχρι και την τιμή 4.0. Ειδικότερα, κατά τη φάση της αρχικοποίησης του πληθυσμού με την πρώτη προσέγγιση αναπαράστασης χρωμοσώματος, δηλαδή με την ελεύθερη επιλογή χαρακτηριστικών, η απόδοση τιμών στο τμήμα των χαρακτηριστικών πραγματοποιείται με τυχαίο τρόπο, όπου κάθε γονίδιο-χαρακτηριστικό έχει ακριβώς 50% πιθανότητα να επιλεγεί (λαμβάνοντας την τιμή 1) και 50% πιθανότητα να μην επιλεγεί (λαμβάνοντας την τιμή 0). Κατά συνέπεια, κάθε άτομο του αρχικού πληθυσμού θα ξεκινάει, κατά μέσο όρο, με επιλεγμένα περίπου 100 από τα συνολικά 202 χαρακτηριστικά.

5.5 Πειραματικό πρωτόκολλο για ταξινόμηση πολλαπλών κλάσεων

Το πειραματικό πρωτόκολλο για την ταξινόμηση πολλαπλών κλάσεων έχει την ίδια φιλοσοφία με το πειραματικό πρωτόκολλο για τη δυαδική ταξινόμηση. Πλέον, ο στόχος των πειραμάτων δεν είναι μόνο η ανίχνευση κανονικής και εχθρικής δικτυακής κίνησης, αλλά και η αναγνώριση συγκεκριμένων κατηγοριών επιθέσεων. Τα πειράματα που θα ακολουθήσουν στοχεύουν στην αξιολόγηση της απόδοσης του προτεινόμενου μοντέλου, συγκρίνοντας τα αποτελέσματα κάθε πειράματος για την αναγνώριση των διαφορετικών κλάσεων. Τα δεδομένα που χρησιμοποιούνται έχουν επεξεργαστεί για να ικανοποιούν τις υψηλές απαιτήσεις των μοντέλων, καθώς η διαδικασία για τη λήψη σωστών αποφάσεων/προβλέψεων των μοντέλων για ταξινόμηση πολλαπλών κλάσεων είναι περισσότερο πολύπλοκη σε σχέση με τη δυαδική ταξινόμηση. Όπως στο πειραματικό πρωτόκολλο για δυαδική ταξινόμηση, έτσι και σε αυτό, γίνεται καθορισμός δύο αρχικών μοντέλων αναφοράς. Έπειτα,

ακολουθούν τα πειράματα Σενάριο 1 και Σενάριο 2 στα οποία γίνεται χρήση του γενετικού αλγορίθμου για την εύρεση των βέλτιστων παραμέτρων που στη συνέχεια θα εφαρμοστούν στην τελική αξιολόγηση.

5.5.1 Μοντέλα αναφοράς

Τα δύο μοντέλα αναφοράς που χρησιμοποιούνται λειτουργούν ως μια αφετηρία για την πορεία βελτιστοποίησης των αποτελεσμάτων με τα επόμενα πειράματα. Το πρώτο μοντέλο αναφοράς είναι ένα μοντέλο ANFIS το οποίο έχει κάποιες βασικές υπερπαραμέτρους, ενώ το δεύτερο μοντέλο αναφοράς είναι με τον αλγόριθμο Random Forest.

Για το πρώτο μοντέλο ANFIS χρησιμοποιείται ο FCM για την αρχικοποίηση και λόγω των υψηλότερων απαιτήσεων ορίστηκαν 10 κανόνες/ομάδες (clusters) και ο εκθέτης ασαφοποίησης (m) με τιμή 2,0. Χρησιμοποιήθηκε Γκαουσιανή συνάρτηση συμμετοχής, ενώ ο ρυθμός μάθησης έχει τιμή 0,01 και οι εποχές εκπαίδευσης είναι 20 με δυνατότητα πρόωρου τερματισμού αν ανιχνευθεί υπερεκπαίδευση.

Το δεύτερο μοντέλο αναφοράς με Random Forest ακολουθεί ακριβώς την ίδια παραμετροποίηση με αυτή που χρησιμοποιήθηκε για το αντίστοιχο μοντέλο στη δυαδική ταξινόμηση. Δηλαδή, γίνεται χρήση 100 δέντρων απόφασης με βάθος 20 επίπεδα. Επίσης, η υλοποίηση έγινε με τη βιβλιοθήκη scikit-learn.

Τα μοντέλα αναφοράς εκπαιδεύονται και αξιολογούνται σε μια μόνο φάση, αντίθετα στα επόμενα πειράματα Σενάριο 1 και Σενάριο 2, σε πρώτη φάση υλοποιείται ο γενετικός αλγόριθμος και στη συνέχεια ακολουθεί το τελικό στάδιο αξιολόγησης.

5.5.2 Υλοποίηση γενετικού αλγορίθμου για ταξινόμηση πολλαπλών κλάσεων

Η υλοποίηση του γενετικού αλγορίθμου για την ταξινόμηση πολλαπλών κλάσεων, όπως έχει ήδη αναφερθεί, γίνεται με τη χρήση της τεχνικής FGA με αντιπροσωπευτικά υποσύνολα δεδομένων για την επιτάχυνση της εξελικτικής διαδικασίας. Λόγω της αυξημένης πολυπλοκότητας της συγκεκριμένης ταξινόμησης, οι παράμετροι του γενετικού αλγορίθμου προσαρμόστηκαν αναλόγως. Ειδικότερα, στο Σενάριο 1 ο γενετικός αλγόριθμος ορίστηκε να παράγει πληθυσμό 10 χρωμοσωμάτων για 20 γενιές. Αντίστοιχα, για το Σενάριο 2, προκειμένου να καλυφθεί ο μεγαλύτερος χώρος αναζήτησης που περιλαμβάνει και την επιλογή χαρακτηριστικών, αυξήθηκε ο πληθυσμός με 20 χρωμοσώματα για 50 γενιές. Η επιλογή γονέων παραμένει με τη μέθοδο Tournament Selection με 3 τυχαία άτομα-χρωμοσώματα του πληθυσμού. Ανεξάρτητα από τους επιμέρους γενετικούς τελεστές, η γενική πιθανότητα διασταύρωσης ορίστηκε στο 50%, ενώ η πιθανότητα μετάλλαξης αυξήθηκε στο 50% σε σχέση με το 30% για τη δυαδική ταξινόμηση, ώστε να ενισχυθεί η εξερεύνηση για νέες λύσεις. Ωστόσο, οι διαδικασίες της διασταύρωσης και μετάλλαξης με βάση τα συγκεκριμένα ποσοστά είναι ανάλογες με αυτές του πειραματικού πρωτοκόλλου δυαδικής ταξινόμησης.

Τα όρια των παραμέτρων του γενετικού αλγορίθμου προσαρμόστηκαν επίσης στις απαιτήσεις της ταξινόμησης πολλαπλών κλάσεων. Τα όρια αυτά αφορούν τον περιορισμό των τιμών που μπορούν να πάρουν τα γονίδια των χρωμοσωμάτων, κατά την αρχικοποίηση του πληθυσμού αλλά και κατά την πορεία εξέλιξης, ώστε η εξελικτική διαδικασία να μην ακολουθήσει ανεπιθύμητη πορεία. Επίσης, έχει πραγματοποιηθεί κατηγορική κωδικοποίηση για τον τύπο συνάρτησης συμμετοχής.

Η αναπαράσταση των χρωμοσωμάτων για το Σενάριο 1 γίνεται με ένα ενιαίο χρωμόσωμα πραγματικών και ακέραιων τιμών 5 γονιδίων, ενώ για το Σενάριο 2 χρησιμοποιείται η υβριδική αναπαράσταση με χρωμοσώματος που αφορά την επιλογή χαρακτηριστικών και την υπερπαραμετροποίηση του ANFIS. Όπως και στον γενετικό αλγόριθμο για τη δυαδική ταξινόμηση, έτσι και εδώ η παραμετροποίηση του γενετικού αλγορίθμου έγινε με τις κλασικές διαδικασίες διασταύρωσης και μετάλλαξης για το Σενάριο

1, ενώ για το Σενάριο 2 έγινε χρήση των εξατομικευμένων διαδικασιών που αναλύονται στην υποενότητα 5.3 και εξυπηρετούν την υβριδική αναπαράσταση χρωμοσώματος.

5.5.3 Σενάριο 1 – Βελτιστοποίηση υπερπαραμέτρων

Στόχος στο Σενάριο 1 είναι η εύρεση των βέλτιστων υπερπαραμέτρων για το ANFIS χρησιμοποιώντας και τα 202 χαρακτηριστικά. Ουσιαστικά, η υλοποίηση είναι η ίδια με το αντίστοιχο πείραμα για τη δυαδική ταξινόμηση, όπου χρησιμοποιείται ένα χρωμόσωμα 5 γονιδίων που αντιστοιχούν στις υπερπαραμέτρους. Η διαφορά των δύο πειραμάτων βρίσκεται στα εύρη τιμών των γονιδίων, καθώς η πολυπλοκότητα της συγκεκριμένης ταξινόμησης είναι αυξημένη συγκριτικά με τη δυαδική ταξινόμηση. Συγκεκριμένα, τα εύρη τιμών κατά την αρχικοποίηση του πληθυσμού για τα clusters είναι στο [3, 15] (με επέκταση έως 18 κατά την εξέλιξη), για τον εκθέτη ασαφοποίησης [1,1 , 2.5] (έως 3,5 κατά την εξέλιξη), για τον ρυθμό μάθησης [0,001 , 0,03] (έως 0,04), για τις εποχές [5, 25] (έως 30) και για τον τύπο συνάρτησης συμμετοχής το εύρος είναι σταθερό στο [0, 1]. Επίσης, η τιμή καταλληλότητας ενός χρωμοσώματος αποτελείται από το μέσο F1-score όλων των κλάσεων.

5.5.4 Σενάριο 2 – Βελτιστοποίηση υπερπαραμέτρων και επιλογή χαρακτηριστικών

Στο Σενάριο 2 υλοποιείται η πλήρης αρχιτεκτονική του προτεινόμενου μοντέλου και για την ταξινόμηση πολλαπλών κλάσεων. Όπως στο αντίστοιχο πείραμα της δυαδικής ταξινόμησης, και σε αυτό ο γενετικός αλγόριθμος έχει διττό ρόλο, με την επιλογή χαρακτηριστικών αλλά και των υπερπαραμέτρων του ANFIS. Δηλαδή, η αναπαράσταση των χρωμοσωμάτων αποτελείται από δύο θεωρητικά τμήματα μέσα σε ένα χρωμόσωμα. Για αυτό το πείραμα χρησιμοποιήθηκαν οι δύο διαφορετικές προσεγγίσεις σχετικά με την ελεύθερη και την προκαθορισμένου πλήθους επιλογή χαρακτηριστικών. Η τιμή καταλληλότητας κάθε χρωμοσώματος της προσέγγισης προκαθορισμένου πλήθους επιλογής, προέρχεται από το μέσο F1-score όλων των κλάσεων που υπολογίζεται κατά την υλοποίηση του εσωτερικού μοντέλου αξιολόγησης ANFIS. Η τιμή καταλληλότητας της προσέγγισης ελεύθερης επιλογής υπολογίζεται με την πρόσθεση του μέσου F1-score και της μεταβλητής *sparsity_bonus*, ακριβώς όπως γίνεται και στο αντίστοιχο πείραμα της δυαδικής ταξινόμησης και παρουσιάζεται στη σχέση (5.1).

Τα όρια τιμών των γονιδίων που αφορούν τις υπερπαραμέτρους, είναι αυξημένα συγκριτικά με το Σενάριο 1 ακόμα και από την αρχικοποίηση του πληθυσμού. Το εύρος τιμών για τα clusters είναι [7, 15] (με επέκταση έως 22 κατά την εξέλιξη), για τον εκθέτη ασαφοποίησης [1,2 , 3,0] (έως 5,0), για τον ρυθμό μάθησης [0,001 , 0,02] (έως 0,03), για τις εποχές [5, 35] (έως 50) και ο τύπος συνάρτησης συμμετοχής όπως είναι γνωστό, παραμένει σταθερά στο εύρος [0, 1]. Κατά την αρχικοποίηση του πληθυσμού με την προσέγγιση ελεύθερου πλήθους επιλογής χαρακτηριστικών, η απόδοση τιμών στο πρώτο τμήμα που αφορά τα χαρακτηριστικά γίνεται όπως και στο αντίστοιχο πείραμα δυαδικής ταξινόμησης με πιθανότητα 50%. Επομένως, κατά μέσο όρο ο αρχικός πληθυσμός θα έχει περίπου 100 χαρακτηριστικά επιλεγμένα. Στη συνέχεια με την εξελικτική διαδικασία το πλήθος αλλά και η επιλογή συγκεκριμένων χαρακτηριστικών θα προσαρμοστούν κατάλληλα.

5.6 Τελική διαδικασία εκπαίδευσης και αξιολόγησης για Σενάριο 1 & Σενάριο 2

Η τελική φάση εκπαίδευσης και αξιολόγησης είναι όμοια για το πρωτόκολλο πειραμάτων δυαδικής ταξινόμησης και για το πρωτόκολλο ταξινόμησης πολλαπλών κλάσεων. Η εξελικτική διαδικασία για το Σενάριο 1 αλλά και για το Σενάριο 2, που ακολουθεί την παραμετροποίηση που αναλύθηκε στις προηγούμενες υποενότητες, ολοκληρώνεται με το πέρας της εκτέλεσης όλων των προκαθορισμένων γενεών. Έπειτα, αποθηκεύεται το χρωμόσωμα που είχε την καλύτερη τιμή καταλληλότητας και τα

γονίδια που προορίζονται για την υπερπαραμετροποίηση και την επιλογή συγκεκριμένων χαρακτηριστικών χρησιμοποιούνται για την υλοποίηση του τελικού ANFIS με τα πλήρη υποσύνολα εκπαίδευσης (Training), επικύρωσης (Validation) και ελέγχου (Final Test), ώστε να γίνεται εκμετάλλευση όλης της διαθέσιμης πληροφορίας. Σε αυτό το τελικό ANFIS, η αρχικοποίηση με FCM γίνεται με την ίδια φιλοσοφία που ακολουθεί το εσωτερικό «ελαφρύ» ANFIS στον γενετικό αλγόριθμο, όπου δηλαδή γίνεται χρήση του 10% του εκάστοτε διαθέσιμου υποσυνόλου εκπαίδευσης. Η αρχικοποίηση και εκπαίδευση του μοντέλου πραγματοποιείται με συγκεκριμένες τιμές που προέρχονται από το βέλτιστο χρωμόσωμα και στη συνέχεια γίνεται η αξιολόγηση του μοντέλου με το υποσύνολο τελικού ελέγχου (Final Test set), το οποίο χρησιμοποιείται για πρώτη και μοναδική φορά σε όλη τη διαδικασία του προτεινόμενου μοντέλου, διασφαλίζοντας την αποφυγή του φαινομένου διαρροής δεδομένων. Για την τελική αξιολόγηση χρησιμοποιούνται εξειδικευμένες μετρικές προσαρμοσμένες στη φύση των δύο διαφορετικών ταξινομήσεων, οι οποίες αναλύονται εκτενώς στο επόμενο κεφάλαιο.

Κεφάλαιο 6ο: Πειραματικά αποτελέσματα και ανάλυση

6.1 Μετρικές αξιολόγησης

Για τον υπολογισμό και την εξαγωγή των μετρικών αξιολόγησης του προτεινόμενου μοντέλου αλλά και των υπόλοιπων πειραμάτων, έγινε χρήση της βιβλιοθήκης `scikit-learn`. Η αξιολόγηση των πειραμάτων, τόσο της δυαδικής ταξινόμησης όσο και της ταξινόμησης πολλαπλών κλάσεων, ως προς την ικανότητά τους να διαχωρίζουν και να ταξινομούν σωστά έγινε με τις ακόλουθες μετρικές.

Η πρώτη μετρική είναι η Ορθότητα (Accuracy), η οποία είναι μια γενική μετρική και εκφράζει το ποσοστό των συνολικά σωστών προβλέψεων κάθε κλάσης. Αυτή αποτελεί την πιο δημοφιλή μετρική απόδοσης μοντέλων στον τομέα της Μηχανικής Μάθησης. Ωστόσο, σε ανισορροπημένα δεδομένα, όπως αυτά του τομέα των IDS, το Accuracy δεν αποτελεί τη χρησιμότερη και πιο αξιόπιστη μετρική αξιολόγησης [51].

Άλλη μετρική είναι η Ακρίβεια (Precision), με την οποία εκφράζεται η αξιοπιστία του μοντέλου όταν κάνει πρόβλεψη μιας συγκεκριμένης κλάσης και η εγγραφή αντιστοιχεί πραγματικά σε αυτή την κλάση. Δηλαδή για παράδειγμα, η ακρίβεια εκφράζει το κατά πόσο σίγουρο είναι το μοντέλο ότι οι ταξινομήσεις που κάνει στην κατηγορία επιθέσεων Fuzzers είναι και στην πραγματικότητα κατηγορίας Fuzzers.

Η Ανάκληση (Recall) είναι μια επιπλέον μετρική, με την οποία αντικατοπτρίζεται η ευαισθησία του μοντέλου στην ανίχνευση συγκεκριμένων κλάσεων. Υποδηλώνει το ποσοστό εγγραφών που το μοντέλο κατάφερε να εντοπίσει σωστά, ανεξάρτητα από το αν ταξινόμησε επιπλέον λανθασμένα εγγραφές που δεν αντιστοιχούν πραγματικά στη ζητούμενη κλάση.

Ο συνδυασμός των δύο προηγούμενων μετρικών, της ακρίβειας και της ανάκλησης, μπορεί να εκφραστεί σε μία μετρική με την ονομασία F1-score. Η συγκεκριμένη αποτελεί μια από τις σημαντικότερες μετρικές, καθώς η υψηλή τιμή F1-score προϋποθέτει ότι το μοντέλο έχει ταυτόχρονα υψηλή ακρίβεια και υψηλή ανάκληση. Αξίζει να αναφερθεί πως στα πειράματα χρησιμοποιήθηκε το macro-average F1-score λόγω της ανισορροπίας των δεδομένων μεταξύ των διαφορετικών κλάσεων. Η macro μέθοδος δίνει ίση βαρύτητα σε όλες τις κλάσεις, διασφαλίζοντας ότι η απόδοση στις σπάνια εμφανιζόμενες κλάσεις δεν επισκιάζεται από τις κλάσεις με συχνότερη εμφάνιση.

Η μετρική PR-AUC (Precision-Recall Area Under Curve) προσφέρει μια συνολική εικόνα σταθερότητας και απόδοσης των μοντέλων στα πειράματα δυαδικής ταξινόμησης. Αποτελεί ιδανικό δείκτη για μη ισορροπημένα σύνολα δεδομένων, όπου η σωστή ανίχνευση της μειονοτικής κλάσης (επιθέσεις) είναι ζωτικής σημασίας.

Για την πλήρη κατανόηση της συμπεριφοράς των μοντέλων, μελετώνται επίσης δύο κρίσιμα μεγέθη που προκύπτουν από τον Πίνακα Σύγχυσης (Confusion Matrix). Τα Ψευδώς Θετικά (False Positives) αντιπροσωπεύουν τη φυσιολογική δικτυακή κίνηση που έχει ταξινομηθεί λανθασμένα ως επίθεση, δημιουργώντας ψευδείς συναγερμούς (False Alarms). Αντίθετα, τα Ψευδώς Αρνητικά (False Negatives) είναι οι πραγματικές επιθέσεις τις οποίες το μοντέλο ταξινόμησε ως κανονική δικτυακή κίνηση, αποτυγχάνοντας έτσι να τις εντοπίσει (Missed Attacks). Η ελαχιστοποίηση των ψευδών αρνητικών αποτελεί την ύψιστη προτεραιότητα ενός συστήματος ανίχνευσης εισβολών, καθώς μια «χαμένη» επίθεση μπορεί να αποβεί μοιραία για την ασφάλεια και τη λειτουργία ενός δικτύου.

6.2 Αποτελέσματα δυαδικής ταξινόμησης

Σε αυτή την ενότητα παρουσιάζονται και αναλύονται τα αποτελέσματα των πειραμάτων της δυαδικής ταξινόμησης. Υπενθυμίζεται, πως ο στόχος σε αυτή τη ταξινόμηση είναι ο διαχωρισμός της δικτυακής κίνησης σε φυσιολογική (normal) και επίθεση (attack).

6.2.1 Συγκριτική αξιολόγηση μοντέλων

Παρακάτω, στον Πίνακα 6.1 παρουσιάζονται οι υπερπαράμετροι του τελικού μοντέλου ANFIS για τα πειράματα. Υπενθυμίζεται ότι στο δεύτερο μοντέλο αναφοράς με Random Forest χρησιμοποιήθηκαν 100 δέντρα απόφασης με 20 επίπεδα. Στον Πίνακα 6.2 παρουσιάζονται οι μετρικές απόδοσης των δύο μοντέλων αναφοράς καθώς και των Σεναρίων 1 και 2. Η διαδικασία υλοποίησης του τελικού ANFIS εκτελέστηκε για 10 ανεξάρτητες φορές και τα αποτελέσματα του Πίνακα 6.2 προέρχονται από την εκτέλεση με την κορυφαία επίδοση (best run) κάθε πειραματικής διάταξης (setup). Επίσης, είναι σημαντικό να αναφερθεί πως ο γενετικός αλγόριθμος έχει εκτελεστεί τουλάχιστον τρεις φορές για κάθε πείραμα, ώστε να επιβεβαιωθεί η σταθερότητα των αποτελεσμάτων.

Από τον Πίνακα 6.1 φαίνεται πως ο γενετικός αλγόριθμος σε όλα τα setup κατέληξε με περισσότερα από 2 clusters, που ορίστηκαν για το πρώτο μοντέλο αναφοράς. Παράλληλα και ο εκθέτης ασαφοποίησης (m) έχει τιμές πάνω από το 2, κυμαινόμενος από 2,22 μέχρι 3,51. Το φαινόμενο αυτό υποδηλώνει ότι για την καλύτερη διαχείριση των δεδομένων, που οδηγεί σε καλύτερα αποτελέσματα, χρειάστηκε να χρησιμοποιηθούν περισσότεροι ασαφείς κανόνες-ομάδες αλλά και μεγαλύτερη ασάφεια μεταξύ των ομάδων αυτών από το αρχικό μοντέλο αναφοράς ANFIS.

Πίνακας 6.1: Υπερπαράμετροι ANFIS των πειραμάτων δυαδικής ταξινόμησης

Ονομασία Setup	Προσέγγιση Χρωμοσώματος	Υπερπαράμετροι ANFIS				
		Clusters	m	Ρυθμός Μάθησης	Εποχές	Συνάρτηση Συμμετοχής
1^ο Μοντέλο Αναφοράς Κλασικό ANFIS	-	2	2,0	0,01	20	Gaussian
2^ο Μοντέλο Αναφοράς Random Forest	-	-	-	-	-	-
Σενάριο 1	-	4	2,22	0,03	10	Gaussian
Σενάριο 2 (Fixed 10)	Προκαθορισμένο Πλήθος 10	6	2,54	0,0108	22	Triangular
Σενάριο 2 (Fixed 45)	Προκαθορισμένο Πλήθος 45	10	3,34	0,014	25	Gaussian
Σενάριο 2 (Fixed 60)	Προκαθορισμένο Πλήθος 60	3	2,75	0,013	9	Gaussian
Σενάριο 2 (Fixed 120)	Προκαθορισμένο Πλήθος 120	9	3,51	0,035	10	Triangular
Σενάριο 2 (Free 69)	Ελεύθερο Πλήθος 69	9	2,73	0,002	30	Gaussian
Σενάριο 2 (Free 100)	Ελεύθερο Πλήθος 100	5	2,60	0,004	18	Triangular

Πίνακας 6.2: Αποτελέσματα δυαδικής ταξινόμησης (βέλτιστες εκτελέσεις)

Setup	Accuracy	F1-score	PR-AUC	False Positives	False Negatives	Χρόνος Εκπαίδευσης
1^ο Μοντέλο Αναφοράς Κλασικό ANFIS	0,9858	0,9469	0,8992	7114	83	147 sec
2^ο Μοντέλο Αναφοράς Random Forest	0,9894	0,9597	0,9982	5312	83	61 sec
Σενάριο 1	0,9862	0,9481	0,9013	6953	78	570 sec
Σενάριο 2 (Fixed 10)	0,9861	0,9476	0,9006	6830	245	167 sec
Σενάριο 2 (Fixed 45)	0,9877	0,9535	0,9111	6210	58	579 sec
Σενάριο 2 (Fixed 60)	0,9848	0,9433	0,8927	7644	77	165 sec
Σενάριο 2 (Fixed 120)	0,9875	0,953	0,9102	6308	32	1004 sec
Σενάριο 2 (Free 69)	0,9863	0,9487	0,9025	6877	59	914 sec
Σενάριο 2 (Free 100)	0,9877	0,9536	0,9113	6237	18	846 sec

Παρατηρώντας τα αποτελέσματα του Πίνακα 6.2, είναι φανερό ότι ο αλγόριθμος Random Forest έχει τα καλύτερα αποτελέσματα, τα οποία αποτυπώνονται με τη μεγάλη διαφορά της μετρικής PR-AUC, η οποία έχει τιμή πολύ κοντά στο απόλυτο 1. Ωστόσο, όπως προαναφέρθηκε, το Random Forest χρησιμοποιείται ως ισχυρό μοντέλο αναφοράς το οποίο έχει περιορισμένη εγγενή ερμηνευσιμότητα και συνήθως απαιτεί μεταγενέστερες τεχνικές επεξήγησης. Αντίθετα, το ANFIS παρέχει εγγενή ερμηνευσιμότητα, επειδή η εσωτερική του λειτουργία μπορεί να εξεταστεί άμεσα μέσω των ασαφών κανόνων.

Εξετάζοντας τα υπόλοιπα μοντέλα που βασίζονται στο ANFIS, η συμβολή του γενετικού αλγορίθμου είναι φανερή. Η βελτιστοποίηση των υπερπαραμέτρων με το Σενάριο 1 κατάφερε να βελτιώσει την απόδοση σε σχέση με το ANFIS μοντέλο αναφοράς, παρουσιάζοντας οριακές αυξήσεις στο Accuracy, στο F1-score, PR-AUC και μια μικρή μείωση στα False Positives και False Negatives.

Η πραγματική δυναμική του προτεινόμενου μοντέλου αποκαλύπτεται από τα αποτελέσματα των πειραμάτων με το Σενάριο 2, όπου η βελτιστοποίηση των υπερπαραμέτρων συνδυάζεται με την επιλογή χαρακτηριστικών. Το βέλτιστο μοντέλο συνολικά από τα πειράματα του Σεναρίου 2 προέκυψε από την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών και συγκεκριμένα το setup που κατέληξε με 100 χαρακτηριστικά. Το συγκεκριμένο μοντέλο πέτυχε τις μεγαλύτερες τιμές σε Accuracy, F1-score και PR-AUC μεταξύ των μοντέλων που χρησιμοποιούν ANFIS, παρουσιάζοντας σημαντική βελτίωση. Οι τιμές των μετρικών αυτού του μοντέλου προσεγγίζουν αυτές του Random Forest χωρίς όμως να τις ξεπερνούν. Ωστόσο, υπάρχει μια καθοριστική διαφορά στα αποτελέσματα του βέλτιστου μοντέλου σε σχέση με αυτά του Random Forest, η οποία αφορά τα False Negatives και θα αναλυθεί στις επόμενες ενότητες.

6.2.2 Ανάλυση επιλογής χαρακτηριστικών

Γενικά στα συστήματα ανίχνευσης εισβολών, η μεγάλη διαστατικότητα των δεδομένων μπορεί να έχει καθοριστικό ρόλο στη λειτουργία τους, καθώς επηρεάζει την εκπαίδευση των μοντέλων Μηχανικής Μάθησης προσθέτοντας υπολογιστικό κόστος και «θόρυβος» που αλλοιώνει την τελική τους απόδοση. Στο Σενάριο 2 εξετάστηκαν δύο διαφορετικές προσεγγίσεις επιλογής χαρακτηριστικών (Feature Selection) μέσω του γενετικού αλγορίθμου, των οποίων οι λεπτομέρειες αναλύονται στην ενότητα 5.3.

Αναλύοντας τα αποτελέσματα, τα δύο πειράματα με την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών έχουν πολύ καλή απόδοση. Ουσιαστικά, ο ίδιος γενετικός αλγόριθμος σε δύο διαφορετικές εκτελέσεις επέλεξε 69 και 100 χαρακτηριστικά αντίστοιχα, ενώ παράλληλα με τις υπερπαραμέτρους από τα χρωμοσώματα αυτών εκτελέστηκαν ξεχωριστά οι τελικές αξιολογήσεις. Από τα δύο πειράματα, αυτό με τα 100 χαρακτηριστικά έχει τα καλύτερα αποτελέσματα σε όλες τις μετρικές. Σε ολική σύγκριση του συγκεκριμένου πειράματος με όλα τα υπόλοιπα πειράματα με ANFIS, συμπεριλαμβανομένων των πειραμάτων με την προσέγγιση προκαθορισμένου πλήθους χαρακτηριστικών, φαίνεται πως είναι το ολικά καλύτερο σε απόδοση μοντέλο. Η επιλογή των 100 χαρακτηριστικών μείωσε τη διαστατικότητα του συνόλου δεδομένων σχεδόν στο μισό, απορρίπτοντας 102 χαρακτηριστικά. Έτσι, η απομάκρυνση της περιττής πληροφορίας διευκολύνει την εκπαίδευση του τελικού μοντέλου ANFIS.

Πέραν όμως του πειράματος «νικητή», τα πειράματα με προσέγγιση προκαθορισμένου πλήθους ανέδειξαν σημαντικά ευρήματα ως προς τον περιορισμό του πλήθους χαρακτηριστικών και τη συμπίεση του χώρου διαθέσιμης πληροφορίας. Τα πειράματα έγιναν με προκαθορισμένες τιμές πλήθους 10, 45, 60 και 120 χαρακτηριστικών. Το πείραμα με το πλήθος 10 χαρακτηριστικών χρησιμοποιεί περίπου το 5% του πλήθους των αρχικών χαρακτηριστικών και παρόλο που δεν έχει τα καλύτερα αποτελέσματα μεταξύ των πειραμάτων αυτής της προσέγγισης, εμφανίζει σημαντική ανθεκτικότητα παρουσιάζοντας βελτίωση σε Accuracy, F1-score και PR-AUC σε σύγκριση με το πρώτο μοντέλο αναφοράς που χρησιμοποιεί και τα 202 διαθέσιμα χαρακτηριστικά, παρουσιάζοντας όμως περισσότερες χαμένες επιθέσεις. Αυτό το εύρημα αναδεικνύει την ικανότητα του γενετικού αλγορίθμου να εντοπίζει σε έναν αρκετά περιορισμένο χώρο τα σημαντικότερα χαρακτηριστικά, τα οποία όμως δεν είναι τελικά αρκετά για συνολικά ικανοποιητικές τιμές απόδοσης, καθώς υπάρχει μεγάλος αριθμός των False Negatives. Στο αντίθετο άκρο, το πείραμα με πλήθος 120 χαρακτηριστικών αναγκάζει τον γενετικό αλγόριθμο να αναζητήσει περισσότερα χαρακτηριστικά από αυτά του πειράματος ελεύθερης επιλογής, οδηγώντας σε καλύτερες επιδόσεις από αυτές του πειράματος 10 χαρακτηριστικών, δημιουργώντας μεγάλο υπολογιστικό κόστος το οποίο αποτυπώνεται στον υψηλό χρόνο εκπαίδευσης του τελικού ANFIS. Τα πειράματα με πλήθος 60 και 45 χαρακτηριστικών αποτελούν μια μέση λύση, με το πείραμα των 45 χαρακτηριστικών να έχει τα καλύτερα αποτελέσματα μεταξύ των πειραμάτων προκαθορισμένου πλήθους. Το συγκεκριμένο πείραμα, ξεπερνάει τα αποτελέσματα του «βαρύτερου» μοντέλου 120 χαρακτηριστικών σε Accuracy και F1-score, ενώ ισοφαρίζει και το Accuracy του μοντέλου ελεύθερης επιλογής με 100 χαρακτηριστικά. Ωστόσο, δεν μπορεί το «ελαφρύτερο» μοντέλο με τα 45 χαρακτηριστικά να θεωρηθεί αυτό με τη συνολικά κορυφαία απόδοση, καθώς εμφανίζει μια σημαντική διαφορά με το μοντέλο των 100 χαρακτηριστικών η οποία σχετίζεται με τα Ψευδώς Αρνητικά και θα αναλυθεί στην επόμενη υποενότητα.

6.2.3 Ανάλυση False Positives και False Negatives

Όπως είναι γνωστό, στα συστήματα ανίχνευσης εισβολών, ύψιστη προτεραιότητα έχει η ανίχνευση όσο το δυνατόν περισσότερων εισβολών. Συνεπώς, οι μετρικές Accuracy, F1-score και PR-AUC δεν αρκούν

από μόνες τους για να αποδείξουν την ανωτερότητα ενός μοντέλου. Για τον λόγο αυτό, στα αποτελέσματα συμπεριλαμβάνονται τα Ψευδώς Θετικά (False Positives – Ψευδείς Συναγερμοί) και τα Ψευδώς Αρνητικά (False Negatives – Χαμένες επιθέσεις).

Από τα αποτελέσματα του Πίνακα 6.2 αναδεικνύεται η πραγματική υπεροχή του Σεναρίου 2 έναντι του Random Forest στον τομέα των False Negatives. Το Random Forest, αν και έχει τις καλύτερες τιμές από όλα τα μοντέλα στις μετρικές Accuracy, F1-score και PR-AUC, απέτυχε να αναγνωρίσει 83 πραγματικές επιθέσεις, τις οποίες ταξινόμησε λανθασμένα ως φυσιολογική κίνηση δικτύου.

Αντίθετα, το μοντέλο του Σεναρίου 2 με την ελεύθερη επιλογή και 100 χαρακτηριστικά μείωσε δραματικά το πλήθος των False Negatives. Στη βέλτιστη εκτέλεση το συγκεκριμένο μοντέλο περιόρισε τις χαμένες επιθέσεις σε μόλις 18, επιτυγχάνοντας εξαιρετική ευαισθησία στην αναγνώριση επιθέσεων, ενώ η μέση συμπεριφορά των 10 εκτελέσεων παρουσιάζεται αναλυτικά στον αντίστοιχο πίνακα στατιστικής ανάλυσης της επόμενης ενότητας. Η μείωση των False Negatives συνοδεύτηκε από την αναμενόμενη αύξηση των False Positives (6.237 έναντι 5.312 του Random Forest). Ωστόσο, σε ένα πραγματικό περιβάλλον κυβερνοασφάλειας, αυτός ο συμβιβασμός (trade-off) κρίνεται αποδεκτός και αποτελεί ορθή πρακτική. Η συγκεκριμένη προσέγγιση, όπως αναφέρει και το [5], αποτελεί μια καθιερωμένη πρακτική για πολλούς οργανισμούς, οι οποίοι δίνουν προτεραιότητα στη μείωση των False Negatives, ακόμα και αν αυτό οδηγήσει σε αναπόφευκτη αύξηση των False Positives. Αυτό συμβαίνει διότι σε πραγματικά περιβάλλοντα είναι προτιμότερο να δαπανηθούν επιπλέον πόροι για την επαλήθευση των ψευδών συναγερμών, παρά μια πραγματική επίθεση να περάσει απαρατήρητη από το σύστημα. Ο πίνακας σύγχυσης του Σεναρίου 2 ελεύθερης επιλογής 100 χαρακτηριστικών παρουσιάζεται στο Σχήμα 6.1.

ANFIS Binary Confusion Matrix - Scenario 2 (Free 100)

ΠΡΑΓΜΑΤΙΚΗ ΚΛΑΣΗ	Normal	437516	6237
	Attack	18	64239
		Normal	Attack

ΠΡΟΒΛΕΨΗ

Σχήμα 6.1: Πίνακας σύγχυσης δυαδικής ταξινόμησης

Αξιοσημείωτη είναι και η συμπεριφορά του μοντέλου προκαθορισμένου πλήθους 45 χαρακτηριστικών. Το συγκεκριμένο μοντέλο σημειώνει καλύτερη τιμή False Negatives από το Random Forest, με 58 έναντι 83 χαμένων επιθέσεων. Βέβαια, παρουσιάζει περισσότερα False Positives, ωστόσο η ικανότητά του να εντοπίζει περισσότερες πραγματικές επιθέσεις χρησιμοποιώντας λιγότερο από το ένα τέταρτο των αρχικών χαρακτηριστικών του συνόλου δεδομένων UNSW-NB15 αναδεικνύει τη δυναμική του προτεινόμενου υβριδικού μοντέλου ANFIS-GA.

6.3 Αποτελέσματα ταξινόμησης πολλαπλών κλάσεων

Σε αυτή την ενότητα παρουσιάζονται τα αποτελέσματα των πειραμάτων ταξινόμησης πολλαπλών κλάσεων, όπου υπάρχουν 10 (9 επιθέσεων + 1 φυσιολογική κίνηση) διαφορετικές κλάσεις για να αναγνωριστούν από τα μοντέλα. Όπως έχει ήδη αναφερθεί η συγκεκριμένη ταξινόμηση είναι πιο πολύπλοκη από τη δυαδική, καθώς πλέον τα μοντέλα καλούνται όχι μόνο να εντοπίσουν μια επίθεση, αλλά και να την ταξινομήσουν με τη μεγαλύτερη δυνατή ακρίβεια στη σωστή κατηγορία επιθέσεων.

6.3.1 Συγκριτική αξιολόγηση μοντέλων

Στον Πίνακα 6.3 παρουσιάζονται οι υπερπαραμέτροι που προέκυψαν από τον γενετικό αλγόριθμο για τα Σενάρια 1 και 2, καθώς και οι σταθερές υπερπαραμέτροι για τα μοντέλα αναφοράς. Με αυτές πραγματοποιείται η υλοποίηση των τελικών μοντέλων ANFIS. Σε αυτό το πειραματικό πρωτόκολλο της συγκεκριμένης ταξινόμησης, ο γενετικός αλγόριθμος εκτελέστηκε τουλάχιστον τρεις φορές για την επιβεβαίωση της σταθερότητας των αποτελεσμάτων του.

Παρατηρώντας τον Πίνακα 6.3, γίνεται άμεσα αντιληπτή η αυξημένη πολυπλοκότητα της ταξινόμησης πολλαπλών κλάσεων. Ο γενετικός αλγόριθμος συνέκλινε σε σημαντικά μεγαλύτερο αριθμό κανόνων-ομάδων (clusters), φτάνοντας μέχρι και 19, ενώ αντίστοιχα στη δυαδική ταξινόμηση ο μέγιστος αριθμός ήταν 10. Παράλληλα, ο εκθέτης ασαφοποίησης (m) παρουσίαζε σημαντικά χαμηλότερες τιμές (από 1,26 έως 2,22). Ο συνδυασμός των πολλών clusters και της χαμηλής τιμής του m υποδηλώνει ότι ο διαχωρισμός των πολλαπλών κλάσεων απαιτεί μεγάλο αριθμό ασαφών κανόνων με πιο «αυστηρά» και ξεκάθαρα όρια μεταξύ αυτών, προκειμένου τα μοντέλα να μπορέσουν να διακρίνουν τις διαφορετικές κατηγορίες κλάσεων. Υπενθυμίζεται ότι στο Random Forest χρησιμοποιούνται 100 δέντρα απόφασης με 20 επίπεδα.

Στον πίνακα 6.4 εμφανίζονται συγκεντρωτικά οι μετρικές των κορυφαίων, μεταξύ 10 εκτελέσεων για κάθε setup, επιδόσεων. Όπως έχει αναφερθεί και στη μεθοδολογία, η μετρική F1-score που χρησιμοποιείται στο Σενάριο 1 και 2 είναι η μέση τιμή F1-score των επιμέρους κλάσεων, δηλαδή το Macro-Average F1-score. Με μια πρώτη εκτίμηση των αποτελεσμάτων, επιβεβαιώνεται η δυσκολία της ταξινόμησης πολλαπλών κλάσεων από τις τιμές που έχει το Macro-Average F1-score ανεξάρτητα από το μοντέλο και το setup, σε σύγκριση πάντα με τα αποτελέσματα της δυαδικής ταξινόμησης. Το Accuracy διατηρεί υψηλές τιμές, οι οποίες όμως οφείλονται στην κυριαρχία της κλάσης Normal η οποία έχει το μεγαλύτερο πλήθος εγγραφών στο υποσύνολο τελικού ελέγχου (Final test set), το οποίο είναι ανέπαφο από την προεπεξεργασία με SMOTE και αντικατοπτρίζει τις πραγματικές αναλογίες που παρουσιάζονται σε δικτυακή κίνηση. Ωστόσο, η ικανότητα των μοντέλων να ταξινομούν σωστά δοκιμάζεται έντονα.

Το 2^ο μοντέλο αναφοράς με τον αλγόριθμο Random Forest στην ταξινόμηση πολλαπλών κλάσεων εμφανίζει τα καλύτερα αποτελέσματα συγκριτικά με όλα τα μοντέλα, σημειώνοντας 0 χαμένες επιθέσεις. Ωστόσο, όπως έχει αναφερθεί, η εσωτερική λογική του Random Forest δεν είναι άμεσα αναγνώσιμη σε μορφή ενός συνόλου κανόνων και η ερμηνεία του απαιτεί πρόσθετες τεχνικές επεξήγησης.

Ανάμεσα στα μοντέλα που χρησιμοποιούν το ANFIS, ο γενετικός αλγόριθμος συμβάλλει στη βελτίωση της απόδοσης και αυτό παρουσιάζεται ακόμα και από τα αποτελέσματα του Σεναρίου 1. Το Σενάριο 1 με τη βελτιστοποίηση των υπερπαραμέτρων του ANFIS κατάφερε να βελτιώσει ελαφρώς την απόδοση σε σχέση με το 1^ο μοντέλο αναφοράς σε όλες τις μετρικές και μειώνοντας παράλληλα τις χαμένες επιθέσεις από 59 σε 33. Παράλληλα, η βελτίωση αυτή επιτεύχθηκε χρησιμοποιώντας λιγότερα clusters, γεγονός που μείωσε το υπολογιστικό κόστος, το οποίο αντικατοπτρίζεται στον χαμηλότερο χρόνο εκπαίδευσης.

Στο Σενάριο 2, όπου γίνεται επιλογή χαρακτηριστικών παράλληλα με τις υπερπαραμέτρους, παρουσιάζεται σημαντική βελτίωση στις μετρικές Accuracy, F1-score και Recall. Το βέλτιστο μοντέλο συνολικά από τα μοντέλα που χρησιμοποιούν ANFIS είναι το Σενάριο 2 (Free 42). Το συγκεκριμένο μοντέλο πέτυχε το υψηλότερο F1-score (0,4556) μεταξύ των μοντέλων ANFIS, πλησιάζοντας αισθητά την τιμή F1-score του Random Forest. Παράλληλα, πέτυχε ένα εξαιρετικά κρίσιμο πλεονέκτημα

ασφαλείας, περιορίζοντας τις χαμένες επιθέσεις σε μόλις 14 στη βέλτιστη εκτέλεση, με τον μέσο όρο των πολλαπλών εκτελέσεων να αποτυπώνεται επίσης σε ιδιαίτερα ικανοποιητικά επίπεδα, όπως αναλύεται στην επόμενη ενότητα. Αυτή η επίδοση προήλθε από ένα μοντέλο που χρησιμοποιεί περίπου το 20% των αρχικά διαθέσιμων χαρακτηριστικών, με τον γενετικό αλγόριθμο να έχει επιλέξει τα σημαντικότερα. Επίσης, ο χρόνος εκπαίδευσης του τελικού ANFIS παρουσιάζει σημαντική μείωση σε σύγκριση με το 1^ο μοντέλο αναφοράς, μειώνοντάς τον από 4.215 δευτερόλεπτα (περίπου 1 ώρα και 10 λεπτά) σε μόλις 788 δευτερόλεπτα (περίπου 13 λεπτά).

Πίνακας 6.3: Υπερπαράμετροι ANFIS των πειραμάτων ταξινόμησης πολλαπλών κλάσεων

Ονομασία Setup	Προσέγγιση Χρωμοσώματος	Υπερπαράμετροι ANFIS				
		Clusters	m	Ρυθμός Μάθησης	Εποχές	Συνάρτηση Συμμετοχής
1^ο Μοντέλο Αναφοράς Κλασικό ANFIS	-	10	2,0	0,01	20	Gaussian
2^ο Μοντέλο Αναφοράς Random Forest	-	-	-	-	-	-
Σενάριο 1	-	6	2,18	0,016	15	Gaussian
Σενάριο 2 (Fixed 30)	Προκαθορισμένο Πλήθος 30	10	2,06	0,008	25	Triangular
Σενάριο 2 (Fixed 45)	Προκαθορισμένο Πλήθος 45	19	1,26	0,001	16	Gaussian
Σενάριο 2 (Fixed 60)	Προκαθορισμένο Πλήθος 60	12	1,54	0,001	30	Gaussian
Σενάριο 2 (Fixed 120)	Προκαθορισμένο Πλήθος 120	16	1,66	0,005	29	Triangular
Σενάριο 2 (Free 42)	Ελεύθερο Πλήθος 42	10	2,22	0,006	30	Triangular

Πίνακας 6.4: Αποτελέσματα ταξινόμησης πολλαπλών κλάσεων (βέλτιστες εκτελέσεις)

Setup	Accuracy	Macro F1-score	Macro Precision	Macro Recall	Χαμένες Επιθέσεις	Χρόνος Εκπαίδευσης
1^ο Μοντέλο Αναφοράς Κλασικό ANFIS	0,9618	0,4143	0,4147	0,6040	59	4215 sec
2^ο Μοντέλο Αναφοράς Random Forest	0,9681	0,4874	0,4626	0,6782	0	86 sec
Σενάριο 1	0,9621	0,4287	0,4287	0,6208	33	1339 sec
Σενάριο 2 (Fixed 30)	0,9596	0,4293	0,4196	0,6359	58	249 sec
Σενάριο 2 (Fixed 45)	0,9604	0,4548	0,4330	0,6639	22	2451 sec
Σενάριο 2 (Fixed 60)	0,9612	0,4416	0,4241	0,6537	113	2078 sec
Σενάριο 2 (Fixed 120)	0,9646	0,4455	0,4350	0,6311	28	1990 sec
Σενάριο 2 (Free 42)	0,9623	0,4556	0,4373	0,6559	14	788 sec

6.3.2 Ανάλυση επιλογής χαρακτηριστικών

Στο απαιτητικό πλαίσιο της ταξινόμησης πολλαπλών κλάσεων, η μεγάλη διαστατικότητα των δεδομένων αποτελεί πολύ σημαντικό παράγοντα για την απόδοση των μοντέλων. Η χρησιμοποίηση χαρακτηριστικών που περιλαμβάνουν είτε περιττές είτε θορυβώδεις πληροφορίες συχνά προκαλούν σύγχυση στη διαδικασία τελικής ταξινόμησης-απόφασης του μοντέλου. Αυτό το φαινόμενο είναι περισσότερο εμφανές στην ταξινόμηση πολλαπλών κλάσεων λόγω της αυξημένης πολυπλοκότητάς της.

Στην προσέγγιση ελεύθερης επιλογής χαρακτηριστικών, ο γενετικός αλγόριθμος επέλεξε 42 χαρακτηριστικά, με τα οποία κατέληξε να έχει την καλύτερη απόδοση μεταξύ των μοντέλων με ANFIS. Ο γενετικός αλγόριθμος επέλεξε αυτά τα 42 χαρακτηριστικά, απορρίπτοντας τα υπόλοιπα 160 και συμπιέζοντας έτσι την πληροφορία και τη διαστατικότητα. Ως αποτέλεσμα, το συγκεκριμένο μοντέλο κατάφερε να πετύχει την καλύτερη τιμή F1-score αλλά και τη μικρότερη τιμή χαμένων επιθέσεων ανάμεσα στα μοντέλα με ANFIS, αποτυγχάνοντας να αναγνωρίσει μόνο 14 από αυτές.

Τα αποτελέσματα της προσέγγισης προκαθορισμένου πλήθους χαρακτηριστικών και συγκεκριμένα το μοντέλο με τα 45 χαρακτηριστικά, είναι παρόμοια με αυτά της ελεύθερης επιλογής με 42 χαρακτηριστικά. Βέβαια, οι χαμένες επιθέσεις είναι ελαφρώς αυξημένες, ενώ και ο χρόνος εκπαίδευσης είναι σημαντικά μεγαλύτερος. Ο λόγος είναι τα περισσότερα clusters που χρησιμοποιεί το μοντέλο των προκαθορισμένων 45 χαρακτηριστικών σε σύγκριση με το ελεύθερο των 42 (19 έναντι 10 αντίστοιχα). Τα υπόλοιπα setup της προσέγγισης προκαθορισμένου πλήθους εμφανίζουν και αυτά ικανοποιητικά αποτελέσματα, ωστόσο δεν ξεπερνούν το setup των 45 χαρακτηριστικών. Συγκεκριμένα, το setup των 120 χαρακτηριστικών, παρότι έχει πρόσβαση σε περισσότερες πληροφορίες, σημείωσε χαμηλότερο F1-score από τα μοντέλα με 42 και 45 χαρακτηριστικά. Αντίθετα, το setup με πλήθος 30 χαρακτηριστικών σημείωσε αρκετά χαμηλότερο F1-score από τα υπόλοιπα του Σεναρίου 2, ενώ παράλληλα δεν απέχει πολύ από την απόδοση του Σεναρίου 1 και του 1^{ου} μοντέλου αναφοράς. Επομένως, το πλήθος των 30 χαρακτηριστικών φαίνεται να μην περιέχει αρκετή πληροφορία για τη σωστή ταξινόμηση πολλαπλών κλάσεων στο συγκεκριμένο πρόβλημα.

Το γεγονός ότι το μοντέλο με προκαθορισμένο πλήθος 45 χαρακτηριστικών σημείωσε εξίσου υψηλή απόδοση επιβεβαιώνει πως το εύρος [40, 45] αποτελεί το βέλτιστο πλήθος χαρακτηριστικών, το οποίο είναι αρκετό για να συμπεριλάβει την απαραίτητη πληροφορία που οδηγεί στη σωστή ταξινόμηση κλάσεων, χωρίς περιττό «θόρυβο». Συνοψίζοντας, ως βέλτιστο μοντέλο για την ταξινόμηση πολλαπλών κλάσεων, εκτός του Random Forest, κρίθηκε το Σενάριο 2 με την ελεύθερη επιλογή, τα 42 χαρακτηριστικά και τις υπερπαραμέτρους που εμφανίζονται στον Πίνακα 6.3, λόγω της υψηλής του απόδοσης και του μικρού πλήθους χαμένων επιθέσεων.

6.3.3 Ανάλυση σπάνιων επιθέσεων

Για την πλήρη κατανόηση της συμπεριφοράς του βέλτιστου μοντέλου (Σενάριο 2 – Free 42), παρουσιάζονται στον Πίνακα 6.5 οι μετρικές Precision, Recall, F1-score και το πλήθος των εγγραφών κάθε κλάσης (Total records). Τα αποτελέσματα αυτά προέρχονται από τη βέλτιστη εκτέλεση του τελικού ANFIS αυτού του μοντέλου και αποτελούν μια πιο λεπτομερή ανάλυση ανά κλάση των αντίστοιχων αποτελεσμάτων του Πίνακα 6.4.

Αναλύοντας τα αποτελέσματα, διαπιστώνεται αρχικά η εξαιρετική απόδοση του μοντέλου στις πλειοψηφικές κλάσεις. Η κλάση φυσιολογικής δικτυακής κίνησης (Normal) έχει σχεδόν άριστη απόδοση με την τιμή του Precision να είναι το απόλυτο 1, ενώ και η τιμή F1-score είναι εξαιρετικά κοντά στο 1. Η κλάση Generic αποτελεί τη συνηθέστερη κατηγορία επίθεσης και παρουσιάζει εξίσου

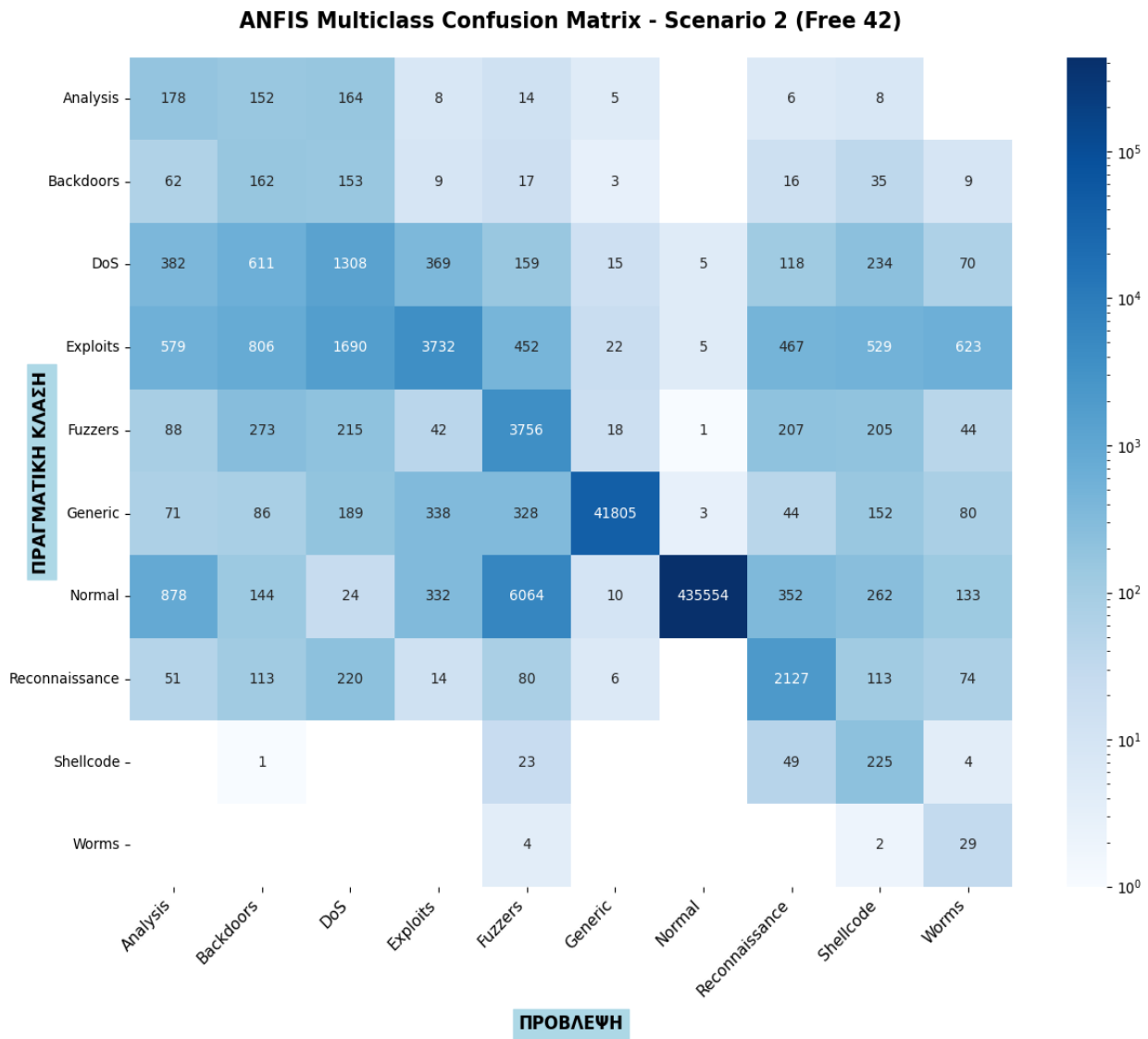
ικανοποιητικά αποτελέσματα. Οι επιδόσεις του μοντέλου σε αυτές τις δύο κλάσεις, οι οποίες αποτελούν περίπου το 95% του υποσυνόλου τελικού ελέγχου (Final Test set), δικαιολογούν την υψηλή τιμή του Accuracy (96,23%) του μοντέλου στον Πίνακα 6.4. Συνεπώς, το μοντέλο έχει αναγνωρίσει με επιτυχία τα κλασικά και πιο συνήθη μοτίβα της δικτυακής κίνησης.

Στις κλάσεις κατηγοριών επιθέσεων με μέση συχνότητα εμφάνισης, η συμπεριφορά του μοντέλου έχει αρκετές διακυμάνσεις. Η κλάση Reconnaissance, η οποία εμφανίζεται 2.798 φορές, παρουσιάζει πολύ καλή ισορροπία με το F1-score να έχει τιμή 0,6879. Οι κλάσεις Exploits και Fuzzers παρουσιάζουν ελαφρώς μειωμένες τιμές F1-score, όμως η καθεμία για διαφορετικό λόγο. Όπως είχε αναφερθεί, η μετρική F1-score υπολογίζεται με βάση το Precision και το Recall, δημιουργώντας μια μέση τιμή αναφοράς. Το μοντέλο είναι πιο «προσεκτικό» για την κλάση Exploits, καθώς εμφανίζει μεγάλο Precision (0,7704) και σχετικά μικρότερο Recall (0,4191). Αντίθετα, για την κλάση Fuzzers, το μοντέλο έχει πιο επιθετική συμπεριφορά, εντοπίζοντας τις περισσότερες πραγματικές εγγραφές της συγκεκριμένης κλάσης με υψηλό Recall (0,7746) αλλά με χαμηλότερη ακρίβεια (Precision 0,3447) και δημιουργώντας περισσότερες λανθασμένες ταξινομήσεις. Το πιο αδύναμο σημείο των μέτρια εμφανιζόμενων κλάσεων είναι η κλάση DoS, όπου το F1-score με τιμή 0,3616 είναι σχετικά χαμηλό.

Οι κλάσεις Analysis, Backdoors, Shellcode και Worms είναι οι πιο σπάνια εμφανιζόμενες. Όλες αυτές, εμφανίζουν πολύ χαμηλές τιμές F1-score (0,05 έως 0,21). Παρά τις χαμηλές τιμές F1-score, οι τιμές του Recall για τις κλάσεις Shellcode και Worms είναι εξαιρετικές. Το μοντέλο καταφέρνει να εντοπίσει το 74,5% των επιθέσεων Shellcode και το 82,86% των επιθέσεων Worms. Σε ρεαλιστικές συνθήκες, η ικανότητα του μοντέλου να ανιχνεύει σπάνιες κλάσεις, όπως Worms που έχει μόνο 35 εγγραφές, σε ένα σύνολο περίπου 500 χιλιάδων εγγραφών, είναι πολύ σημαντική. Το πολύ χαμηλό F1-score των Shellcode και Worms οφείλεται στις πολύ χαμηλές τιμές ακρίβειας (Precision), που σημαίνει ότι το μοντέλο ταξινομεί λανθασμένα άλλες κλάσεις ως Shellcode ή Worms αντίστοιχα. Οι κλάσεις Analysis και Backdoors παρουσίασαν γενικά χαμηλές μετρικές απόδοσης. Ωστόσο, αν και υπάρχουν αρκετές λανθασμένες ταξινομήσεις ανάμεσα σε κλάσεις επιθέσεων, η πλειοψηφία αυτών ταξινομείται ως άλλη κλάση επίθεσης χωρίς να υπάρχουν πολλές χαμένες επιθέσεις, οι οποίες στο συγκεκριμένο πείραμα ήταν μόνο 14. Στο Σχήμα 6.2 απεικονίζονται στο διάγραμμα σύγχυσης (Confusion Matrix) οι ταξινομήσεις του μοντέλου.

Πίνακας 6.5: Αναλυτικά στατιστικά αξιολόγησης για κάθε κλάση στο Σενάριο 2 (Free 42)

Κλάση	Precision	Recall	F1-score	Total records
Normal	1,0000	0,9815	0,9907	443.753
Analysis	0,0778	0,3327	0,1261	535
Backdoors	0,0690	0,3476	0,1151	466
DoS	0,3301	0,3999	0,3616	3.271
Exploits	0,7704	0,4191	0,5429	8.905
Fuzzers	0,3447	0,7746	0,4771	4.849
Generic	0,9981	0,9700	0,9839	43.096
Reconnaissance	0,6282	0,7602	0,6879	2.798
Shellcode	0,1275	0,7450	0,2177	302
Worms	0,0272	0,8286	0,0527	35



Σχήμα 6.2: Πίνακας σύγχυσης ταξινόμησης πολλαπλών κλάσεων

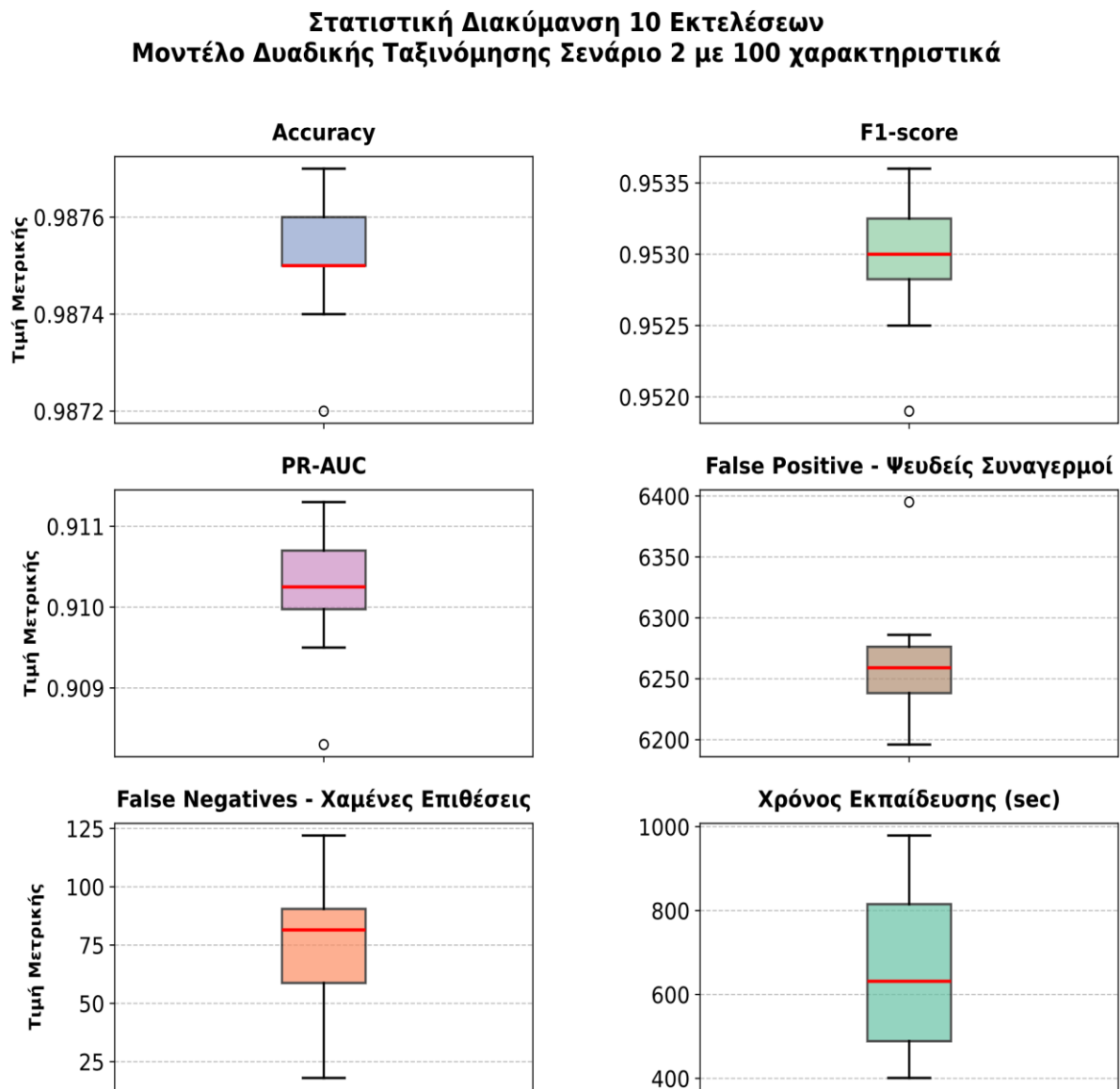
6.4 Στατιστική ανάλυση αποτελεσμάτων

6.4.1 Πολλαπλές εκτελέσεις πειραμάτων και περιγραφικά στατιστικά

Τα αποτελέσματα των βέλτιστων μοντέλων που επιλέχθηκαν για δυαδική ταξινόμηση και ταξινόμηση πολλαπλών κλάσεων, αλλά και όλων των άλλων πειραμάτων του Πίνακα 6.2 και Πίνακα 6.4 αντίστοιχα, όπως έχει ήδη αναφερθεί, προέρχονται από την καλύτερη απόδοση μίας συγκεκριμένης εκτέλεσης ανάμεσα σε 10 ανεξάρτητες, χρησιμοποιώντας διαφορετική τυχαία αρχικοποίηση (random seed). Η πραγματοποίηση των πολλαπλών εκτελέσεων έγινε για την απόδειξη πως τα αποτελέσματα της τελικής αξιολόγησης του κάθε μοντέλου δεν είναι τυχαία. Για τα μοντέλα των δύο διαφορετικών ταξινομήσεων πραγματοποιήθηκε στατιστική επεξεργασία των αποτελεσμάτων των 10 εκτελέσεων, με τα αποτελέσματα αυτής της επεξεργασίας να εμφανίζονται στον Πίνακα 6.6 και Πίνακα 6.7. Συγκεκριμένα, υπολογίστηκαν ο μέσος όρος (Average), η τυπική απόκλιση (Standard Deviation), η ελάχιστη (Minimum) και μέγιστη (Maximum) τιμή κάθε μετρικής.

Πίνακας 6.6: Αποτελέσματα πολλαπλών εκτελέσεων του πειράματος δυαδικής ταξινόμησης
Σενάριο 2 - 100 χαρακτηριστικών

	Accuracy	F1-score	PR-AUC	False Positives – Ψευδείς Συναγερμοί	False Negatives - Χαμένες Επιθέσεις	Χρόνος Εκπαίδευσης
Average	0,9875	0,9529	0,9102	6263,5	75,5	661,2 sec
Standard Deviation	0,00014	0,00047	0,00083	54,5	30,1	205,5 sec
Minimum	0,9872	0,9519	0,9083	6196	18	401 sec
Maximum	0,9877	0,9536	0,9113	6395	122	979 sec



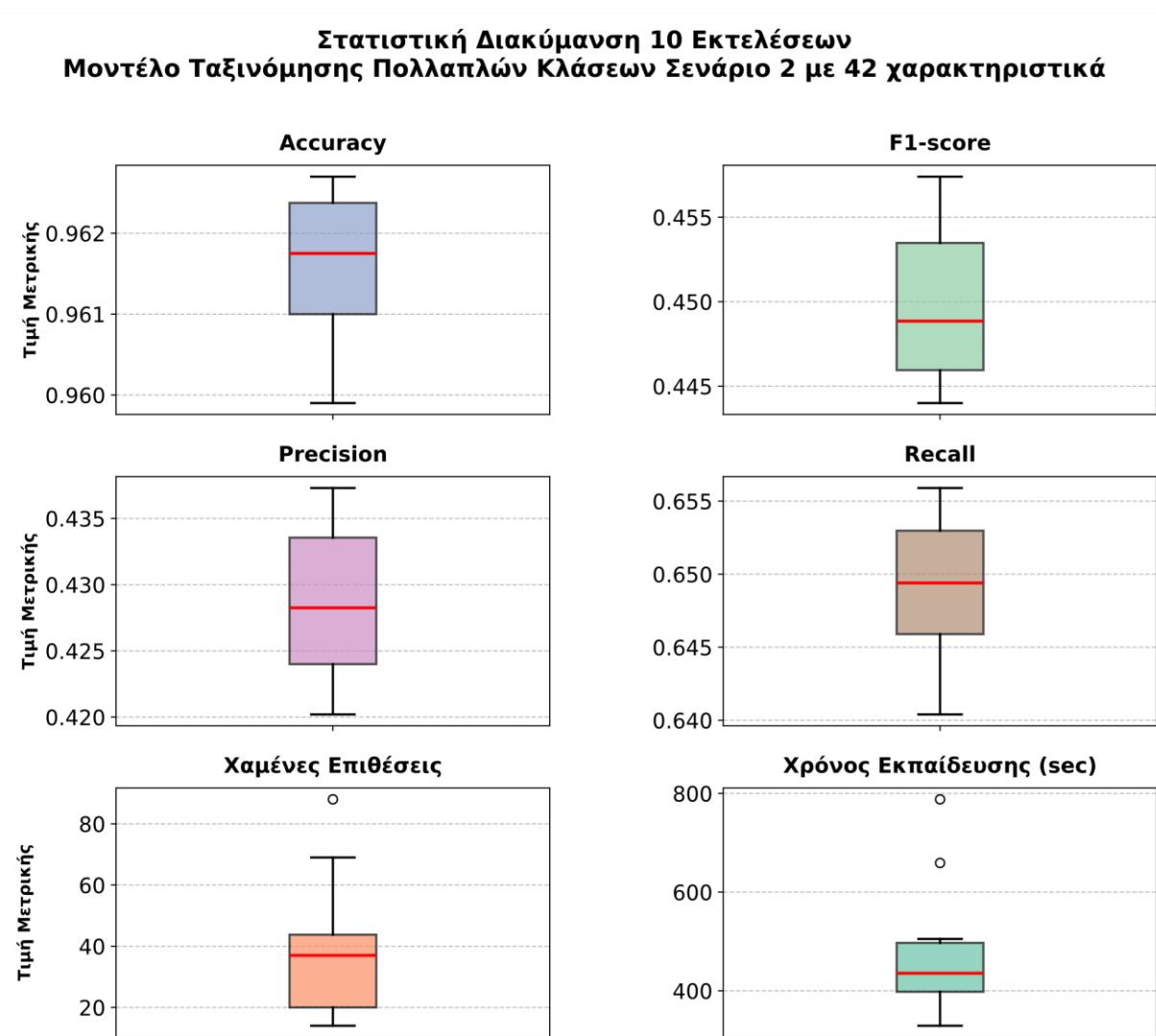
Σχήμα 6.3: Στατιστική διακύμανση εκτελέσεων του μοντέλου για δυαδική ταξινόμηση

Τα αποτελέσματα του Πίνακα 6.6 αποτυπώνονται και στο Σχήμα 6.3, στο οποίο χρησιμοποιείται θηκογράμματα για κάθε μετρική του πίνακα. Όπως δείχνουν τα αποτελέσματα για το συγκεκριμένο μοντέλο δυαδικής ταξινόμησης, σε γενικές γραμμές η διακύμανση των τιμών είναι εξαιρετικά μικρή. Το Accuracy, το F1-score και το PR-AUC εμφανίζουν τόσο μικρή τυπική απόκλιση που η διαφορά βρίσκεται στα τρίτα δεκαδικά ψηφία. Αυτή η σταθερότητα των τιμών απεικονίζεται και στα αντίστοιχα θηκογράμματα των μετρικών. Όσον αφορά τις χαμένες επιθέσεις, αν και έχουν μεγάλο μέσο όρο (75,5) σε αναλογία με το εύρος τιμών τους, η ελάχιστη τιμή των μόλις 18 χαμένων επιθέσεων δείχνει τις πραγματικές δυνατότητες του μοντέλου. Οι ψευδείς συναγερμοί έχουν μικρή τυπική απόκλιση, κρατώντας σταθερές τιμές. Ο χρόνος εκπαίδευσης αν και εμφανίζει αρκετές διακυμάνσεις, έχει μέσο όρο 661,2 δευτερόλεπτα, ενώ οι μεγάλες αποκλίσεις οφείλονται στον πρόωρο τερματισμό της εκπαίδευσης για την πρόληψη υπερεκπαίδευσης. Αυτή η λειτουργία του early stopping υποδεικνύει ότι ο γενετικός αλγόριθμος έχει αναζητήσει τα όρια των βέλτιστων υπερπαραμέτρων καθώς και τα σωστά χαρακτηριστικά. Επομένως, η διαδικασία των πολλαπλών εκτελέσεων αποδεικνύει τη συνολική σταθερότητα του συγκεκριμένου μοντέλου για τη δυαδική ταξινόμηση.

Αντίστοιχα, και για το επιλεγμένο μοντέλο Σενάριο 2 (42 χαρακτηριστικών) ταξινόμησης πολλαπλών κλάσεων, πραγματοποιούνται 10 εκτελέσεις του τελικού ANFIS αλλάζοντας σε κάθε εκτέλεση το random seed, το οποίο μεταβάλλει την τυχαία αρχικοποίηση του FCM αλλά και την εισαγωγή δεδομένων στην εκπαίδευση. Τα συνοπτικά αποτελέσματα των 10 εκτελέσεων του μοντέλου για την πολύπλοκη ταξινόμηση πολλαπλών κλάσεων εμφανίζονται παρακάτω στον Πίνακα 6.7. Το συγκεκριμένο μοντέλο είναι περισσότερο σταθερό από το αντίστοιχο για τη δυαδική ταξινόμηση. Συγκεκριμένα, το Accuracy, F1-score, Precision και Recall έχουν εξαιρετικά χαμηλές τιμές τυπικής απόκλισης, ενώ οι μέσοι όροι αυτών παραμένουν σε ικανοποιητικά υψηλά επίπεδα. Το γεγονός ότι η μέγιστη τιμή των χαμένων επιθέσεων είναι 88, υποδηλώνει ότι το μοντέλο παραμένει εξαιρετικά ευαίσθητο στην ανίχνευση πραγματικών επιθέσεων. Η μέση τιμή χαμένων επιθέσεων παραμένει ικανοποιητικά χαμηλή με τιμή 38,7 και η ελάχιστη είναι στις 14, ενώ και η τυπική απόκλιση είναι σχετικά μικρή για το εύρος τιμών. Ο χρόνος εκπαίδευσης, όπως και στο μοντέλο δυαδικής ταξινόμησης, έχει κάποιες μικρές μεταβολές οι οποίες οφείλονται στη λειτουργία πρόωρου τερματισμού για την αποφυγή υπερεκπαίδευσης. Τέλος, στο Σχήμα 6.4, τα θηκογράμματα απεικονίζουν τη συνολική σταθερότητα και αξιοπιστία του συγκεκριμένου μοντέλου ταξινόμησης πολλαπλών κλάσεων. Συνοψίζοντας, και τα δύο μοντέλα ταξινόμησης που επιλέχθηκαν ως βέλτιστα, αποδεικνύονται αξιόπιστα λόγω της σταθερότητας που εμφανίζουν τα αποτελέσματα των πολλαπλών εκτελέσεών τους.

Πίνακας 6.7: Αποτελέσματα πολλαπλών εκτελέσεων του πειράματος ταξινόμησης πολλαπλών κλάσεων Σενάριο 2 - 42 χαρακτηριστικών

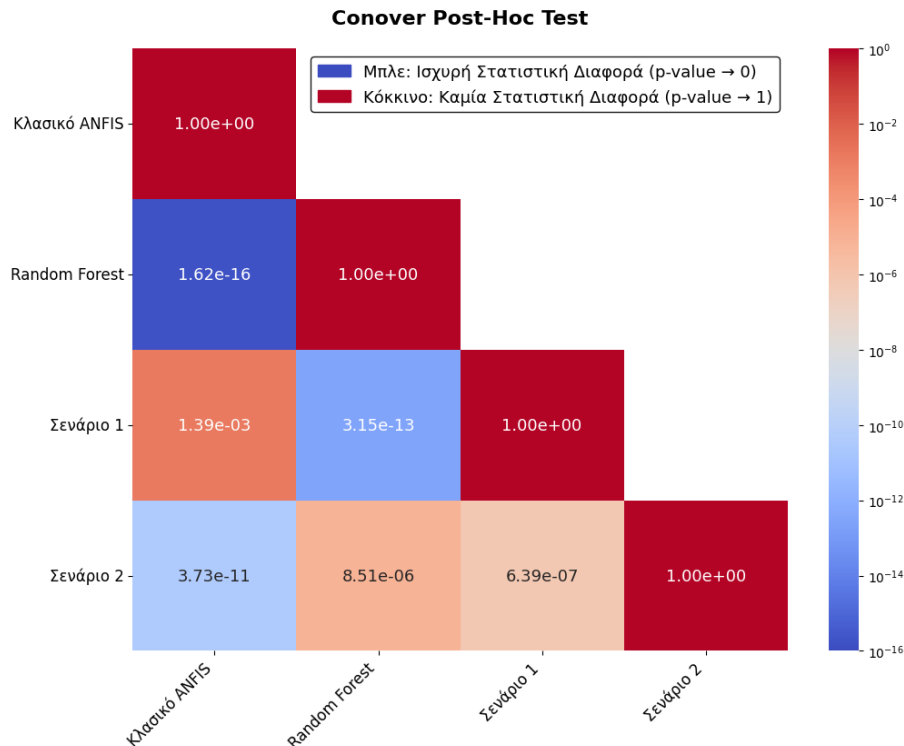
	Accuracy	Macro F1-score	Macro Precision	Macro Recall	Χαμένες Επιθέσεις	Χρόνος Εκπαίδευσης
Average	0,9616	0,4499	0,4284	0,649	38,7	481,1 sec
Standard Deviation	0,0009	0,0047	0,006	0,05	24,3	140,7 sec
Minimum	0,9599	0,444	0,4202	0,6404	14	329 sec
Maximum	0,9627	0,4574	0,4373	0,6559	88	788 sec



Σχήμα 6.4 : Στατιστική διακύμανση εκτελέσεων του μοντέλου για ταξινόμηση πολλαπλών κλάσεων

6.4.2 Έλεγχος Friedman και μεταγενέστεροι στατιστικού έλεγχοι

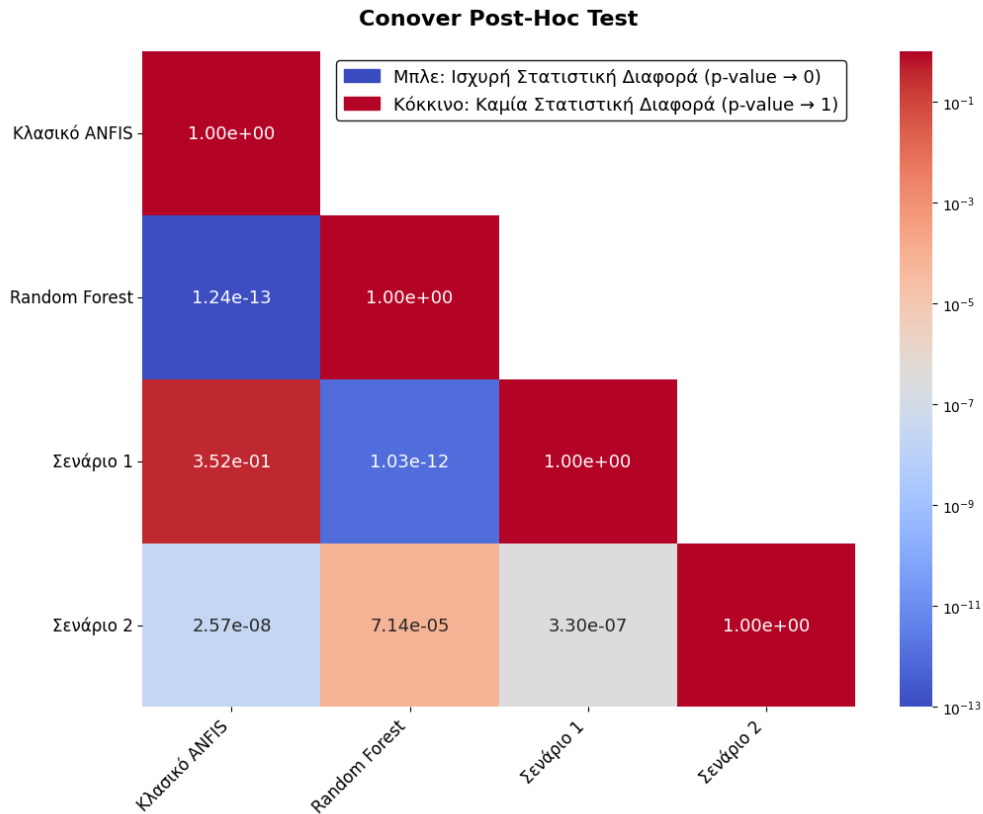
Προκειμένου να τεκμηριωθούν στατιστικά οι διαφορές στην απόδοση μεταξύ των μοντέλων του πειραματικού πρωτοκόλλου και να αποδειχθεί πως τα αποτελέσματα δεν οφείλονται σε τυχαίους παράγοντες, διεξήχθη αρχικά ο μη παραμετρικός στατιστικός έλεγχος Friedman. Για τον έλεγχο, χρησιμοποιήθηκαν τα αποτελέσματα των 10 ανεξάρτητων εκτελέσεων (συγκεκριμένα οι τιμές του F1-score) για τα 4 συγκρινόμενα μοντέλα (1ο Μοντέλο Αναφοράς, 2ο Μοντέλο Αναφοράς, Σενάριο 1, Σενάριο 2). Υπενθυμίζεται πως το βέλτιστο μοντέλο του Σεναρίου 2 δυαδικής ταξινόμησης είναι αυτό της ελεύθερης επιλογής με 100 χαρακτηριστικά, ενώ για την ταξινόμηση πολλαπλών κλάσεων το βέλτιστο μοντέλο για το Σενάριο 2 είναι εκείνο της ελεύθερης επιλογής με 42 χαρακτηριστικά και τις αντίστοιχες υπερπαραμέτρους που εμφανίζονται στον Πίνακα 6.1 και Πίνακα 6.3 αντίστοιχα. Τα αποτελέσματα του ελέγχου Friedman επιβεβαίωσαν την ύπαρξη στατιστικά σημαντικών διαφορών μεταξύ των μοντέλων ($p < 0,05$) και για τις δύο ταξινομήσεις. Ωστόσο, ο συγκεκριμένος έλεγχος αναδεικνύει ότι υπάρχει γενικά διαφορά στα αποτελέσματα των μοντέλων χωρίς να κάνει σύγκριση μεταξύ αυτών.



Σχήμα 6.5 : Post-Hoc έλεγχος Conover στη δυαδική ταξινόμηση

Η απόδειξη της ύπαρξης σημαντικής διαφοράς με τον έλεγχο Friedman επιτρέπει την περαιτέρω διερεύνηση μέσω μεταγενέστερων (Post-Hoc) ελέγχων ανά ζεύγη μοντέλων. Για αυτό τον σκοπό, χρησιμοποιήθηκε ο Post-Hoc έλεγχος Conover, με τον οποίο ελέγχεται η διαφορά ανάμεσα σε ζευγάρια μοντέλων. Σημειώνεται ότι κατά τη διεξαγωγή του ελέγχου Conover δεν εφαρμόστηκε κάποια επιπλέον διόρθωση για πολλαπλές συγκρίσεις (όπως π.χ. διόρθωση Holm ή Bonferroni). Ο έλεγχος Nemenyi δεν χρησιμοποιήθηκε, καθώς παρέχει χαμηλή στατιστική ισχύ [52]. Με τον έλεγχο Conover αποδεικνύεται η στατιστική διαφορά μεταξύ δύο μοντέλων, όμως ο έλεγχος δεν αναδεικνύει ποιο μοντέλο από τα δύο είναι ανώτερο. Αυτό το συμπέρασμα, μπορεί να εκφραστεί με συνδυασμό των αποτελεσμάτων απόδοσης των μοντέλων από τους προγενέστερους πίνακες.

Στο Σχήμα 6.5 εμφανίζονται τα αποτελέσματα του ελέγχου Conover για τη δυαδική ταξινόμηση. Στον θερμικό χάρτη (heatmap), παρουσιάζονται οι τιμές σημαντικότητας (p-values) για κάθε πιθανή σύγκριση μεταξύ των μοντέλων. Αρχικά, το γεγονός πως καμία σύγκριση δεν έχει τιμή p-value μεγαλύτερη από το όριο 0,05, το οποίο αποτελεί το επίπεδο σημαντικότητας (α), είναι σημαντικό εύρημα ισχυρής στατιστικής διαφοράς ανάμεσα στα μοντέλα. Τα ζεύγη του Random Forest με το βασικό ANFIS αλλά και με το Σενάριο 1, εμφανίζουν μεγάλες στατιστικές διαφορές (μπλε χρώμα). Η διαφορά μεταξύ αυτών των ζευγαριών σύμφωνα με προηγούμενα αποτελέσματα απόδοσης, υποδεικνύει την υπεροχή του μοντέλου Random Forest. Το Σενάριο 1 παρουσιάζει μικρή στατιστική διαφορά με το Κλασικό ANFIS, η οποία εμφανίστηκε και στα αποτελέσματα του Πίνακα 6.2. Εστιάζοντας στο προτεινόμενο μοντέλο (Σενάριο 2 - 100 χαρακτηριστικών), το χαμηλό p-value στη σύγκρισή του με το βασικό ANFIS αλλά και η ελαφρώς μικρότερη στατιστική διαφορά (μεγαλύτερο p-value) με το Σενάριο 1, επιβεβαιώνουν την ανωτερότητα του προτεινόμενου μοντέλου ανάμεσα στα υπόλοιπα που χρησιμοποιούν ANFIS. Στο ζευγάρι Random Forest με Σενάριο 2 υπάρχει μεγάλη στατιστική διαφορά, η οποία εκφράζει την καλύτερη απόδοση του Random Forest. Ωστόσο, αυτή η διαφορά Random Forest–Σενάριο 2 δεν είναι τόσο ισχυρή όσο των υπόλοιπων ζευγών του Σεναρίου 2, άρα η απόδοση του Σεναρίου 2 πλησιάζει αυτή του Random Forest.



Σχήμα 6.6 : Post-Hoc έλεγχος Conover στην ταξινόμηση πολλαπλών κλάσεων

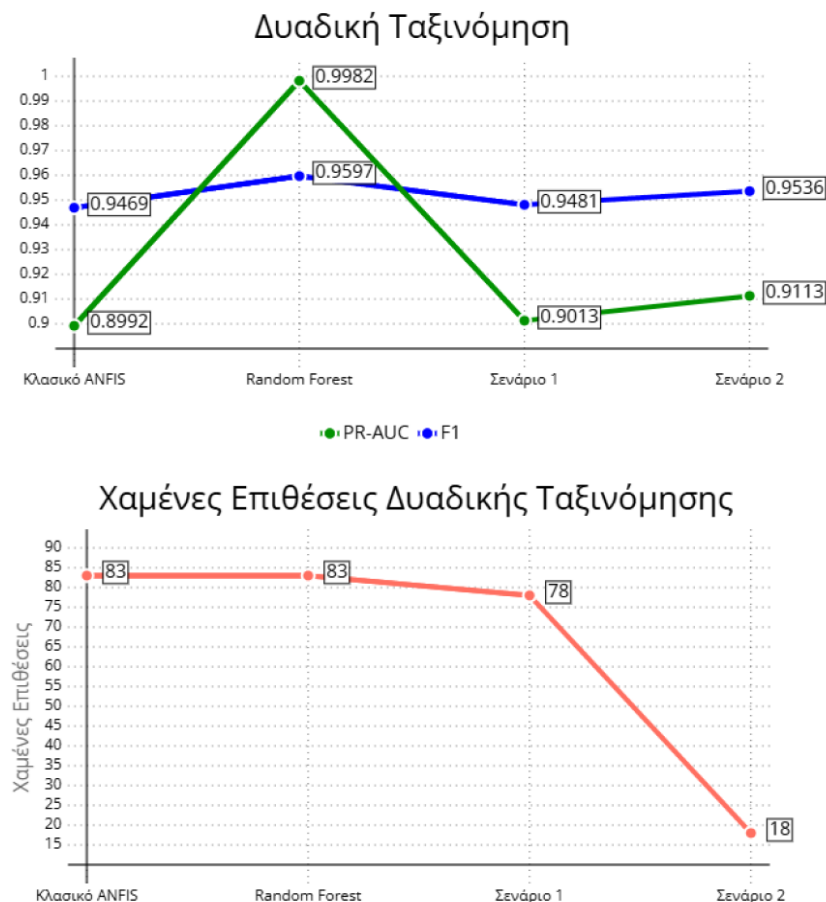
Τα αποτελέσματα του ελέγχου Conover για την ταξινόμηση πολλαπλών κλάσεων παρουσιάζονται στο Σχήμα 6.6, στο οποίο εμφανίζονται και οι αντίστοιχες τιμές των p-values. Η πολυπλοκότητα ταξινόμησης πολλαπλών κλάσεων, απεικονίζεται στις αρκετά μικρότερες τιμές των p-values των ζευγών Random Forest με το Κλασικό ANFIS και Σενάριο 1, οι οποίες πλησιάζουν περισσότερο το 0 και υποδηλώνουν τη μεγαλύτερη υπεροχή του Random Forest. Το Σενάριο 1 εμφανίζει πολύ υψηλή τιμή p-value ($0,352 > \alpha=0,05$) στη σύγκρισή του με το βασικό ANFIS, το οποίο σημαίνει πως μόνο η βελτιστοποίηση των υπερπαραμέτρων ANFIS του Σεναρίου 1 δεν είναι αρκετή για την απόδειξη της μη-τυχαίας βελτίωσης απόδοσης του μοντέλου έναντι του Κλασικού ANFIS μοντέλου αναφοράς. Το Σενάριο 2, όπως και στον αντίστοιχο έλεγχο για τη δυαδική ταξινόμηση, παρουσιάζει ισχυρές στατιστικές διαφορές στα ζεύγη με το βασικό ANFIS και Σενάριο 1. Το ζεύγος Σεναρίου 2 και Random Forest έχει επίσης σημαντική στατιστική διαφορά υποδεικνύοντας την ανώτερη απόδοση του Random Forest. Ουσιαστικά, το Σενάριο 2 είναι αποδεδειγμένα το ανώτερο μοντέλο μεταξύ των μοντέλων με ANFIS.

Συνοψίζοντας, ο έλεγχος Friedman έδειξε ότι υπάρχουν στατιστικά σημαντικές διαφορές μεταξύ των συγκρινόμενων προσεγγίσεων, ενώ οι μεταγενέστερες συγκρίσεις Conover ανέδειξαν τα ζεύγη μοντέλων μεταξύ των οποίων εμφανίζονται οι διαφορές. Στη σύγκριση του Σεναρίου 2 με το Random Forest, η ανωτερότητα του τελευταίου είναι γνωστή όμως όπως έχει αναφερθεί στερείται την ερμηνευσιμότητα που προσφέρει το ANFIS.

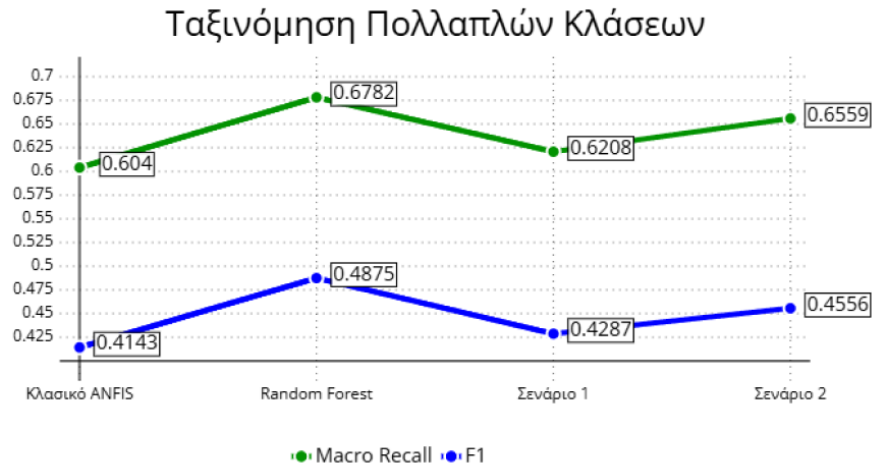
6.4.3 Σύγκριση αποτελεσμάτων δυαδικής ταξινόμησης και ταξινόμησης πολλαπλών κλάσεων

Έχοντας πλέον ορίσει το βέλτιστο μοντέλο και για τις δύο ταξινομήσεις, το οποίο είναι το Σενάριο 2 με την ελεύθερη επιλογή χαρακτηριστικών και βελτιστοποίηση υπερπαραμέτρων του ANFIS, ακολουθεί η σύγκριση μεταξύ των δύο μοντέλων για τις διαφορετικές ταξινομήσεις. Η συνολική οπτική απεικόνιση της σύγκρισης των επιδόσεων του Σεναρίου 2 έναντι των υπολοίπων μοντέλων της μελέτης παρουσιάζεται συγκεντρωτικά στο Σχήμα 6.7 για τη δυαδική ταξινόμηση και στο Σχήμα 6.8 για την ταξινόμηση πολλαπλών κλάσεων.

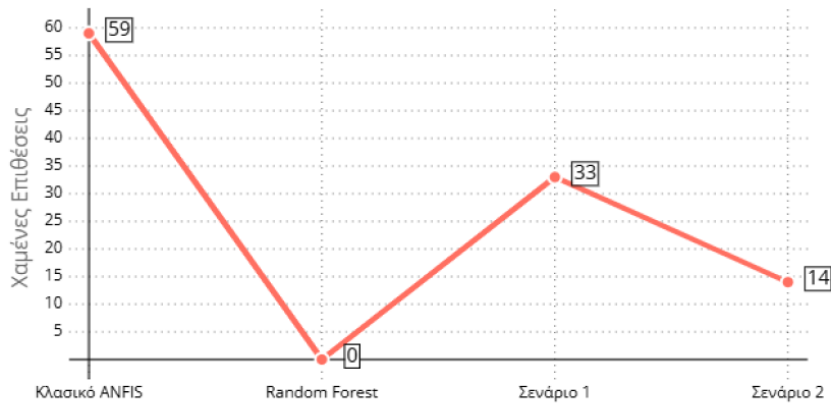
Αρχικά η μεγαλύτερη διαφορά στα αποτελέσματα των δύο ταξινομήσεων, εμφανίζεται στο F1-score. Συγκεκριμένα το Σενάριο 2 με την ελεύθερη επιλογή των 100 χαρακτηριστικών της δυαδικής ταξινόμησης, έχει σημαντικά μεγαλύτερη τιμή F1-score από το Σενάριο 2 ελεύθερης επιλογής των 42 χαρακτηριστικών. Για τη δυαδική ταξινόμηση η τιμή είναι 0,9536, ενώ για την ταξινόμηση πολλαπλών κλάσεων είναι 0,4536. Η διαφορά μεταξύ των δύο τιμών οφείλεται στη διαφορετική πολυπλοκότητα και αρχιτεκτονική κάθε ταξινόμησης. Η δυαδική ταξινόμηση καλείται να ταξινομήσει δύο μόνο διαφορετικές κλάσεις. Από την άλλη, στην ταξινόμηση πολλαπλών κλάσεων, το μοντέλο καλείται να ταξινομήσει 10 διαφορετικές κλάσεις που μπορεί να έχουν παρόμοια χαρακτηριστικά. Επομένως, αυτή η διαφορά 0,5 μονάδων στις τιμές του F1-score είναι δικαιολογημένη, λόγω της αυξημένης πολυπλοκότητας της ταξινόμησης πολλαπλών κλάσεων, χωρίς να σημαίνει πως η χαμηλότερη τιμή F1-score στη ταξινόμηση πολλαπλών κλάσεων έχει συνολικά χαμηλότερη απόδοση από τη δυαδική.



Σχήμα 6.7: Γενική σύγκριση αποτελεσμάτων δυαδικής ταξινόμησης



Χαμένες Επιθέσεις Ταξινόμησης Πολλαπλών Κλάσεων



Σχήμα 6.8: Γενική σύγκριση αποτελεσμάτων ταξινόμησης πολλαπλών κλάσεων

Σημαντικό εύρημα από τα δύο μοντέλα ταξινόμησης, αποτελεί η ικανότητα αναγνώρισης των επιθέσεων και συγκεκριμένα ο χαμηλός αριθμός των χαμένων επιθέσεων. Για τη δυαδική ταξινόμηση η μέση τιμή των χαμένων επιθέσεων από τις 10 εκτελέσεις, που εμφανίζεται και στον Πίνακα 6.6, είναι 75,5 και η ελάχιστη είναι 18, ξεπερνώντας ακόμα και την αντίστοιχη τιμή που είχε το Random Forest στις 83 χαμένες επιθέσεις. Η περισσότερο πολύπλοκη ταξινόμηση πολλαπλών κλάσεων, εμφανίζει μέση τιμή 38,7 και ελάχιστη τιμή 14 χαμένων επιθέσεων. Αυτή η φαινομενικά παράδοξη συμπεριφορά όπου στο πολύπλοκότερο πρόβλημα εμφανίζονται καλύτερα αποτελέσματα τεκμηριώνεται από τη διαφορετική προεπεξεργασία των δεδομένων για τις δύο ταξινομήσεις. Όπως έχει αναλυθεί και στην αντίστοιχη ενότητα του Κεφαλαίου 4, για τη δυαδική ταξινόμηση, τα δεδομένα του υποσυνόλου εκπαίδευσης έχουν χωριστεί σε 50% εγγραφές φυσιολογικής δικτυακής κίνησης και 50% εγγραφές που αντιστοιχούν σε επιθέσεις. Ωστόσο, στο κομμάτι που αντιστοιχεί στις επιθέσεις υπάρχει ανισορροπία κλάσεων επιθέσεων, καθώς ο στόχος της συγκεκριμένης ταξινόμησης είναι ο δυαδικός διαχωρισμός σε φυσιολογική εγγραφή ή επίθεση. Αντίθετα, για την ταξινόμηση πολλαπλών κλάσεων, το υποσύνολο εκπαίδευσης αποτελείται από εγγραφές όπου κάθε κλάση καταλαμβάνει το 10% του συνολικού υποσυνόλου. Έτσι οι 9 κλάσεις επίθεσης και 1 κλάση φυσιολογικής δικτυακής κίνησης αποτελούν το ισορροπημένο υποσύνολο δεδομένων εκπαίδευσης για το μοντέλο ταξινόμησης πολλαπλών κλάσεων. Τα ισορροπημένα δεδομένα επέτρεψαν στο μοντέλο να εκπαιδευτεί και να αναγνωρίζει με μεγαλύτερη επιτυχία τις σπανιότερα εμφανιζόμενες κλάσεις.

Τέλος, και στις δύο ταξινομήσεις η εφαρμογή πρόωρου τερματισμού (early stopping) περιορίσε τον κίνδυνο της υπερεκπαίδευσης, ενώ παράλληλα η αξιολόγηση του μοντέλου με τα αντίστοιχα υποσύνολα απέδωσε απολύτως ικανοποιητικά αποτελέσματα, αποδεικνύοντας την ικανότητα γενίκευσης απέναντι σε άγνωστα δεδομένα. Επίσης, στο βέλτιστο μοντέλο ταξινόμησης πολλαπλών κλάσεων ο γενετικός αλγόριθμος κατέληξε σε 10 clusters-ασαφείς κανόνες, ενώ για τη δυαδική ταξινόμηση σε 5 clusters. Ωστόσο, με λιγότερα clusters η δυαδική ταξινόμηση έχει ελαφρώς αυξημένο χρόνο εκπαίδευσης από την ταξινόμηση πολλαπλών κλάσεων, με το γεγονός αυτό να οφείλεται στη διαφορά του πλήθους χαρακτηριστικών που χρησιμοποιούν τα δύο. Το δυαδικό μοντέλο χρησιμοποιεί 100 χαρακτηριστικά, ενώ αντίθετα το μοντέλο ταξινόμησης πολλαπλών κλάσεων χρησιμοποιεί 42 χαρακτηριστικά. Η διαφορά των 58 χαρακτηριστικών έχει σημαντικό αντίκτυπο στο υπολογιστικό κόστος για την εκπαίδευση του ANFIS. Επίσης, στην εξελικτική διαδικασία και των δύο ταξινομήσεων υπήρξε εναλλαγή ανάμεσα σε Τριγωνικές και Γκαουσιανές συναρτήσεις συμμετοχής. Όμως, το πλεονέκτημα το έχουν οι Τριγωνικές, με αυτές να επιλέγονται στα πειράματα με τις βέλτιστες αποδόσεις του Σεναρίου 2. Η Τριγωνική συνάρτηση συμμετοχής προσφέρει πιο ξεκάθαρα όρια από τη Γκαουσιανή, γεγονός το οποίο φαίνεται να έχει σημαντικό ρόλο.

6.5 Αξιολόγηση και επιλογή τελικού μοντέλου

Με τη σύνοψη όλης της εκτενούς πειραματικής διαδικασίας και την ανάλυση των αποτελεσμάτων που παρουσιάζονται στις προηγούμενες ενότητες, αποδεικνύεται πως μεταξύ των προσεγγίσεων που βασίζονται στο ANFIS, το Σενάριο 2 με ελεύθερη επιλογή χαρακτηριστικών παρουσίασε την καλύτερη συνολική ισορροπία μεταξύ προβλεπτικής απόδοσης, μείωσης διαστατικότητας και ερμηνευσιμότητας. Το Random Forest διατήρησε υψηλότερη συνολική προβλεπτική απόδοση και μικρότερο χρόνο εκπαίδευσης. Η επιλογή της συγκεκριμένης υβριδικής προσέγγισης τεκμηριώνεται από τα εξής πέντε βασικά σημεία:

- **Απόδοση και Αξιοπιστία (Performance):** Η αρχιτεκτονική του συγκεκριμένου Σεναρίου 2 σημείωσε τις υψηλότερες τιμές F1-score ανάμεσα σε όλα τα πειραματικά μοντέλα στα οποία χρησιμοποιήθηκε το ANFIS. Επίσης, σημαντικό επίτευγμα είναι η δραστική μείωση των χαμένων επιθέσεων (False Negatives). Σε ένα σύστημα IDS, ο εντοπισμός του μέγιστου δυνατού αριθμού πραγματικών απειλών αποτελεί προτεραιότητα, με τα μοντέλα της συγκεκριμένης αρχιτεκτονικής να ανταποκρίνονται με απόλυτη επιτυχία.
- **Στατιστική Συμπεριφορά:** Όπως αναλύθηκε, η ανωτερότητα των μοντέλων δεν προέκυψε από αποτελέσματα τυχαίων διακυμάνσεων κατά την εκπαίδευση. Η πραγματοποίηση 10 διαφορετικών εκτελέσεων για κάθε πείραμα και των δύο ταξινομήσεων, απέδειξε τη σταθερότητα των αποτελεσμάτων. Με την εφαρμογή των ελέγχων Friedman και των μεταγενέστερων (post-hoc) ελέγχων Conover, το Σενάριο 2 παρουσίασε ισχυρή στατιστική διαφορά και καθαρή υπεροχή απέναντι στο βασικό μοντέλο αναφοράς ANFIS καθώς και απέναντι στο Σενάριο 1.
- **Μείωση διαστατικότητας:** Η αρχιτεκτονική με τον γενετικό αλγόριθμο και την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών απέδωσε τα καλύτερα αποτελέσματα. Η μείωση των χαρακτηριστικών σε 100 για τη δυαδική ταξινόμηση και 42 για την ταξινόμηση πολλαπλών κλάσεων, απέδειξε την ικανότητα του προτεινόμενου μοντέλου να χρησιμοποιεί μόνο τα χαρακτηριστικά με τις απαραίτητες πληροφορίες που οδηγούν σε σωστές προβλέψεις.
- **Υπολογιστική πολυπλοκότητα:** Η μείωση των χαρακτηριστικών εισόδου του ANFIS μειώνει τη διαστατικότητα των δεδομένων εκπαίδευσης και επιτρέπει τη γρηγορότερη και με χαμηλότερο υπολογιστικό κόστος εκπαίδευση του μοντέλου. Αυτή η μείωση σε συνδυασμό με την εφαρμογή πρόωρου τερματισμού εκπαίδευσης διατήρησε την υπολογιστική πολυπλοκότητα σε βιώσιμο επίπεδο, το οποίο αναδεικνύεται και από τον σχετικά χαμηλότερο χρόνο εκπαίδευσης.

- **Ερμηνευσιμότητα:** Η ικανότητα ερμηνείας των αποφάσεων του μοντέλου αποτελεί τον σημαντικότερο παράγοντα για την τελική επιλογή του Σεναρίου 2 ως το βέλτιστο μοντέλο. Στη συγκριτική αξιολόγηση των αποτελεσμάτων και των δύο ταξινομήσεων, το Random Forest παρουσίασε μεγαλύτερη γενική απόδοση. Ωστόσο, το Random Forest, αν και αποτελεί ένα ισχυρότατο μοντέλο αναφοράς, διαθέτει περιορισμένη εγγενή ερμηνευσιμότητα και η εσωτερική διαδικασία λήψης αποφάσεων του απαιτεί εξειδικευμένες τεχνικές (όπως SHAP, LIME, ανάλυση σημαντικότητας χαρακτηριστικών) για να ερμηνευθεί. Στον τομέα της κυβερνοασφάλειας, η ανάδειξη ενός συναγερμού δεν είναι αρκετή. Η επαλήθευση του συναγερμού αυτού πρέπει πολλές φορές να γίνεται και από ορισμένους αναλυτές, οι οποίοι πρέπει να γνωρίζουν τη μεθοδολογία που χρησιμοποίησε το σύστημα για να παραγάγει τον συναγερμό. Το προτεινόμενο μοντέλο με τη χρήση του ANFIS προσφέρει ένα ενδιαφέρον και ισορροπημένο πλαίσιο, το οποίο συνδυάζει τη στατιστικά τεκμηριωμένη ανταγωνιστική απόδοση με την ικανότητα ερμηνείας των ασαφών κανόνων, παρέχοντας διαφάνεια στη λειτουργία του μοντέλου.

6.6 Ερμηνεία και ανάλυση των κανόνων ANFIS

Η συγκεκριμένη ενότητα εστιάζει στην ικανότητα ερμηνείας της λειτουργίας του ANFIS για τη λήψη των αποφάσεών του. Η ανάλυση που ακολουθεί βασίζεται στους κανόνες που παρήχθησαν από το μοντέλο Σενάριο 2 με την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών της ταξινόμησης πολλαπλών κλάσεων, δηλαδή το βέλτιστο τελικό μοντέλο που αναδείχθηκε στην προηγούμενη ενότητα.

6.6.1 Ανάλυση Ασαφών Κανόνων

Η αρχιτεκτονική του ANFIS μπορεί να εξάγει τη «γνώση» που απέκτησε κατά την εκπαίδευση με τη μορφή ασαφών κανόνων. Κάθε τέτοιος κανόνας αποτελείται από το υποθετικό τμήμα (IF), το οποίο περιλαμβάνει συνδυασμό συγκεκριμένων συνθηκών με τις τιμές των χαρακτηριστικών εισόδου του συστήματος, και το τμήμα που αντιστοιχεί στο συμπέρασμα (THEN), το οποίο υποδεικνύει πως το μοτίβο συμπεριφοράς του υποθετικού τμήματος έχει μεγάλη δυνατότητα να αντιστοιχεί σε μια συγκεκριμένη κλάση.

Για τη σωστή ερμηνεία των τιμών του υποθετικού τμήματος, είναι σημαντικό να τονιστεί ο τρόπος προεπεξεργασίας των δεδομένων. Όπως έχει αναλυθεί στο Κεφάλαιο 4, τα δεδομένα εισόδου έχουν υποστεί κανονικοποίηση με την τεχνική Z-score. Συνεπώς, οι τιμές στους κανόνες δεν αντιστοιχούν στις πραγματικές μονάδες μέτρησης των χαρακτηριστικών, αλλά εκφράζουν τυπικές αποκλίσεις από τον μέσο όρο του συνόλου κάθε ξεχωριστού χαρακτηριστικού. Η τιμή 0 αποτελεί τον μέσο όρο, ενώ μια θετική τιμή υποδεικνύει ότι αυτή βρίσκεται πάνω από τον μέσο όρο και αντίστοιχα μια αρνητική τιμή βρίσκεται κάτω από αυτόν. Αυτές οι κανονικοποιημένες τιμές αποτελούν το πεδίο στο οποίο ορίζονται οι συναρτήσεις συμμετοχής, επιτρέποντας την αντιστοίχισή τους σε λεκτικές όρους (linguistic terms) της ασαφούς λογικής όπως για παράδειγμα «χαμηλό», «μεσαίο», «υψηλό», οι οποίοι μπορούν να διευκολύνουν την ερμηνεία των ασαφών κανόνων από τον χρήστη.

6.6.2 Συσχέτιση κανόνων με συμπεριφορές επιθέσεων

Ακολουθεί η ανάλυση τεσσάρων ασαφών κανόνων από τους συνολικά δέκα που παρήχθησαν από το βέλτιστο μοντέλο ταξινόμησης πολλαπλών κλάσεων, Σενάριο 2 με 42 χαρακτηριστικά. Για λόγους ευκρίνειας παρουσιάζονται στη συνέχεια τέσσερις ενδεικτικοί κανόνες, οι οποίοι θεωρήθηκαν οι πλέον αντιπροσωπευτικοί για την ερμηνεία της λειτουργίας του μοντέλου. Οι υπόλοιποι κανόνες παρατίθενται στο Παράρτημα Α. Υπενθυμίζεται πως οι παρακάτω τιμές των κανόνων είναι κανονικοποιημένες με Z-score:

- **Κλάση Generic:** Οι επιθέσεις τύπου Generic, αποτελούν μια ευρεία κατηγορία κακόβουλων ενεργειών. Σύμφωνα με το σχετικό άρθρο του συνόλου δεδομένων [39], οι επιθέσεις τους συγκεκριμένου τύπου ορίζονται στοχευμένα ως τεχνικές παραβίασης που λειτουργούν εναντίον αλγορίθμων κρυπτογράφησης (block-cipher).
 - **Κανόνας:** IF ($dur \approx -0.82$) AND ($dloss \approx 1.696$) AND ($Dpkts \approx 1.708$) AND (...) THEN=>Generic
 - **Κύρια χαρακτηριστικά:** Ο κανόνας ενεργοποιείται όταν η συνολική διάρκεια της σύνδεσης καταγράφεται ως αισθητά μειωμένη (Record total duration, $dur \approx -0.82$ - χαμηλή τιμή) σε σχέση με τον μέσο όρο, ενώ ταυτόχρονα παρατηρείται σημαντικά αυξημένη απώλεια πακέτων προς τον στόχο-δίκτυο (Destination packets retransmitted or dropped, $dloss \approx 1.696$ - πολύ υψηλή τιμή) και απότομη αύξηση στα πακέτα που επιστρέφονται στον επιτιθέμενο (Destination to Source packet count, $Dpkts \approx 1.708$ - πολύ υψηλή τιμή).
 - **Πιθανή ερμηνεία:** Η ταυτόχρονη ύπαρξη ελάχιστου χρόνου σύνδεσης σε συνδυασμό με μεγάλο όγκο χαμένων αλλά και επιστρεφόμενων πακέτων μπορεί να υποδεικνύει την απότομη και μαζική αποστολή δεδομένων.
 - **Σύνδεση με Network Behavior:** Οι επιθέσεις Generic παρουσιάζουν παρόμοια συμπεριφορά, καθώς προσομοιάζουν τεχνικές εξαντλητικής αναζήτησης (Brute Force) ή επιθέσεις Codebook ενάντια σε αλγορίθμους κρυπτογράφησης (όπως ο AES). Οι τεχνικές αυτές απαιτούν συστηματικές και επαναλαμβανόμενες δοκιμές για την αποκρυπτογράφηση των πακέτων. Η μαζική και συνεχόμενη αποστολή αυτών των δοκιμαστικών πακέτων μπορεί να προκαλέσει υψηλή απώλεια και αναμετάδοση πακέτων μέσα σε σύντομο χρονικό διάστημα.
- **Κλάση Fuzzers:** Οι επιθέσεις τύπου Fuzzers ορίζονται ως προσπάθεια αναστολής ή κατάρρευσης της λειτουργίας του δικτύου, τροφοδοτώντας το δίκτυο συνεχώς με τυχαία ή μη έγκυρα δεδομένα.
 - **Κανόνας:** IF ($dur \approx 0.812$) AND ($sttl \approx 1.257$) AND ($dloss \approx 1.768$) AND ($Dpkts \approx 1.680$) AND (...) THEN=>Fuzzers
 - **Κύρια χαρακτηριστικά:** Ο συγκεκριμένος κανόνας ανιχνεύει τη μεγάλη διάρκεια σύνδεσης ($dur \approx 0.812$ - υψηλή τιμή), καθώς και τον αυξημένο χρόνο ζωής πακέτων από την πλευρά του επιτιθέμενου προς το δίκτυο (Source to Destination Time to Live, $sttl \approx 1.257$ - υψηλή τιμή). Ταυτόχρονα, καταγράφεται μεγάλη απώλεια και αναμετάδοση πακέτων από το δίκτυο ($dloss \approx 1.768$ - πολύ υψηλή τιμή), συνοδευόμενη από αύξηση πακέτων που επιστρέφονται από το δίκτυο προς τον επιτιθέμενο ($Dpkts \approx 1.680$ - πολύ υψηλή τιμή).
 - **Πιθανή ερμηνεία:** Η παρουσία μακροχρόνιας σύνδεσης σε συνδυασμό με τον παρατεταμένο χρόνο ζωής των πακέτων, κατά την οποία το δίκτυο αντιμετωπίζει συνεχή σφάλματα και μεγάλο όγκο αναμεταδιδόμενων πακέτων, ενδέχεται να υποδηλώνει μια προβληματική επικοινωνία.
 - **Σύνδεση με Network Behavior:** Το συγκεκριμένο μοτίβο δικτυακής κίνησης φαίνεται να συνδέεται με την παρουσία επιθέσεων τύπου Fuzzers. Ο επιτιθέμενος αποστέλλει μεγάλο όγκο τυχαίων και απρόβλεπτων δεδομένων, διατηρώντας τη σύνδεση ενεργή για μεγάλο χρονικό διάστημα. Το δίκτυο από τη μεριά του αδυνατεί να επεξεργαστεί τα λανθασμένα διαμορφωμένα δεδομένα, με αποτέλεσμα να παράγει συνεχώς σφάλματα. Λόγω της προσπάθειας επαναποστολής των λανθασμένων πακέτων πίσω στον επιτιθέμενο ή της οριστικής απόρριψής τους, παρουσιάζεται η αυξημένη τιμή $dloss$. Η διαδικασία αυτή ενδέχεται να προκαλέσει υπερφόρτωση και πιθανώς την αναστολή λειτουργίας του δικτύου, εξυπηρετώντας τον στόχο του συγκεκριμένου τύπου επίθεσης. Επομένως, η σύνδεση του κανόνα με τη συγκεκριμένη ερμηνεία μπορεί σε σημαντικό βαθμό να επαληθεύσει το μοτίβο που ανιχνεύθηκε.
- **Κλάση Reconnaissance:** Η κατηγορία Reconnaissance περιλαμβάνει «χαρτογράφηση» του δικτύου, ανίχνευση εύλωτων θυρών (ports) και υπηρεσιών που μπορούν να αποκαλυφθούν. Γενικά, περιέχει κακόβουλες ενέργειες με στόχο τη συλλογή δεδομένων και την ανίχνευση αδυναμιών του δικτύου.

- **Κανόνας:** IF ($dur \approx -0.629$) AND ($Spkts \approx -0.548$) AND ($swin \approx -0.644$) AND ($dwin \approx -0.657$) AND ($sloss \approx 0.558$) AND (...) THEN=>Reconnaissance
- **Κύρια χαρακτηριστικά:** Ο κανόνας ανιχνεύει συνδέσεις με μικρή χρονική διάρκεια ($dur \approx -0.629$ - χαμηλή τιμή) και μειωμένο όγκο απεσταλμένων πακέτων από τον επιτιθέμενο προς το δίκτυο (Source to Destination packet count, $Spkts \approx -0.548$ - χαμηλή τιμή). Παρά τον μικρό όγκο πακέτων που στέλνονται προς το δίκτυο, εντοπίζεται δυσανάλογα αυξημένη απώλεια πακέτων του επιτιθέμενου ($sloss \approx 0.558$ - υψηλή τιμή). Επιπρόσθετα, καταγράφονται αισθητά μειωμένα μεγέθη στα παράθυρα TCP, τόσο από την πλευρά του αποστολέα (Source TCP window, $swin \approx -0.644$ - χαμηλή τιμή) όσο και του παραλήπτη (Destination TCP window, $dwin \approx -0.657$ - χαμηλή τιμή).
- **Πιθανή ερμηνεία:** Η αποστολή ενός μικρού αριθμού πακέτων σε μικρό χρονικό διάστημα, εκ των οποίων ένα μεγάλο ποσοστό δεν λαμβάνει απάντηση ή καταγράφεται ως χαμένο, μπορεί να υποδηλώνει μια σύντομη αναγνωριστική αλληλεπίδραση χωρίς σταθερή και επιτυχή επικοινωνία.
- **Σύνδεση με Network Behavior:** Αυτό το μοτίβο είναι πιθανό να προσομοιάζει με τεχνικές δικτυακής σάρωσης και διερεύνησης (Network Scanning and Probing), καθώς και σάρωσης θυρών, οι οποίες εκτελούνται συχνά με εργαλεία τύπου Nmap. Σύμφωνα με το [53], τεχνικές όπως το TCP SYN Scan αποστέλλουν μεμονωμένα πακέτα όπου η τυπική χειραψία τριών σταδίων (3-way handshake) του TCP διακόπτεται εσκεμμένα πριν από την ολοκλήρωσή της, γεγονός που αποτυπώνεται και στα μικρότερα μεγέθη των TCP παραθύρων. Αυτό μπορεί να εξηγήσει τη μικρή διάρκεια σύνδεσης καθώς και τον περιορισμένο όγκο απεσταλμένων πακέτων, ο οποίος υποδεικνύει απουσία κανονικής επικοινωνίας για τη μεταφορά δεδομένων (payload). Επίσης, η αυξημένη τιμή της απώλειας πακέτων, όπως αναφέρεται και στο [53], πιθανώς αντικατοπτρίζει την προσπάθεια του επιτιθέμενου να σαρώσει πολλές θύρες, με αποτέλεσμα ορισμένα πακέτα να φιλτράρονται και να απορρίπτονται ή να αποστέλλονται σε μη διαθέσιμες θύρες.
- **Κλάση Backdoors:** Οι επιθέσεις τύπου Backdoors αφορούν τεχνικές όπου οι μηχανισμοί ασφαλείας ενός συστήματος παρακάμπτονται με τέτοιο τρόπο ώστε να μην ανιχνευθεί η παραβίαση. Στόχος αυτών των επιθέσεων είναι η πρόσβαση σε κρίσιμα δεδομένα με τη δημιουργία κρυφών διόδων επικοινωνίας για την απομακρυσμένη διαχείρισή τους.
 - **Κανόνας:** IF ($dur \approx 0.764$) AND ($Spkts \approx 0.82$) AND ($sloss \approx -0.485$) AND ($dttl \approx 1.796$) AND (...) THEN=>Backdoors
 - **Κύρια χαρακτηριστικά:** Ο κανόνας ανιχνεύει συνδέσεις με αυξημένη χρονική διάρκεια ($dur \approx 0.764$ - υψηλή τιμή) και αυξημένο όγκο πακέτων από τον επιτιθέμενο προς το δίκτυο ($Spkts \approx 0.82$ - υψηλή τιμή). Παράλληλα, παρατηρείται μεγάλη αύξηση στον χρόνο ζωής των πακέτων που επιστρέφει το δίκτυο στον επιτιθέμενο (Destination to Source time to live, $dttl \approx 1.796$ - πολύ υψηλή τιμή). Αντίθετα, καταγράφεται μείωση στην απώλεια πακέτων από την πλευρά του επιτιθέμενου ($sloss \approx -0.485$ - χαμηλή τιμή).
 - **Πιθανή ερμηνεία:** Η διατήρηση παρατεταμένης σύνδεσης με λίγες απώλειες πακέτων, παράλληλα αυξημένη ροή πακέτων προς το δίκτυο και επιστροφή πακέτων με υψηλό χρόνο ζωής (Time To Live - TTL), ενδέχεται να υποδηλώνει την ύπαρξη ενός καναλιού επικοινωνίας.
 - **Σύνδεση με Network Behavior:** Η συγκεκριμένη συμπεριφορά φαίνεται να είναι συμβατή με πρακτικές τύπου Backdoor που ενσωματώνουν επικοινωνία C2 (Command and Control). Η αυξημένη διάρκεια σύνδεσης και το μικρότερο μέγεθος απώλειας πακέτων, δικαιολογούνται από τη διατήρηση μιας σταθερής σύνδεσης με το παραβιασμένο δίκτυο-σύστημα. Σύμφωνα με το [54], οι Backdoor συνδέσεις στοχεύουν στην παροχή σταθερής και διαδραστικής πρόσβασης (Interactive Access) μέσω απομακρυσμένων εντολών. Η αυξημένη ροή πακέτων από τον επιτιθέμενο μπορεί να ερμηνευθεί από τη συνεχή αποστολή αυτών των απομακρυσμένων εντολών. Παράλληλα, η εξαιρετικά υψηλότερη τιμή του χρόνου ζωής των πακέτων από το δίκτυο προς τον επιτιθέμενο, είναι πιθανό να υποδεικνύει την επιτυχή και συνεχή μετάδοση

πακέτων-απαντήσεων από το δίκτυο. Ο συνδυασμός όλων των παραπάνω είναι πολύ πιθανό να δικαιολογεί τη δομή του συγκεκριμένου κανόνα.

6.6.3 Ερμηνευσιμότητα του μοντέλου

Στον τομέα της ανίχνευσης δικτυακών εισβολών (IDS), η υψηλή ορθότητα (Accuracy) ενός μοντέλου δεν είναι πάντα επαρκής από μόνη της. Σε γενικές γραμμές, τα μοντέλα βαθιάς μάθησης ή μοντέλα όπως το Random Forest έχουν εξαιρετικές επιδόσεις σε προβλήματα ταξινόμησης ακόμα και σε δεδομένα με μεγάλη διαστατικότητα. Ωστόσο, τα συγκεκριμένα διαθέτουν περιορισμένη εγγενή ερμηνευσιμότητα. Η διαδικασία λήψης αποφάσεων που ακολουθούν είναι εξαιρετικά πολύπλοκη και η εσωτερική τους λογική δεν είναι άμεσα αναγνώσιμη σε μορφή ενός συνόλου κανόνων. Κατά συνέπεια, απαιτούνται συνήθως πρόσθετες μεταγενέστερες τεχνικές επεξήγησης (post-hoc explainability) από τον αναλυτή ασφαλείας για να κατανοήσει τον λόγο για τον οποίο το μοντέλο ταξινόμησε μια εγγραφή σε μια συγκεκριμένη κλάση. Αυτή η έλλειψη άμεσης διαφάνειας μπορεί να προκαλέσει προβλήματα εμπιστοσύνης και δυσκολεύει την εξαγωγή χρήσιμων και αξιοποιήσιμων πληροφοριών [55][56].

Το ANFIS αποτελεί μια αρχιτεκτονική που γεφυρώνει το χάσμα μεταξύ των υψηλών επιδόσεων και της διαφάνειας λειτουργίας του μοντέλου. Οι ασαφείς κανόνες που παράγονται από το μοντέλο παρέχουν μια περισσότερο άμεσα αναγνώσιμη αναπαράσταση της εσωτερικής λογικής του μοντέλου σε σύγκριση με μοντέλα περιορισμένης εγγενούς ερμηνευσιμότητας, καθώς μεταφράζονται σε γλωσσικούς κανόνες της μορφής AN-TOTE (IF-THEN). Το γεγονός αυτό καθιστά το μοντέλο ANFIS ένα πιο διάφανο λογικό μοντέλο, όπου οι ασαφείς κανόνες αποκαλύπτουν τον τρόπο που το μοντέλο καταλήγει σε μια συγκεκριμένη απόφαση.

Είναι σημαντικό να αναφερθεί πως αυτοί οι κανόνες ακολουθούν την ασαφή λογική, η οποία διαφοροποιείται από την κλασική δίτιμη λογική. Σε ένα μοντέλο δέντρων απόφασης, ένας κανόνας εκφράζει με απόλυτη βεβαιότητα αν ικανοποιείται η συνθήκη του τμήματος IF, με το αποτέλεσμα του τμήματος THEN να ισχύει χωρίς περιθώριο αμφισβήτησης. Αντίθετα, στο μοντέλο ANFIS, οι τιμές των χαρακτηριστικών δεν ανήκουν απόλυτα σε ένα συγκεκριμένο σύνολο, αλλά χαρακτηρίζονται από έναν βαθμό συμμετοχής. Επομένως, όταν ένας ασαφής κανόνας ενεργοποιείται, η ταξινόμηση που πραγματοποιεί στο THEN τμήμα δεν αποτελεί μια απόλυτη πρόβλεψη. Ουσιαστικά, το αποτέλεσμα του τμήματος THEN είναι σταθμισμένο (weighted) από τη βαρύτητα (ή αλλιώς την πεποίθηση - confidence) του κανόνα, η οποία εξαρτάται αποκλειστικά από τον βαθμό ικανοποίησης των συνθηκών του τμήματος IF μέσω των συναρτήσεων συμμετοχής. Η κλάση που τελικά αποδίδεται στο τμήμα THEN ενός κανόνα, προκύπτει επειδή το μοτίβο συμπεριφοράς του IF τμήματος έχει τη μεγαλύτερη δυνατότητα συμβατότητας με τον συγκεκριμένο κανόνα μέσα στο αντίστοιχο cluster. Δηλαδή, ένας κανόνας υποδεικνύει την κυρίαρχη τάση των δεδομένων προς μια συγκεκριμένη κλάση, αναγνωρίζοντας παράλληλα την πιθανή συνύπαρξη άλλων κλάσεων με μικρότερη δυνατότητα συμβατότητας στο ίδιο cluster. Η τελική απόφαση του μοντέλου για την ταξινόμηση μιας συγκεκριμένης εγγραφής του συνόλου δεδομένων δεν βασίζεται αποκλειστικά σε έναν μόνο κανόνα. Αυτή προκύπτει από τη διαδικασία της ασαφούς συνάθροισης (fuzzy aggregation), κατά την οποία συνδυάζονται τα αποτελέσματα-συμπεράσματα όλων των κανόνων που ενεργοποιήθηκαν και στη συνέχεια με τη διαδικασία αποασαφοποίησης (defuzzification), το αποτέλεσμα μετατρέπεται σε μια καθαρή τελική τιμή εξόδου.

Στην προηγούμενη υποενότητα 6.6.2, αναλύονται μερικοί από τους ασαφείς κανόνες που παρήγαγε το μοντέλο. Η ικανότητα της ερμηνείας και τεκμηρίωσης συγκεκριμένων κανόνων προσφέρει την απαραίτητη πληροφορία που χρειάζονται τα μέλη της ομάδας SOC, επιτρέποντας έτσι την εξάλειψη του προβλήματος της περιορισμένης εμπιστοσύνης προς την αξιοπιστία του μοντέλου. Η άμεση

ερμηνεία και σύνδεση των κανόνων αυτών δεν παρέχεται εγγενώς από ένα μοντέλο όπως το Random Forest, το οποίο σε περίπτωση λανθασμένης απόφασης θα απαιτούσε από τον αναλυτή την εφαρμογή πρόσθετων μεταγενέστερων τεχνικών ανάλυσης για να αιτιολογήσει το αποτέλεσμα.

Σε αυτό το σημείο, πρέπει να γίνει μια κρίσιμη διάκριση ανάμεσα στη δυνατότητα εξαγωγής των κανόνων και στην πλήρη κατανόηση ολόκληρου του μοντέλου. Το βέλτιστο μοντέλο ταξινόμησης πολλαπλών κλάσεων χρησιμοποιεί 42 χαρακτηριστικά και 10 κανόνες, γεγονός που καθιστά τους πλήρεις κανόνες αρκετά εκτενείς. Επομένως, είναι αναγκαίο να διευκρινιστεί ότι η παρούσα εργασία αναδεικνύει κυρίως τη δυνατότητα εξαγωγής και ποιοτικής ερμηνείας επιλεγμένων κανόνων (ή τμημάτων αυτών) και των κυρίαρχων χαρακτηριστικών τους, και όχι την πλήρη και άμεση ερμηνεία όλων των συνθηκών του μοντέλου ταυτόχρονα από έναν αναλυτή ασφαλείας.

Κεφάλαιο 7ο: Συζήτηση αποτελεσμάτων

7.1 Συνολική αξιολόγηση του προτεινόμενου μοντέλου

Το προτεινόμενο υβριδικό μοντέλο ANFIS-GA, έπειτα από μια εκτενή πειραματική διαδικασία με πολλούς ελέγχους αξιολόγησης, αναδείχθηκε ως η καλύτερη προσέγγιση μεταξύ των μοντέλων ANFIS που εξετάστηκαν. Πιο συγκεκριμένα, το Σενάριο 2 με την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών αποτελεί μια υποσχόμενη αρχιτεκτονική επιλογή για το συγκεκριμένο σύνολο δεδομένων και τη χρήση του μοντέλου ως IDS. Οι έλεγχοι Friedman και Conover απέδειξαν τη διαφοροποίηση που παρουσιάζει το προτεινόμενο μοντέλο από τα υπόλοιπα πειράματα. Επίσης, η διαφορά και η υπεροχή του Σεναρίου 2 φαίνεται οπτικά από το Σχήμα 6.7 και Σχήμα 6.8 για τις δύο ταξινομήσεις. Εξαιρετικά σημαντική είναι η ικανότητα του ANFIS να εξάγει τους ασαφείς κανόνες που παράγει, προσφέροντας διαφάνεια στον τρόπο λειτουργίας του, ενώ παράλληλα παρουσιάζει αρκετά καλές επιδόσεις. Το Random Forest, ως το ισχυρό μοντέλο αναφοράς, αν και παρουσίασε σχετικά καλύτερα αποτελέσματα, διαθέτει περιορισμένη εγγενή ερμηνευσιμότητα, γεγονός που απαιτεί επιπλέον ενέργειες επεξήγησης για να εδραιωθεί εμπιστοσύνη προς το μοντέλο σε πραγματικά περιβάλλοντα κυβερνοασφάλειας.

7.2 Συμβιβασμοί μεταξύ απόδοσης και ερμηνευσιμότητας

Όπως αναφέρθηκε και στην προηγούμενη ενότητα, το Random Forest δεν επιλέχθηκε ως βέλτιστο στο πλαίσιο του συμβιβασμού μεταξύ της απόδοσης και της ερμηνευσιμότητας, όπου δόθηκε προτεραιότητα στην ερμηνευσιμότητα. Σύμφωνα με τα αποτελέσματα του Κεφαλαίου 6, το ANFIS παρουσιάζει μεγαλύτερη χρονική διάρκεια εκπαίδευσης και συνεπώς μεγαλύτερο υπολογιστικό κόστος, ενώ παράλληλα εμφανίζει σχετικά χαμηλότερα επίπεδα επιδόσεων με βάση τις μετρικές αξιολόγησης. Αντίθετα, το Random Forest εμφανίζει εξαιρετικές επιδόσεις καθώς και αποδοτικότητα λόγω του εξαιρετικά χαμηλότερου χρόνου εκπαίδευσης. Ωστόσο, όπως έχει ήδη αναλυθεί, το Random Forest δεν παρέχει την ικανότητα ερμηνείας της λήψης αποφάσεών του, σε αντίθεση με το ANFIS που έχει τη συγκεκριμένη ικανότητα, ενώ παράλληλα παρουσιάζει χαμηλότερες, αλλά όχι με μεγάλη διαφορά, επιδόσεις σε σύγκριση με το Random Forest. Έτσι, εφαρμόζεται ο συμβιβασμός μεταξύ της καλύτερης απόδοσης και της ερμηνευσιμότητας, όπου επιλέγεται το μοντέλο ANFIS λόγω της διαφάνειας που προσφέρει. Στον τομέα της κυβερνοασφάλειας και των IDS, αυτός ο συμβιβασμός είναι ευρέως αποδεκτός. Ένα μοντέλο που παρέχει ικανοποιητική απόδοση με ταυτόχρονη αιτιολόγηση των αποφάσεών του υπερτερεί έναντι ενός άλλου μοντέλου black-box με καλύτερη απόδοση αλλά μηδενική διαφάνεια του τρόπου λειτουργίας του.

7.3 Επίδραση της επιλογής χαρακτηριστικών (feature selection)

Η εφαρμογή του γενετικού αλγορίθμου για την επιλογή χαρακτηριστικών από το σύνολο δεδομένων UNSW NB-15, αποδείχθηκε πως έχει καθοριστικό ρόλο στη συνολική απόδοση. Το Σενάριο 2, όπου γίνεται η ταυτόχρονη επιλογή χαρακτηριστικών και η βέλτιστη υπερπαραμετροποίηση του ANFIS, αποτελεί την υλοποίηση με την οποία εμφανίστηκαν τα εξαιρετικά αποτελέσματα που παρουσιάζονται στους αντίστοιχους πίνακες του Κεφαλαίου 6.

Οι προσεγγίσεις σχετικά με την επιλογή χαρακτηριστικών του Σεναρίου 2, διαχωρίστηκαν σε ελεύθερου πλήθους και σε προκαθορισμένου πλήθους επιλογή χαρακτηριστικών. Τα αποτελέσματα των πειραμάτων μεταξύ των δύο προσεγγίσεων απέδειξαν πως η επιλογή χαρακτηριστικών επηρεάζει την

απόδοση με αντίκτυπο στις μετρικές αξιολόγησης αλλά και στην πολυπλοκότητα. Η προσέγγιση προκαθορισμένου πλήθους ανάγκασε τον γενετικό αλγόριθμο να κρατάει συγκεκριμένο πλήθος επιλεγμένων χαρακτηριστικών το οποίο επέτρεψε τον πειραματισμό με ακραίες τιμές πλήθους χαρακτηριστικών. Τα πειράματα με πολύ χαμηλό πλήθος χαρακτηριστικών γενικά εμφάνισαν καλύτερο F1-score από το αντίστοιχο μοντέλο αναφοράς ANFIS, χρησιμοποιώντας πολύ λιγότερα χαρακτηριστικά από τα συνολικά 202. Τα πειράματα με υψηλό πλήθος επιλεγμένων χαρακτηριστικών, εμφάνισαν καλύτερα αποτελέσματα από τα αντίστοιχα πειράματα με εξαιρετικά χαμηλό πλήθος, όμως με μεγαλύτερη πολυπλοκότητα και χρόνο εκπαίδευσης. Συνολικά, και για τις δύο ταξινομήσεις τα πειράματα προκαθορισμένου πλήθους που έχουν τιμή πλήθους χαρακτηριστικών πιο κοντά στο πλήθος χαρακτηριστικών των πειραμάτων της ελεύθερης επιλογής, παρουσιάζουν τα καλύτερα αποτελέσματα συγκριτικά με τα υπόλοιπα πειράματα προκαθορισμένου πλήθους. Αυτό το γεγονός επιβεβαιώνει και επαληθεύει κατά κάποιο τρόπο το βέλτιστο εύρος πλήθους επιλεγμένων χαρακτηριστικών για κάθε ταξινόμηση. Βέβαια, τα βέλτιστα αποτελέσματα κάθε ταξινόμησης προέρχονται από την προσέγγιση ελεύθερης επιλογής, ωστόσο και τα αντίστοιχα πειράματα προκαθορισμένου πλήθους εμφανίζουν ικανοποιητικά αποτελέσματα.

Η μείωση της διαστατικότητας των δεδομένων έχει άμεση επίδραση στην απόδοση του προτεινόμενου μοντέλου. Με τον περιορισμό του όγκου και την αύξηση της ποιότητας των δεδομένων που καλείται να διαχειριστεί το ANFIS, αυξάνεται η συνολική απόδοση του μοντέλου. Όσον αφορά τις απαιτήσεις σε πόρους, πρέπει να διευκρινιστεί ότι η επιλογή χαρακτηριστικών μειώνει το υπολογιστικό κόστος του τελικού ANFIS, αλλά η συνολική διαδικασία βελτιστοποίησης παραμένει υπολογιστικά απαιτητική λόγω των πολλαπλών εκπαιδεύσεων του εσωτερικού ANFIS μέσα στον εξελικτικό αλγόριθμο. Πρακτικά, η αυξημένη απαίτηση πόρων μετατοπίζεται αποκλειστικά στο αρχικό στάδιο της ανάπτυξης, αποδίδοντας τελικά ένα ελαφρύ, γρήγορο και αποδοτικό μοντέλο, έτοιμο για επιχειρησιακή χρήση. Το προτεινόμενο μοντέλο παρουσιάζει τις συνολικά βέλτιστες τιμές του F1-score ανάμεσα στα μοντέλα των πειραμάτων που χρησιμοποιούν το ANFIS. Αυτό αποδεικνύεται και από τις πολλαπλές εκτελέσεις των πειραμάτων που τροφοδοτήθηκαν στον έλεγχο Friedman καθώς και στον έλεγχο Conover, όπου το Σενάριο 2 παρουσιάζει ισχυρή στατιστική διαφορά (βελτίωση) από τα πειράματα με το 1ο μοντέλο αναφοράς και το Σενάριο 1. Έτσι, με την επιλογή χαρακτηριστικών μέσω του γενετικού αλγορίθμου, το ANFIS μετατρέπεται σε ένα πιο αποδοτικό και επιχειρησιακά ταχύτερο μοντέλο.

7.4 Ανάλυση ψευδώς θετικών και ψευδώς αρνητικών αποτελεσμάτων

Στο πλαίσιο των IDS, η ανάλυση των λανθασμένων ταξινομήσεων είναι εξίσου σημαντική με την ανάλυση των υπόλοιπων μετρικών, όπως το F1-score που συνδυάζει την ακρίβεια και την ανάκληση. Ωστόσο, μια λανθασμένα ταξινομημένη εγγραφή του συνόλου δεδομένων δεν σημαίνει απαραίτητα και ολική αποτυχία.

Για τη δυαδική ταξινόμηση, το προτεινόμενο μοντέλο κατάφερε να καταγράψει μόλις 18 Ψευδώς Αρνητικά (χαμένες επιθέσεις) στη βέλτιστη εκτέλεσή του (με τη μέση τιμή των 10 εκτελέσεων να παραμένει επίσης σε χαμηλότερα επίπεδα από το Random Forest), ενώ το Random Forest κατέγραψε 83 Ψευδώς Αρνητικά αποτελέσματα. Αυτό αναδεικνύει την ευαισθησία του προτεινόμενου μοντέλου, καθώς αποτελεί ένα ολικά πιο αυστηρό μοντέλο ως IDS, όπου δεν υπάρχουν πολλές επιθέσεις τις οποίες δεν αναγνώρισε. Βέβαια, το πιο αυστηρό μοντέλο έχει ως επίπτωση την αύξηση των Ψευδώς Θετικών (False Positives), τα οποία αντιστοιχούν σε λανθασμένες ταξινομήσεις εγγραφών πραγματικής φυσιολογικής δικτυακής κίνησης ως επιθέσεις. Ως αποτέλεσμα, το σύστημα μπορεί να προκαλέσει «κόπωση συναγερμών» (alert fatigue) στους διαχειριστές του δικτύου λόγω των πολλαπλών λανθασμένων υποτιθέμενων επιθέσεων. Ωστόσο, αυτή η τεχνική αποτελεί ένα γνωστό συμβιβασμό

στον τομέα των IDS. Όπως αναφέρει και το [5], συνηθίζεται στα IDS να προτιμάται ένα μοντέλο το οποίο έχει περισσότερους λανθασμένους συναγερμούς και πολύ λιγότερες χαμένες επιθέσεις, ακόμα και αν αυτό οδηγεί σε μεγαλύτερο υπολογιστικό και επιχειρησιακό κόστος.

Στην ταξινόμηση πολλαπλών κλάσεων, η αξιολόγηση των λανθασμένων ταξινομήσεων αποκτά διαφορετική διάσταση. Σε μια λανθασμένη ταξινόμηση εγγραφής, έχει σημασία αν αυτή ανήκει σε άλλη κατηγορία επίθεσης ή αν αντιστοιχεί στη Normal κλάση. Αν η επίθεση ταξινομηθεί λανθασμένα στην κλάση Normal, τότε θεωρείται ως Ψευδώς Αρνητικό, δηλαδή μια επίθεση την οποία δεν ανίχνευσε το μοντέλο. Όμως, αν η λανθασμένη ταξινόμηση γίνει σε κλάση η οποία αντιστοιχεί σε άλλη επίθεση, τότε αυτό δεν αποτελεί σημαντικό πρόβλημα για το IDS καθώς έχει αναγνωρίσει την εγγραφή ως επίθεση ούτως ή άλλως. Η ταξινόμηση σε λανθασμένη κλάση επίθεσης, θα είχε πολύ σημαντικό ρόλο αν αντί για IDS το μοντέλο προοριζόταν για IPS (Intrusion Prevention System). Το IPS ουσιαστικά αποτελείται από IDS με επιπλέον δυνατότητες διαχείρισης για την αντιμετώπιση ή τον περιορισμό της επίθεσης. Σε ένα τέτοιο σύστημα, μια εγγραφή πρέπει να είναι απολύτως σωστά ταξινομημένη ώστε να επιλεγεί ο ανάλογος τρόπος για την αντιμετώπισή της. Οι διαφορετικού τύπου επιθέσεις έχουν και διαφορετικούς τρόπους διαχείρισης, ωστόσο στην προκειμένη περίπτωση της εργασίας όπου υλοποιείται ένα IDS, δεν δίνεται τόσο μεγάλη βαρύτητα για τη συγκεκριμένη κατάσταση. Τα αποτελέσματα των θεωρητικά Ψευδώς Αρνητικών (χαμένων επιθέσεων) στη συγκεκριμένη ταξινόμηση ήταν ανώτερα και ξεπέρασαν αυτά της δυαδικής ταξινόμησης, ταξινομώντας μόνο 14 πραγματικές επιθέσεις ως Normal στη βέλτιστη εκτέλεση (με τον μέσο όρο των 10 εκτελέσεων να διατηρείται εξίσου χαμηλά). Το συγκεκριμένο φαινόμενο, όπου η ταξινόμηση με μεγαλύτερο βαθμό πολυπλοκότητας εμφανίζει καλύτερα αποτελέσματα, οφείλεται στη διαφορετική προεπεξεργασία των δεδομένων εκπαίδευσης, όπως έχει ήδη αναλυθεί.

7.5 Συμπεριφορά σε σπάνιες επιθέσεις

Στην ταξινόμηση πολλαπλών κλάσεων, οι σπάνια εμφανιζόμενες κλάσεις επιθέσεων παρουσιάζουν ιδιαίτερα ενδιαφέροντα αποτελέσματα. Όπως καταγράφηκε στα αποτελέσματα του Κεφαλαίου 6, το προτεινόμενο μοντέλο διατήρησε αρκετά ικανοποιητικά επίπεδα αναγνώρισης αυτών των κλάσεων, αποφεύγοντας την πλήρη επισκίασή τους από τις υπόλοιπες πλειοψηφικές σε εγγραφές κλάσεις. Συγκεκριμένα, η κλάση επίθεσης Worms στο αρχικό σύνολο δεδομένων παρουσιάζεται μόνο 174 φορές, οι οποίες είναι σημαντικά λιγότερες από τις υπόλοιπες κλάσεις. Ωστόσο, παρά τις αρχικά λιγοστές εμφανίσεις της συγκεκριμένης κλάσης, το προτεινόμενο μοντέλο ANFIS-GA κατάφερε να αναγνωρίσει (recall) το 83% των πραγματικών εγγραφών της κλάσης Worms, βέβαια με πολύ μικρή ακρίβεια (precision), η οποία προκάλεσε λανθασμένες ταξινομήσεις εγγραφών άλλων κλάσεων ως Worms. Ίδια συμπεριφορά υψηλού ποσοστού αναγνώρισης εμφανίζεται και για τις εγγραφές της αρχικά σπάνια εμφανιζόμενης κλάσης Shellcode, με επίσης μειωμένη ακρίβεια.

Η σημαντική απόδοση του προτεινόμενου μοντέλου σε αυτές τις σπάνια εμφανιζόμενες κλάσεις, οφείλεται στην προεπεξεργασία που έχουν υποστεί τα δεδομένα εκπαίδευσης του μοντέλου. Όπως έχει αναφερθεί στο Κεφάλαιο 4, έχει γίνει χρήση της τεχνικής SMOTE με την οποία αντιμετωπίζεται το πρόβλημα ακραίας ανισορροπίας κλάσεων. Η χρήση της τεχνικής επιτρέπει την αύξηση των εγγραφών σπάνιων κλάσεων την οποία αξιοποιεί το μοντέλο για την εκπαίδευσή του με στόχο την καλύτερη αναγνώριση και διαχωρισμό αυτών από τις υπόλοιπες πλειοψηφικές κλάσεις. Χωρίς κάποια τεχνική αντιμετώπισης της ανισορροπίας, θα αναμενόταν σημαντική υποβάθμιση της αναγνώρισης των σπάνιων κλάσεων που θα οδηγούσε σε συνολικά χαμηλότερη απόδοση του μοντέλου.

Η εμφάνιση της χαμηλής ακρίβειας των προβλέψεων στις σπάνιες κλάσεις προκαλείται επίσης λόγω της χρήσης του SMOTE. Τα επιπλέον συνθετικά δεδομένα που παράγει το SMOTE για τη βελτίωση της αναγνώρισης των κλάσεων, οδηγούν το μοντέλο να εκπαιδευτεί μεν με περισσότερη πληροφορία για την κλάση, όμως παράλληλα αυξάνεται η συχνότητα εμφάνισης στα δεδομένα εκπαίδευσης, το οποίο επιβεβαιώνεται και από τον Πίνακα 4.3 του Κεφαλαίου 4. Αυτό το φαινόμενο ονομάζεται αλλοίωση των πιθανοτήτων (prior probability shift), όπου το μοντέλο εκπαιδευτεί με δυσανάλογες πιθανότητες εμφάνισης μιας συγκεκριμένης κλάσης που δεν αντιστοιχούν στις πραγματικές συχνότητες [57]. Για παράδειγμα, όπως αποτυπώθηκε και στην ανάλυση του συνόλου εκπαίδευσης, η κλάση Worms ενισχύθηκε τεχνητά από μόλις 104 πραγματικές εγγραφές (λιγότερο από το 0,01% της αρχικής κίνησης) σε 80.000 εγγραφές, φτάνοντας να αποτελεί το 10% του τελικού συνόλου εκπαίδευσης. Το μοντέλο, λόγω αυτής της προκατάληψης (bias), είναι πιθανό να οδηγείται στην ταξινόμηση μιας εγγραφής σε συγκεκριμένη κλάση λόγω της αυξημένης τεχνητής συχνότητας που εμφανιζόταν κατά την εκπαίδευση. Το συγκεκριμένο πρόβλημα μπορεί να οδηγήσει και σε επισκίαση αναγνώρισης άλλων επιθέσεων. Ωστόσο, και το συγκεκριμένο πρόβλημα αποτελεί έναν επιπλέον συμβιβασμό ανάμεσα στην ακρίβεια και την αναγνώριση, ο οποίος είναι αποδεκτός καθώς το μοντέλο προορίζεται για IDS, όπου η αναγνώριση του μέγιστου δυνατού πλήθους επιθέσεων αποτελεί ύψιστη προτεραιότητα.

7.6 Πιθανή ευελιξία σε Zero-day επιθέσεις

Οι επιθέσεις τύπου Zero-day αποτελούν απειλές με στόχο την εκμετάλλευση αδυναμιών ή ευπαθειών ενός συστήματος ή δικτύου, η ύπαρξη των οποίων είναι ακόμη άγνωστη στους διαχειριστές. Οι συγκεκριμένες επιθέσεις είναι εξαιρετικά επικίνδυνες καθώς μπορούν να επηρεάσουν το γνωστό τρίπτυχο της κυβερνοασφάλειας που αποτελείται από την εμπιστευτικότητα, την ακεραιότητα και τη διαθεσιμότητα. Τα περισσότερα κλασικά μοντέλα Μηχανικής Μάθησης χρησιμοποιούν απόλυτα όρια απόφασης (hard decision boundaries) για τη λήψη των αποφάσεών τους λόγω της αρχιτεκτονικής τους. Αυτό σημαίνει πως όταν εμφανιστεί μια zero-day επίθεση, η οποία παρουσιάζει συμπεριφορά που δεν αντιστοιχεί σε κανένα από τα γνωστά μοτίβα της εκπαίδευσης, το μοντέλο έχει μειωμένες πιθανότητες να πραγματοποιήσει σωστή αναγνώριση και ταξινόμηση.

Αν και η παρούσα εργασία δεν αξιολογεί πειραματικά την ανίχνευση zero-day επιθέσεων, η αρχιτεκτονική του ANFIS παρουσιάζει ένα σημαντικό θεωρητικό πλεονέκτημα. Αντίθετα με τα περισσότερα κλασικά μοντέλα Μηχανικής Μάθησης, στην ασαφή λογική τα όρια απόφασης είναι ευέλικτα, καθώς δεν αναγκάζουν μια εγγραφή των δεδομένων να ανήκει αποκλειστικά σε μία μόνο κλάση. Αν μια Zero-day επίθεση αναγνωριστεί ως ύποπτη μέσω των ασαφών κανόνων του ANFIS, τότε το μοντέλο θα μπορούσε να την ταξινομήσει προσεγγιστικά στην κλάση που παρουσιάζει μεγαλύτερη ομοιότητα ανάλογα με τον βαθμό συμμετοχής. Έτσι, υπάρχουν περισσότερες πιθανότητες να ξεχωρίσει μια εγγραφή κλάσης Normal από μια Zero-day επίθεση, σε αντίθεση με το μοντέλο Random Forest που εξάγει αυστηρά απόλυτες ταξινομήσεις, αναγκάζοντας την εγγραφή να ταξινομηθεί σε απόλυτο βαθμό σε μία αποκλειστική κλάση, γεγονός που μπορεί να οδηγήσει σε λανθασμένα αποτελέσματα. Επομένως, το ANFIS ενδέχεται να προσφέρει μεγαλύτερη ευελιξία και θα μπορούσε να συμβάλει στην αντιμετώπιση πιθανών zero-day επιθέσεων, λόγω της χρήσης ασαφούς λογικής σε αντίθεση με ένα κλασικό μοντέλο Μηχανικής Μάθησης όπως το Random Forest με απόλυτα (crisp) όρια απόφασης [58][59].

7.7 Περιορισμοί της εργασίας

Η παρούσα εργασία ανέδειξε την αποτελεσματικότητα και τη διαφάνεια του προτεινόμενου υβριδικού μοντέλου ANFIS-GA για την ανίχνευση δικτυακών εισβολών. Ωστόσο, για την πλήρη και

αντικειμενική αξιολόγηση του προτεινόμενου μοντέλου, είναι κρίσιμο να αναγνωριστούν ορισμένοι περιορισμοί της έρευνας, οι οποίοι έχουν επηρεάσει τα τελικά αποτελέσματα.

Ο πρώτος και βασικός περιορισμός σχετίζεται με τα δεδομένα που χρησιμοποιήθηκαν. Αρχικά, η αξιολόγηση πραγματοποιήθηκε σε ένα μόνο σύνολο δεδομένων (UNSW NB-15) και επομένως απαιτείται περαιτέρω διερεύνηση της δυνατότητας γενίκευσης σε άλλα σύνολα δεδομένων. Επιπλέον, η αξιολόγηση πραγματοποιήθηκε offline σε καταγεγραμμένα δεδομένα και όχι σε πραγματικό περιβάλλον δικτύου. Το σύνολο δεδομένων αποτελείται αρχικά από 47 χαρακτηριστικά εισόδου, τα οποία στη συνέχεια με τη χρήση της τεχνικής One-Hot Encoding αυξάνονται σε 202. Η χρήση της συγκεκριμένης τεχνικής έχει σημαντική και επιχειρησιακή σημασία, καθώς χωρίς αυτή, η εκπαίδευση και ορθή λειτουργία του ANFIS θα ήταν αδύνατη. Το γεγονός της σημαντικής αύξησης των χαρακτηριστικών εισόδου προκαλεί εκθετική αύξηση του υπολογιστικού κόστους εκπαίδευσης του ANFIS. Έτσι, η επιλογή χαρακτηριστικών με τη βοήθεια του γενετικού αλγορίθμου ήταν αναγκαία. Επίσης, η παρουσία της ακραίας ανισορροπίας κλάσεων απαίτησε τη χρήση της τεχνικής SMOTE, η οποία, όπως αναλύθηκε στην προηγούμενη ενότητα, επιλύει μεν το πρόβλημα της επισκίασης των σπάνιων επιθέσεων, αλλά έχει ως επίπτωση την τεχνητή μετατόπιση πιθανοτήτων ή αλλιώς την αλλοίωση της συχνότητας εμφάνισης συγκεκριμένων κλάσεων που δεν αντιστοιχεί στην πραγματική και επηρεάζει και την τελική ακρίβεια (precision) του μοντέλου. Ένας επιπλέον μεθοδολογικός περιορισμός αφορά την ίδια την εφαρμογή του SMOTE, η οποία έγινε στο σύνολο των 202 χαρακτηριστικών, συμπεριλαμβάνοντας τα δυαδικά χαρακτηριστικά του One-Hot Encoding. Καθώς ο αλγόριθμος παράγει συνθετικά δεδομένα μέσω γραμμικής παρεμβολής, ενδέχεται να παρήγαγε ενδιάμεσες δεκαδικές τιμές σε στήλες που κανονικά λαμβάνουν αυστηρά δυαδικές τιμές (0 ή 1). Αν και αυτό δεν εμπόδισε την εκπαίδευση, καθώς το ANFIS διαχειρίζεται συνεχείς τιμές, αποτελεί μια σχεδιαστική απόκλιση.

Πέρα από τους περιορισμούς των δεδομένων, σημαντικό ρόλο έχουν και οι πρακτικοί περιορισμοί που αφορούν τις απαιτήσεις σε υπολογιστικούς πόρους. Ο συνδυασμός της εσωτερικής λειτουργίας του ANFIS για κάθε άτομο του πληθυσμού του γενετικού αλγορίθμου, απαιτεί την υλοποίηση εκατοντάδων συστημάτων ANFIS για κάθε πείραμα, δημιουργώντας σημαντικό υπολογιστικό φορτίο. Οι εκτελέσεις των πειραμάτων πραγματοποιήθηκαν με επεξεργαστή (CPU) AMD Ryzen 3 3200G, ο οποίος προσφέρει χαμηλές προς μεσαίες επιδόσεις σε σύγκριση με τους σύγχρονους επεξεργαστές της αγοράς. Επίσης, η έλλειψη μονάδας επεξεργασίας γραφικών (GPU) έπαιξε καθοριστικό ρόλο, καθώς θα επέτρεπε σημαντική μείωση του χρόνου εκπαίδευσης του ANFIS.

Η βιβλιοθήκη X-ANFIS, με την οποία γίνεται η υλοποίηση του ANFIS, βασίζεται στο πλαίσιο PyTorch. Το συγκεκριμένο πλαίσιο υποστηρίζει επιτάχυνση υλικού (hardware acceleration) μέσω της GPU, χρησιμοποιώντας τεχνολογίες όπως το CUDA της Nvidia ή το ROCm της AMD, με τις οποίες δίνεται η δυνατότητα παράλληλων υπολογισμών (Data Parallelism) [47]. Επομένως, η απουσία GPU στέρησε τη δυνατότητα ταχύτερων πράξεων που μπορεί να προσφέρει η βιβλιοθήκη για τη γρηγορότερη εκπαίδευση του ANFIS. Ως αποτέλεσμα, η αποκλειστική χρήση του συγκεκριμένου επεξεργαστή χαμηλών επιδόσεων δημιούργησε σημαντικό υπολογιστικό περιορισμό (bottleneck), αυξάνοντας τον συνολικό χρόνο εκτέλεσης του γενετικού αλγορίθμου αλλά και του τελικού ANFIS. Συγκεκριμένα, η εκτέλεση του γενετικού αλγορίθμου για τα πειράματα (Σενάριο 1 και Σενάριο 2) απαιτούσε αρκετές ώρες συνεχούς λειτουργίας για κάθε πειραματική διάταξη (setup), ώστε να εκπαιδευτούν και να αξιολογηθούν όλα τα άτομα κάθε γενιάς. Επομένως, η χρήση καλύτερου υπολογιστικού υλικού θα βελτίωνε ξεκάθαρα τις απαιτούμενες χρονικές περιόδους για την πραγματοποίηση των πειραμάτων.

Λόγω του περιορισμού στο υλικό, υπήρξε αναγκαστικός συμβιβασμός κατά τον καθορισμό των παραμέτρων του γενετικού αλγορίθμου. Το μέγεθος του πληθυσμού και ο μέγιστος αριθμός γενεών, διατηρήθηκαν σε συντηρητικά επίπεδα προκειμένου οι χρόνοι εκτέλεσης να παραμείνουν ρεαλιστικοί.

Τέλος, η χρήση early stopping περιόρισε τον κίνδυνο υπερεκπαίδευσης του εσωτερικού ANFIS. Ωστόσο, επειδή ο εξελικτικός αλγόριθμος αξιολογεί επανειλημμένα μεγάλο αριθμό υποψήφιων λύσεων στο ίδιο υποσύνολο αξιολόγησης της καταλληλότητας, δεν μπορεί να αποκλειστεί η έμμεση προσαρμογή της διαδικασίας βελτιστοποίησης στο συγκεκριμένο υποσύνολο (φαινόμενο γνωστό ως validation overfitting ή meta-overfitting).

Κεφάλαιο 8ο: Συμπεράσματα

8.1 Συνολικά συμπεράσματα

Η παρούσα εργασία επικεντρώθηκε στη μελέτη, την ανάπτυξη και την αξιολόγηση ενός υβριδικού συστήματος ανίχνευσης εισβολών σε δίκτυα, συνδυάζοντας την αρχιτεκτονική του Προσαρμοστικού Νευρο-Ασαφούς Συστήματος Συμπερασμού (ANFIS) με τεχνικές βελτιστοποίησης μέσω γενετικού αλγορίθμου (GA). Κύριος στόχος ήταν η αποτελεσματική ανίχνευση απειλών σε δίκτυα, δίνοντας έμφαση όχι μόνο στην υψηλή απόδοση, αλλά πρωτίστως στη διαφάνεια και την ερμηνευσιμότητα της διαδικασίας λήψης αποφάσεων του μοντέλου.

Τα αποτελέσματα αξιολόγησης δείχνουν πως ο στόχος της υλοποίησης ενός διαφανούς μοντέλου με ικανοποιητική απόδοση έχει επιτευχθεί σε μεγάλο βαθμό. Το προτεινόμενο μοντέλο ANFIS-GA αξιολογήθηκε με βάση το στατικό σύνολο δεδομένων UNSW NB-15, το οποίο σε σχέση με άλλα γνωστά σύνολα δεδομένων όπως το KDD Cup '99, θεωρείται πιο πρόσφατο και με πιο επικαιροποιημένους και μοντέρνους τύπους επιθέσεων. Το σύνολο δεδομένων που χρησιμοποιήθηκε επέτρεψε την εξέταση και την απόδειξη της αποτελεσματικότητας του μοντέλου τόσο στη δυαδική ταξινόμηση, όσο και στην ταξινόμηση πολλαπλών κλάσεων με διαφορετικού τύπου επιθέσεις. Η συνδρομή του γενετικού αλγορίθμου υπήρξε καθοριστική, καθώς περιόρισε την πολυπλοκότητα, επιλέγοντας το βέλτιστο υποσύνολο χαρακτηριστικών, ενώ ταυτόχρονα πραγματοποίησε βελτιστοποίηση των υπερπαραμέτρων που αφορούν το ANFIS. Ως αποτέλεσμα, η απόδοση του μοντέλου παρουσίασε σημαντική βελτίωση, ενώ η πολυπλοκότητα μειώθηκε.

Ιδιαίτερα σημαντική είναι η ανάδειξη της αξίας της ερμηνευσιμότητας που προσφέρει το ANFIS. Η διαφάνεια του μοντέλου σε ένα πεδίο όπως η κυβερνοασφάλεια διευκολύνει την κατανόηση του πλαισίου λήψης απόφασης. Αυτό καθίσταται εφικτό μέσω των ασαφών κανόνων IF-THEN, με τους οποίους ο αναλυτής μπορεί να εξετάσει τη λογική των αποφάσεων του μοντέλου, κατανοώντας παράλληλα τη συμπεριφορά που τους ενεργοποίησε. Η προτεινόμενη προσέγγιση παρουσίασε υποσχόμενα αποτελέσματα και αποτελεί μια σταθερή και κατάλληλη βάση για περαιτέρω αξιολόγηση σε διαφορετικά σύνολα δεδομένων και περιβάλλοντα πραγματικού χρόνου (real-time).

8.2 Κύρια ερευνητικά ευρήματα

Η εκτενής μελέτη, η πειραματική διαδικασία και η στατιστική ανάλυση των αποτελεσμάτων ανέδειξαν ορισμένα σημαντικά ευρήματα. Η επιλογή χαρακτηριστικών μέσω του γενετικού αλγορίθμου κατάφερε να μειώσει τα αρχικά 202 χαρακτηριστικά σε σημαντικά μικρότερα υποσύνολα. Αυτή η μείωση χαρακτηριστικών βελτίωσε τη συνολική απόδοση του μοντέλου, η οποία είναι στατιστικά αποδεδειγμένη με τους αντίστοιχους ελέγχους. Ο γενετικός αλγόριθμος επέλεξε τα σημαντικότερα χαρακτηριστικά προσφέροντας την απαραίτητη πληροφορία χωρίς επιπλέον «θόρυβο», γεγονός που επιτρέπει στο τελικό μοντέλο ANFIS να παρουσιάσει βελτιωμένα αποτελέσματα. Επίσης, η απουσία των επιπλέον χαρακτηριστικών, τα οποία δεν προσφέρουν σημαντικές πληροφορίες, επέτρεψε τη μείωση της πολυπλοκότητας η οποία μεταφράζεται στους σχετικά μειωμένους χρόνους εκπαίδευσης αλλά και στο μικρότερο υπολογιστικό κόστος που απαιτήθηκε κατά την εκπαίδευση. Οι δύο διαφορετικές προσεγγίσεις που χρησιμοποιήθηκαν για την ελεύθερη επιλογή και την προκαθορισμένου πλήθους επιλογή χαρακτηριστικών επέτρεψαν την εξέταση των μεταξύ τους διαφορών απόδοσης. Η χρήση των δύο προσεγγίσεων και στις δύο ταξινομήσεις ανέδειξε συγκεκριμένα εύρη πλήθους

χαρακτηριστικών στα οποία το προτεινόμενο μοντέλο παρουσιάζει υψηλότερη απόδοση, γεγονός που επαληθεύει πως η λειτουργία του γενετικού αλγορίθμου δεν απέδωσε τυχαία αποτελέσματα.

Η αντιμετώπιση της ακραίας ανισορροπίας κλάσεων, που εμφάνιζε αρχικά το σύνολο δεδομένων, μέσω της τεχνικής SMOTE, επέτρεψε στο μοντέλο να εντοπίζει σπάνιους τύπους επιθέσεων όπως Shellcode και Worms. Αυτοί οι δύο τύποι επιθέσεων εμφάνισαν ικανοποιητικά αποτελέσματα ανάκλησης (recall), ωστόσο παρατηρήθηκε και μεγάλη πτώση της ακρίβειας (precision) λόγω της αλλοίωσης των πιθανοτήτων εμφάνισής τους (prior probability shift). Το συγκεκριμένο φαινόμενο συχνά αποτελεί έναν αποδεκτό συμβιβασμό για ένα IDS, όπου προτεραιότητα είναι η ανίχνευση όσο το δυνατόν περισσότερων επιθέσεων.

Το μοντέλο αναφοράς Random Forest παρουσίασε συγκριτικά υψηλότερα επίπεδα συνολικής απόδοσης βάσει των μετρικών και αισθητά μικρότερο χρόνο εκπαίδευσης. Ωστόσο, η σύγκριση του Random Forest με το προτεινόμενο μοντέλο ανέδειξε τον συμβιβασμό μεταξύ απόδοσης και ερμηνευσιμότητας. Η εσωτερική λογική του Random Forest δεν είναι άμεσα αναγνώσιμη σε μορφή ενός συνόλου κανόνων και η ερμηνεία του συνήθως απαιτεί πρόσθετες τεχνικές επεξήγησης. Αντίθετα, το ANFIS παρέχει εγγενή ερμηνευσιμότητα, επειδή η λειτουργία του μπορεί να εξεταστεί μέσω των ασαφών κανόνων, ενώ και η απόδοσή του σύμφωνα με τα αποτελέσματα των πειραμάτων δεν έχει σημαντική απόκλιση. Επίσης, αξιοσημείωτο είναι το γεγονός πως το προτεινόμενο μοντέλο στη δυαδική ταξινόμηση παρουσίασε μικρότερο πλήθος χαμένων επιθέσεων από το Random Forest, βελτιώνοντας δηλαδή την ικανότητα ανίχνευσης επιθέσεων. Επομένως, το προτεινόμενο μοντέλο φαίνεται να αποτελεί μια υποσχόμενη και αξιόλογη προσέγγιση για περιβάλλοντα στα οποία, πέρα από την απόδοση, δίνεται αυξημένη βαρύτητα στην ερμηνευσιμότητα και στον περιορισμό των μη ανιχνευμένων επιθέσεων.

8.3 Μελλοντικές επεκτάσεις

Πέρα από την ικανοποιητική επίτευξη των στόχων της παρούσας εργασίας, υπάρχουν μερικές μελλοντικές βελτιώσεις που μπορούν να λύσουν προβλήματα και περιορισμούς που ανιχνεύθηκαν. Παρακάτω αναφέρονται μελλοντικές ενέργειες που μπορούν να βελτιώσουν και να αναβαθμίσουν τη συγκεκριμένη εργασία, χρησιμοποιώντας την ως υπόβαθρο. Οι προτεινόμενες μελλοντικές επεκτάσεις αναφέρονται στα εξής σημεία:

- **Υλικό (Hardware):** Ένας από τους σημαντικότερους περιορισμούς της εργασίας ήταν η περιορισμένη υπολογιστική ισχύς που ήταν διαθέσιμη για την υλοποίηση των πειραμάτων. Η αξιοποίηση ισχυρού υπολογιστικού υλικού (hardware), όπως μονάδα επεξεργασίας γραφικών (GPU) και νεότερης τεχνολογίας επεξεργαστή (CPU), θα επιτρέψει τη δραστική μείωση του χρόνου εκπαίδευσης. Επίσης, η ταχύτερη εκπαίδευση του εσωτερικού ANFIS θα επιτρέψει τον ορισμό μεγαλύτερων πληθυσμών και περισσότερων γενεών του γενετικού αλγορίθμου. Αυτό θα δημιουργήσει την ικανότητα για εξερεύνηση ενός πολύ μεγαλύτερου χώρου αναζήτησης που ενδεχομένως να οδηγήσει σε ακόμη καλύτερη παραμετροποίηση όλου του μοντέλου με συνολικά βελτιωμένα αποτελέσματα.
- **NSGA-II:** Η υλοποίηση του γενετικού αλγορίθμου της εργασίας ακολουθεί μονοκριτηριακή βελτιστοποίηση (single-objective optimization). Δηλαδή, η αξιολόγηση του ατόμου-χρωμοσώματος γίνεται με βάση την τιμή καταλληλότητας. Στην πλειοψηφία των πειραμάτων που πραγματοποιήθηκαν, η τιμή καταλληλότητας εξαρτάται από το F1-score. Ωστόσο, στο Σενάριο 2 και την προσέγγιση ελεύθερης επιλογής χαρακτηριστικών, η τιμή καταλληλότητας, όπως απεικονίζεται και στη Σχέση 5.1, εξαρτάται τόσο από το F1-score όσο και από το *sparsity_bonus*. Η συνάρτηση της μεταβλητής *sparsity_bonus* εμφανίζεται στη σχέση 8.1:

$$sparsity_bonus = \frac{0,05}{\text{πλήθος επιλεγμένων χαρακτηριστικών}} \quad (8.1)$$

Η τιμή του αριθμητή προέκυψε έπειτα από εκτενείς πειραματικές εκτελέσεις του γενετικού αλγορίθμου. Η συγκεκριμένη τιμή κρίθηκε βέλτιστη καθώς αποδίδει σωστή αύξηση της τιμής καταλληλότητας στο πλαίσιο συμβιβασμού ανάμεσα στο F1-score και το πλήθος χαρακτηριστικών. Έτσι, αυτή η μεταβλητή επιβράβευσης ευνοεί εκείνα τα άτομα που διαθέτουν σημαντικά λιγότερα χαρακτηριστικά, επιτρέποντάς τους να ανταγωνιστούν άτομα με οριακά υψηλότερες τιμές F1-score αλλά και πολύ μεγαλύτερα επίπεδα πολυπλοκότητας. Η επαναληπτική διαδικασία της «χειροκίνητης» προσαρμογής του αριθμητή για τη μεταβλητή επιβράβευσης της σχέσης 8.1, μπορεί να αντικατασταθεί από την εφαρμογή της τεχνικής NSGA-II. Με αυτή, μπορεί να γίνει ταυτόχρονα βελτιστοποίηση και των δύο κριτηρίων μεγιστοποίησης ακρίβειας και ελαχιστοποίησης πολυπλοκότητας. Αντί να εξάγει μια μοναδική λύση βάσει μιας προκαθορισμένης μαθηματικής εξίσωσης, το NSGA-II παράγει ένα μέτωπο Pareto (Pareto Front) [60]. Αυτό προσφέρει μια ποικιλία βέλτιστων λύσεων, απαλλάσσοντας την ερευνητική διαδικασία από τη χρονοβόρα εύρεση σταθερών παραμέτρων επιβράβευσης. Επίσης, παρέχει στον χρήστη-αναλυτή την ελευθερία να επιλέξει τον ιδανικό συμβιβασμό για τις εκάστοτε επιχειρησιακές ανάγκες.

- **Αξιολόγηση σε δεδομένα πραγματικού χρόνου (Real-Time IDS):** Το προτεινόμενο μοντέλο εκπαιδεύτηκε και αξιολογήθηκε σε ένα στατικό σύνολο δεδομένων, το UNSW NB-15. Μια μελλοντική επέκταση της εργασίας αποτελεί η αξιολόγηση του ήδη εκπαιδευμένου μοντέλου σε περιβάλλον συνεχούς ροής δεδομένων. Σε πραγματικές συνθήκες, ένα IDS καλείται να ανιχνεύει με μικρή καθυστέρηση (latency) οποιαδήποτε εισβολή και να διαχειρίζεται τον δυναμικό και απρόβλεπτο όγκο δεδομένων. Η μετάβαση σε επιχειρησιακή λειτουργία πραγματικού χρόνου θα επιτρέψει την εξέταση της ανθεκτικότητας και της απόδοσης του προτεινόμενου υβριδικού μοντέλου. Έτσι, μπορεί να εξεταστεί η συμπεριφορά απέναντι στον φυσικό «θόρυβο» του δικτύου και ο χρόνος απόκρισης του μοντέλου. Η υλοποίηση αυτού του σεναρίου απαιτεί τη σύλληψη δεδομένων δικτύου και προεπεξεργασία αυτών σε ελάχιστο χρόνο ώστε να εισαχθούν στο προτεινόμενο μοντέλο με στόχο την άμεση ανίχνευση ενδεχόμενων εισβολών.
- **Αξιολόγηση με άλλα σύνολα δεδομένων:** Το σύνολο δεδομένων που χρησιμοποιήθηκε είναι το UNSW NB-15, το οποίο αποτελεί μια σύγχρονη και ρεαλιστική αναπαράσταση της δικτυακής κίνησης. Ωστόσο, προκειμένου να πιστοποιηθεί η ικανότητα γενίκευσης και απόδοσης του μοντέλου απέναντι σε διαφορετικές κλάσεις, μπορούν μελλοντικά να χρησιμοποιηθούν και άλλα αξιολογικά σύνολα δεδομένων. Σύνολα δεδομένων όπως το CIC-IDS-2017 ή CSE-CIC-IDS2018 αποτελούν σύγχρονα και πολύπλοκα σύνολα, με στόχο την εκπαίδευση μοντέλων για τη χρήση τους σε περιβάλλοντα IDS [61]. Επιπρόσθετα, το προτεινόμενο υβριδικό μοντέλο θα μπορούσε να εφαρμοστεί σε σύνολα δεδομένων που αφορούν διαφορετικούς τομείς, όπως το Διαδίκτυο των Πραγμάτων (Internet of Things - IoT). Σε τέτοια περιβάλλοντα, η ικανότητα του μοντέλου να διατηρεί υψηλή απόδοση με παράλληλη μείωση της πολυπλοκότητας (μέσω της επιλογής χαρακτηριστικών) έχει καθοριστικό ρόλο, καθώς συμβάλλει άμεσα στη μείωση του υπολογιστικού κόστους και της κατανάλωσης ενέργειας.
- **Πειραματική επαλήθευση απόδοσης σε Zero-day επιθέσεις:** Όπως αναφέρθηκε και στο Κεφάλαιο 7, το ANFIS έχει θεωρητικό πλεονέκτημα συγκριτικά με το Random Forest για την ανίχνευση Zero-day επιθέσεων. Η επαλήθευση αυτής της θεωρητικής συμπεριφοράς αποτελεί μια επιπλέον μελλοντική επέκταση. Η εκπαίδευση του μοντέλου σε γνωστές κατηγορίες επιθέσεων και η αξιολόγησή του με αντίστοιχο υποσύνολο το οποίο θα περιέχει άγνωστες κατηγορίες επιθέσεων, θα δώσει τη δυνατότητα τεκμηρίωσης της ικανότητας του μοντέλου να λειτουργεί ως ένας αξιόπιστος μηχανισμός εντοπισμού άγνωστων ανωμαλιών σε διαρκώς μεταβαλλόμενα πλαίσια κυβερνοαπειλών.
- **Αναβάθμιση από IDS σε IPS:** Το προτεινόμενο μοντέλο λειτουργεί ως ένα NIDS. Αυτό σημαίνει ότι η λειτουργία του περιορίζεται στην παρακολούθηση της δικτυακής κίνησης, την ανάλυση και την παραγωγή προβλέψεων και ειδοποιήσεων (alerts) σε περίπτωση αναγνώρισης κακόβουλης δραστηριότητας. Τα IDS δεν παρεμβαίνουν στη λειτουργία των συστημάτων στα οποία λειτουργούν. Μια κρίσιμη μελλοντική επέκταση αποτελεί η αναβάθμιση του Συστήματος Ανίχνευσης Εισβολών (IDS) σε ένα Σύστημα Πρόληψης Εισβολών (IPS). Αρχικά, πριν από αυτή την αναβάθμιση πρέπει να προηγηθεί η εύρεση λύσης στο πρόβλημα της αλλοίωσης των

πιθανοτήτων εμφάνισης εγγραφών (prior probability shift) που δημιουργείται λόγω της προεπεξεργασίας των δεδομένων με την τεχνική SMOTE. Όπως αναφέρει το [57], υπάρχουν συγκεκριμένες τεχνικές που μειώνουν το συγκεκριμένο πρόβλημα. Η επίλυση αυτού του προβλήματος θα επιτρέψει την περαιτέρω βελτίωση της ακρίβειας πρόβλεψης των σπάνια εμφανιζόμενων κλάσεων του συνόλου δεδομένων. Η μεγαλύτερη ακρίβεια πρόβλεψης απαιτείται σε περιβάλλοντα IPS, διότι για διαφορετικούς τύπους επιθέσεων χρησιμοποιούνται και διαφορετικές τεχνικές αντιμετώπισης αυτών. Στο πλαίσιο του IPS, κατά την ανίχνευση ενός μοτίβου επίθεσης, το σύστημα δεν θα περιορίζεται στην ειδοποίηση, αλλά θα λαμβάνει αυτοματοποιημένες αποφάσεις αντιμετώπισης σε πραγματικό χρόνο. Έτσι, η συγκεκριμένη αναβάθμιση μπορεί να προσφέρει ένα ολοκληρωμένο σύστημα το οποίο θα αναγνωρίζει και αντιμετωπίζει απειλές με υψηλή απόδοση.

- **Ενσωμάτωση επεξηγήσιμης τεχνητής νοημοσύνης (XAI) για διασταύρωση και επαλήθευση αποτελεσμάτων:** Οι μέθοδοι XAI χρησιμοποιούνται συνήθως σε black-box μοντέλα μηχανικής μάθησης, ώστε να προσφέρουν την ικανότητα επεξηγησιμότητας που υστερούν. Το ANFIS αποτελεί ένα μοντέλο το οποίο παρέχει ερμηνευσιμότητα μέσω της εξαγωγής των ασαφών κανόνων. Ωστόσο, μια ενδιαφέρουσα μελλοντική επέκταση αποτελεί ο συνδυασμός του προτεινόμενου μοντέλου με post-hoc μεθόδους επεξηγήσιμης τεχνητής νοημοσύνης (XAI), όπως LIME ή SHAP. Η προσέγγιση αυτή μπορεί να δημιουργήσει έναν μηχανισμό διασταύρωσης αποτελεσμάτων, με στόχο την αύξηση της αξιοπιστίας και της διαφάνειας του μοντέλου. Συγκεκριμένα, η ερμηνεία που παρέχουν οι ασαφείς κανόνες μπορεί να συγκριθεί και να επιβεβαιωθεί από τα αποτελέσματα που παράγουν οι XAI μέθοδοι με τους δικούς τους τρόπους επεξήγησης. Η συγκεκριμένη προσέγγιση, σύμφωνα με το [62], επιτρέπει την αυτοματοποιημένη επαλήθευση των αποφάσεων, προσφέροντας ένα σύστημα όπου η διασταύρωση δύο μεθόδων (ασαφών κανόνων ANFIS και XAI) αυξάνει την αξιοπιστία και την εμπιστοσύνη προς το σύστημα.
- **Ανάπτυξη γραφικού περιβάλλοντος χρήστη (GUI):** Η παρούσα εργασία επικεντρώθηκε στο κομμάτι της υλοποίησης και της αξιολόγησης του προτεινόμενου μοντέλου. Μια απαραίτητη προσθήκη για την επιχειρησιακή του αξιοποίηση σε πραγματικές συνθήκες είναι η ανάπτυξη ενός ολοκληρωμένου γραφικού περιβάλλοντος χρήστη (GUI). Μέσω αυτού, οι χρήστες θα έχουν τη δυνατότητα παρακολούθησης της δικτυακής κίνησης σε πραγματικό χρόνο και των ειδοποιήσεων σχετικά με απειλές. Σημαντικό είναι πως η χρήση του GUI θα αξιοποιήσει την ερμηνευσιμότητα που προσφέρει το προτεινόμενο μοντέλο. Κατά την ανίχνευση μιας απειλής, το περιβάλλον μπορεί να εμφανίζει τους εκάστοτε ασαφείς κανόνες που ενεργοποιήθηκαν. Η οπτικοποίηση της λειτουργίας του μοντέλου θα επιτρέψει τη μείωση του χρόνου διερεύνησης και κατανόησης της κατάστασης, ενώ παράλληλα θα διευκολύνει τη λήψη κρίσιμων αποφάσεων από τους χρήστες ώστε να προβούν άμεσα στις κατάλληλες ενέργειες αντιμετώπισης της απειλής.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] M. Aljanabi, M. A. Ismail, and A. H. Ali, "Intrusion Detection Systems, Issues, Challenges, and Needs," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, 2021.
- [2] A. Thakkar and R. Lohiya, "A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions," *Artificial Intelligence Review*, vol. 55, pp. 453-563, 2022.
- [3] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *Journal of Information Security and Applications*, vol. 50, p. 102419, Feb. 2020.
- [4] R. da Silveira Lopes, J. C. Duarte and R. R. Goldschmidt, "False Positive Identification in Intrusion Detection Using XAI," *IEEE Latin America Transactions*, vol. 21, no. 6, pp. 745-751, June 2023.
- [5] K. Scarfone and P. Mell, "Guide to Intrusion Detection and Prevention Systems (IDPS)," *NIST*, Feb. 2007.
- [6] S. Neupane et al., "Explainable Intrusion Detection Systems (X-IDS): A Survey of Current Methods, Challenges, and Opportunities," in *IEEE Access*, vol. 10, pp. 112392-112415, 2022.
- [7] D. E. Denning, "An Intrusion-Detection Model," in *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 222-232, Feb. 1987.
- [8] Y. Otoum and A. Nayak, "AS-IDS: Anomaly and Signature Based IDS for the Internet of Things," *Journal of Network and Systems Management*, vol. 29, no. 3, Mar. 2021.
- [9] M. Fernández-Delgado, E. Cernadas, S. Barro, D. Amorim, and A. Fernández-Delgado, "Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?" *Journal of Machine Learning Research*, vol. 15, pp. 3133–3181, 2014.
- [10] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, June 1965.
- [11] E. H. Mamdani, "Application of fuzzy algorithms for control of simple dynamic plant," *Proceedings of the Institution of Electrical Engineers*, vol. 121, no. 12, pp. 1585–1588, Dec. 1974.
- [12] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 1, pp. 116-132, Jan.-Feb. 1985.
- [13] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [14] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2–3, pp. 191–203, Jan. 1984.
- [15] M. A. Elsabagh, O. E. Emam, M. G. Gafar, and T. Medhat, "Handling uncertainty issue in software defect prediction utilizing a hybrid of ANFIS and turbulent flow of water optimization algorithm," *Neural Computing and Applications*, vol. 36, no. 9, pp. 4583–4602, Dec. 2023.

- [16] H. Enayatollahi, P. Fussey, and B. K. Nguyen, "Modelling evaporator in organic Rankine cycle using hybrid GD-LSE ANFIS and PSO ANFIS techniques," *Thermal Science and Engineering Progress*, vol. 19, no. 5, p. 100570, May 2020.
- [17] M. G. Altarabichi, P. Hilber, S. Pashami, and P. S. Mashhadi, "Fast Genetic Algorithm for feature selection — A qualitative approximation approach," *Expert Systems with Applications*, vol. 211, p. 118528, Aug. 2022.
- [18] J. Demšar, "Statistical Comparisons of Classifiers over Multiple Data Sets," *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.
- [19] P. Parkar and A. Bilimoria, "A Survey on Cyber Security IDS using ML Methods," *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2021.
- [20] I. Ahmad, M. Basher, M. J. Iqbal, and A. Rahim, "Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection," in *IEEE Access*, vol. 6, pp. 33789–33795, 2018.
- [21] A. A. Aburomman and M. Bin Ibne Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," *Applied Soft Computing*, vol. 38, pp. 360–372, Jan. 2016.
- [22] J. Lansky et al., "Deep Learning-Based Intrusion Detection Systems: A Systematic Review," in *IEEE Access*, vol. 9, pp. 101574-101599, 2021.
- [23] N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41-50, Feb. 2018.
- [24] J. Mar, Y. -C. Yeh and I. -F. Hsiao, "An ANFIS-IDS against deauthentication DOS attacks for a WLAN," *2010 International Symposium On Information Theory & Its Applications*, Taichung, Taiwan, 2010, pp. 548-553.
- [25] J. Liu, W. Yinchai, T. C. Siong, X. Li, L. Zhao, and F. Wei, "A hybrid interpretable deep structure based on adaptive neuro-fuzzy inference system, decision tree, and K-means for intrusion detection," *Scientific Reports*, vol. 12, no. 1, Dec. 2022.
- [26] R. Khushbu, K. Megha, "Reviewing a New Optimized an ANFIS-Based Framework for Detecting Intrusion Detection System with Machine Learning Algorithms (Deep Learning Algorithm)," *International Journal of Engineering and Advanced Technology*, pp. 35-42, 2023.
- [27] B. Haznedar, "Training ANFIS Using Genetic Algorithm for Dynamic Systems Identification," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 4, no. Special Issue-1, pp. 44–47, Dec. 2016.
- [28] S. Ansari, K. A. Alnajjar, S. Abdallah, M. Saad and A. A. El-Moursy, "Parameter Tuning of MLP, RBF, and ANFIS Models Using Genetic Algorithm in Modeling and Classification Applications," *2021 International Conference on Information Technology (ICIT)*, pp. 660-666, Amman, Jordan, 2021.
- [29] B. Haznedar and H. C. Kilinc, "A Hybrid ANFIS-GA Approach for Estimation of Hydrological Time Series," *Water Resources Management*, vol. 36, no. 12, pp. 4819–4842, Aug. 2022.

- [30] M. A. Soliman, H. M. Hasanien, H. Z. Azazi, E. E. El-kholy, and S. A. Mahmoud, "Hybrid ANFIS-GA-based control scheme for performance enhancement of a grid-connected wind generator," *IET Renewable Power Generation*, vol. 12, no. 7, pp. 832–843, Apr. 2018.
- [31] Z. Alameer, M. A. Elaziz, A. A. Ewees, H. Ye, and Z. Jianhua, "Forecasting Copper Prices Using Hybrid Adaptive Neuro-Fuzzy Inference System and Genetic Algorithms," *Natural Resources Research*, vol. 28, no. 4, pp. 1385–1401, Mar. 2019.
- [32] M. Esmaeili and M. Zahedi, "Spam detection by ANFIS with feature selection by GA," *Elixir International Journal*, vol. 77, p. 28908, Nov. 2014.
- [33] M. Hamdan, M. Assora, and M. Dakkak, "Performance Evaluation of a Genetic Neuro-Fuzzy Intrusion Detection System across Multiple Datasets," *Journal of Information Security*, vol. 17, no. 03, pp. 209–220, Jan. 2026.
- [34] A. Alazab, M. Hobbs, J. Abawajy and M. Alazab, "Using feature selection for intrusion detection system," *2012 International Symposium on Communications and Information Technologies (ISCIT)*, pp. 296-301, Gold Coast, QLD, Australia, 2012.
- [35] T. Markovic, P. Lidholm, P. E. Strandberg, and M. Leon, "Feature Selection Using Genetic Algorithm for Intrusion Detection on Resource-Constrained Edge Devices," *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1415–1423, July 2025.
- [36] H. Bakır and Ö. Ceviz, "Empirical Enhancement of Intrusion Detection Systems: A Comprehensive Approach with Genetic Algorithm-based Hyperparameter Tuning and Hybrid Feature Selection," *Arabian Journal for Science and Engineering*, Apr. 2024.
- [37] M. Pérez Sevilla, J. Á. Olivas Varela, Á. Herrero Cosio, D. Urda Muñoz, and J. A. Rincón Arango, "An explainable neuro-fuzzy intrusion detection system for industrial IoT," *Internet of Things*, vol. 38, p. 101989, June 2026.
- [38] K. Kulikauskaitė and D. Mažeika, "Neuro-Fuzzy Modeling of Decision-Making in Cyber Defense Exercises Using ANFIS and Synthetic Data Augmentation," *Applied Sciences*, 2026.
- [39] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," *2015 Military Communications and Information Systems Conference (MilCIS)*, pp. 1-6, Canberra, ACT, Australia, 2015.
- [40] N. Moustafa and J. Slay, "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective*, vol. 25, no. 1–3, pp. 18–31, Jan. 2016.
- [41] N. Moustafa, J. Slay and G. Creech, "Novel Geometric Area Analysis Technique for Anomaly Detection Using Trapezoidal Area Estimation on Large-Scale Networks," in *IEEE Transactions on Big Data*, vol. 5, no. 4, pp. 481-494, 1 Dec. 2019.
- [42] N. Moustafa, G. Creech, and J. Slay, "Big Data Analytics for Intrusion Detection System: Statistical Decision-Making Using Finite Dirichlet Mixture Models," *Data Analytics and Decision Support for Cybersecurity*, pp. 127–156, 2017.
- [43] M. Sarhan, S. Layeghy, N. Moustafa, and M. Portmann, "NetFlow Datasets for Machine Learning-Based Network Intrusion Detection Systems," *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (LNICST)*, vol. 371, pp. 117–135, 2021.

- [44] O. Ebrahim, S. Dowaji, and S. Alhammoud, "A lightweight machine learning approach for DDoS detection and classification," *Scientific Reports*, Apr. 2026.
- [45] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinformatics*, vol. 14, no. 1, Mar. 2013.
- [46] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 16, pp. 321–357, June 2002.
- [47] V. T. Nguyen, "X-ANFIS: An Extensible and Cross-Learning ANFIS Framework for Machine Learning Tasks," *Figshare*, Jan. 2025.
- [48] N. V. Thieu and S. A. Mirjalili, "MEALPY: An open-source library for latest meta-heuristic algorithms in Python," *Journal of Systems Architecture*, vol. 139, pp. 102871–102871, June 2023.
- [49] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [50] F.-M. De Rainville, F.-A. Fortin, M.-A. Gardner, M. Parizeau, and C. Gagné, "DEAP: A Python framework for Evolutionary Algorithms," *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 85–92, July 2012.
- [51] D. Berman, A. Buczak, J. Chavis, and C. Corbett, "A Survey of Deep Learning Methods for Cyber Security," *Information*, vol. 10, no. 4, p. 122, Apr. 2019.
- [52] S. García, F. Herrera, and H. Es, "An Extension on 'Statistical Comparisons of Classifiers over Multiple Data Sets' for all Pairwise Comparisons," *Journal of Machine Learning Research*, vol. 9, pp. 2677–2694, 2008.
- [53] G. F. Lyon, *Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning*. Sunnyvale, CA: Insecure.com LLC, 2009.
- [54] Y. Zhang and V. Paxson, "Detecting Backdoors," in *Proc. 9th USENIX Security Symposium*, 2000.
- [55] S. Wali, Y. A. Farrukh, and I. Khan, "Explainable AI and Random Forest based reliable intrusion detection system," *Computers & Security*, vol. 157, p. 104542, Oct. 2025.
- [56] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, Apr. 2018.
- [57] M. Wang, K. Zheng, Y. Yang and X. Wang, "An Explainable Machine Learning Framework for Intrusion Detection Systems," in *IEEE Access*, vol. 8, pp. 73127-73141, 2020.
- [58] A. Isijola, E. Afuadajo, M. Asefon, U. Ogude, J. Akande, and P. Joseph, "Design and Evaluation of a Fuzzy Logic Based Intrusion Detection System for Network Security," *International Journal of Artificial Intelligence*, vol. 12, no. 2, pp. 81–100, Dec. 2025.
- [59] R. K. Sharma, D. K. Singh, A. Kumar, and A. P. Burnwal, "Use of Fuzzy Logic in Cyber Security System: Cyber Threat Intelligence," *International Journal of Students Research in Technology & Management*, vol. 11, no. 3, pp. 10–19, Aug. 2023.

- [60] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," in *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, April 2002.
- [61] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in *Proc. 4th International Conference on Information Systems Security and Privacy (ICISSP)*, pp. 108-116, 2018.
- [62] Y. V. Trofimov, A. D. Lebedev, A. S. Ilin, and A. N. Averkin, "Verified Explainability Core: A GD–ANFIS/SHAP Hybrid Architecture for XAI 2.0," *Automatic Documentation and Mathematical Linguistics*, vol. 59, no. S5, pp. S469-S478, Dec. 2025.

ΠΑΡΑΡΤΗΜΑ Α : Κανόνες IF-THEN

[KANONAS 1]	[KANONAS 2]	[KANONAS 8]	[KANONAS 9]
IF dur ≈ -0.8206 AND sbytes ≈ -0.0360 AND dbytes ≈ -0.3021 AND sttl ≈ 1.2265 AND dttl ≈ -0.1368 AND sloss ≈ 0.2947 AND dloss ≈ 1.6966 AND Sload ≈ 0.1220 AND Dload ≈ -0.0062 AND Spkts ≈ -0.1293 AND Dpkts ≈ 1.7086 AND swin ≈ -0.8766 AND dwin ≈ -0.7650 AND stcpb ≈ -0.0186 AND dtcpb ≈ -0.0294 AND smeansz ≈ -0.0403 AND dmeansz ≈ -0.0181 AND trans_depth ≈ -0.0177 AND res_body_len ≈ -0.0184 AND Sflt ≈ -0.0193 AND Dflt ≈ -0.0179 AND Sintpkt ≈ -0.0174 AND Dintpkt ≈ -0.0174 AND tcprtt ≈ -0.0176 AND synack ≈ -0.0178 AND ackdat ≈ -0.0177 AND is_sm_ips_ports ≈ -0.0188 AND ct_state_ttl ≈ -0.0178 AND ct_flw_http_mthd ≈ -0.0179 AND is_frp_login ≈ -0.0188 AND ct_frp_cmd ≈ -0.0175 AND ct_srv_src ≈ -0.0183 AND ct_srv_dst ≈ -0.0242 AND ct_dst_ltm ≈ -0.0171 AND ct_src_ltm ≈ -0.0180 AND ct_src_dport_ltm ≈ -0.0251 AND ct_dst_sport_ltm ≈ 0.0000 AND ct_dst_src_ltm ≈ -0.0175 AND proto_3pc ≈ -0.0194 AND proto_a/n ≈ -0.0179 AND proto_aes-sp3-d ≈ -0.0846 AND proto_any ≈ -0.0169	IF dur ≈ 0.8124 AND sbytes ≈ -0.0442 AND dbytes ≈ -0.3391 AND sttl ≈ 1.2572 AND dttl ≈ 0.0652 AND sloss ≈ -0.3177 AND dloss ≈ 1.7684 AND Sload ≈ 0.0567 AND Dload ≈ -0.0278 AND Spkts ≈ -0.0939 AND Dpkts ≈ 1.6800 AND swin ≈ -0.9448 AND dwin ≈ -0.8319 AND stcpb ≈ -0.0193 AND dtcpb ≈ -0.0316 AND smeansz ≈ -0.0408 AND dmeansz ≈ -0.0190 AND trans_depth ≈ -0.0182 AND res_body_len ≈ -0.0192 AND Sflt ≈ -0.0200 AND Dflt ≈ -0.0187 AND Sintpkt ≈ -0.0179 AND Dintpkt ≈ -0.0181 AND tcprtt ≈ -0.0183 AND synack ≈ -0.0186 AND ackdat ≈ -0.0183 AND is_sm_ips_ports ≈ -0.0194 AND ct_state_ttl ≈ -0.0184 AND ct_flw_http_mthd ≈ -0.0187 AND is_frp_login ≈ -0.0199 AND ct_frp_cmd ≈ -0.0183 AND ct_srv_src ≈ -0.0191 AND ct_srv_dst ≈ -0.0255 AND ct_dst_ltm ≈ -0.0177 AND ct_src_ltm ≈ -0.0188 AND ct_src_dport_ltm ≈ -0.0266 AND ct_dst_sport_ltm ≈ 0.0000 AND ct_dst_src_ltm ≈ -0.0183 AND proto_3pc ≈ -0.0201 AND proto_a/n ≈ -0.0187 AND proto_aes-sp3-d ≈ -0.0873 AND proto_any ≈ -0.0171	IF dur ≈ 0.7641 AND sbytes ≈ -0.0679 AND dbytes ≈ -0.3261 AND sttl ≈ -0.6614 AND dttl ≈ 1.7964 AND sloss ≈ -0.4858 AND dloss ≈ -0.1259 AND Sload ≈ -0.0771 AND Dload ≈ -0.0214 AND Spkts ≈ 0.8202 AND Dpkts ≈ -0.1491 AND swin ≈ -0.8537 AND dwin ≈ -0.7673 AND stcpb ≈ -0.0169 AND dtcpb ≈ -0.0239 AND smeansz ≈ -0.0401 AND dmeansz ≈ -0.0162 AND trans_depth ≈ -0.0163 AND res_body_len ≈ -0.0164 AND Sflt ≈ -0.0175 AND Dflt ≈ -0.0159 AND Sintpkt ≈ -0.0163 AND Dintpkt ≈ -0.0157 AND tcprtt ≈ -0.0159 AND synack ≈ -0.0158 AND ackdat ≈ -0.0161 AND is_sm_ips_ports ≈ -0.0173 AND ct_state_ttl ≈ -0.0163 AND ct_flw_http_mthd ≈ -0.0158 AND is_frp_login ≈ -0.0162 AND ct_frp_cmd ≈ -0.0156 AND ct_srv_src ≈ -0.0164 AND ct_srv_dst ≈ -0.0213 AND ct_dst_ltm ≈ -0.0156 AND ct_src_ltm ≈ -0.0161 AND ct_src_dport_ltm ≈ -0.0226 AND ct_dst_sport_ltm ≈ 0.0000 AND ct_dst_src_ltm ≈ -0.0155 AND proto_3pc ≈ -0.0179 AND proto_a/n ≈ -0.0162 AND proto_aes-sp3-d ≈ -0.0854 AND proto_any ≈ -0.0169	IF dur ≈ -0.6298 AND sbytes ≈ 0.0437 AND dbytes ≈ 0.1012 AND sttl ≈ 0.7172 AND dttl ≈ 0.1554 AND sloss ≈ 0.5580 AND dloss ≈ 0.0750 AND Sload ≈ 0.0543 AND Dload ≈ -0.0010 AND Spkts ≈ -0.5485 AND Dpkts ≈ 0.0947 AND swin ≈ -0.6444 AND dwin ≈ -0.6576 AND stcpb ≈ -0.0132 AND dtcpb ≈ -0.0120 AND smeansz ≈ -0.0371 AND dmeansz ≈ -0.0120 AND trans_depth ≈ -0.0135 AND res_body_len ≈ -0.0121 AND Sflt ≈ -0.0137 AND Dflt ≈ -0.0115 AND Sintpkt ≈ -0.0138 AND Dintpkt ≈ -0.0122 AND tcprtt ≈ -0.0123 AND synack ≈ -0.0114 AND ackdat ≈ -0.0127 AND is_sm_ips_ports ≈ -0.0142 AND ct_state_ttl ≈ -0.0131 AND ct_flw_http_mthd ≈ -0.0116 AND is_frp_login ≈ -0.0107 AND ct_frp_cmd ≈ -0.0114 AND ct_srv_src ≈ -0.0123 AND ct_srv_dst ≈ -0.0145 AND ct_dst_ltm ≈ -0.0126 AND ct_src_ltm ≈ -0.0122 AND ct_src_dport_ltm ≈ -0.0147 AND ct_dst_sport_ltm ≈ 0.0000 AND ct_dst_src_ltm ≈ -0.0115 AND proto_3pc ≈ -0.0147 AND proto_a/n ≈ -0.0125 AND proto_aes-sp3-d ≈ -0.0645 AND proto_any ≈ -0.0154
THEN => Πρόβλεψη Κάθους; Generic	THEN => Πρόβλεψη Κάθους; Fuzzers	THEN => Πρόβλεψη Κάθους; Backdoors	THEN => Πρόβλεψη Κάθους; Reconnaissance

Σχήμα Π.1: Οι ολόκληροι κανόνες που αναλύθηκαν στο Κεφάλαιο 6