

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανίχνευση Οικονομικής Απάτης και Λογιστικών
Ανωμαλιών με τη Χρήση Υβριδικών Μοντέλων Μη-
Επιβλεπόμενης Μάθησης



Του φοιτητή
Κυριάκου Παπαθεοδωρόπουλου
Αρ. Μητρώου: 2020/133

Επιβλέπων
Παναγιώτης Αδαμίδης
Καθηγητής

Θεσσαλονίκη, Σεπτέμβριος 2026

Τίτλος Δ.Ε.: Ανίχνευση Οικονομικής απάτης και λογιστικών ανωμαλιών με τη χρήση Υβριδικών μοντέλων μη επιβλεπόμενης μάθησης

Κωδικός Δ.Ε.: 26189

Ονοματεπώνυμο φοιτητή: Κυριάκος Παπαθεοδωρόπουλος

Ονοματεπώνυμο εισηγητή: Παναγιώτης Αδαμίδης

Ημερομηνία ανάληψης Δ.Ε.: 14-3-2026

Ημερομηνία περάτωσης Δ.Ε.: 5-9-2026

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 Κυριάκου Παπαθεοδωρόπουλου 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 απαραίτητα και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Αν ξέραμε τι είναι αυτό που κάνουμε δεν θα ονομαζόταν έρευνα, έτσι δεν είναι;»

Albert Einstein

Πρόλογος

Η επιλογή της παρούσας διπλωματικής εργασίας προέκυψε από το ενδιαφέρον μου για τη μηχανική μάθηση, την ανάλυση οικονομικών δεδομένων και την αντιμετώπιση σύνθετων προβλημάτων ανίχνευσης απάτης σε επιχειρησιακά πληροφοριακά συστήματα. Μέσω της παρούσας εργασίας είχα τη δυνατότητα να μελετήσω σύγχρονες τεχνικές ανάλυσης δεδομένων και ανίχνευσης ανωμαλιών, καθώς και να διερευνήσω τον τρόπο με τον οποίο ο συνδυασμός διαφορετικών προσεγγίσεων μπορεί να ενισχύσει την αξιοπιστία ενός συστήματος εντοπισμού ύποπτων συναλλαγών. Παράλληλα, η αξιοποίηση ενός μεγάλου και σύνθετου συνόλου δεδομένων συνέβαλε στην καλύτερη κατανόηση των προκλήσεων που παρουσιάζουν τα έντονα ανισόρροπα δεδομένα.

Η εκπόνηση της εργασίας ενίσχυσε τις γνώσεις μου στην προεπεξεργασία δεδομένων, στον σχεδιασμό πειραμάτων, στην αξιολόγηση μοντέλων και στην κριτική ερμηνεία των αποτελεσμάτων, προσφέροντάς μου παράλληλα σημαντική εμπειρία στην εφαρμοσμένη έρευνα.

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει την ανίχνευση οικονομικής απάτης και λογιστικών ανωμαλιών με τη χρήση μοντέλων μη επιβλεπόμενης μάθησης και τεχνικών συνδυασμού. Αξιολογούνται τέσσερις διαφορετικοί ανιχνευτές ανωμαλιών, τα Isolation Forest, Local Outlier Factor, Histogram-Based Outlier Score και Autoencoder, με στόχο τη διερεύνηση τόσο της ατομικής τους απόδοσης όσο και της μεταξύ τους συμπληρωματικότητας. Η πειραματική αξιολόγηση πραγματοποιείται αρχικά στο σύνολο δεδομένων IEEE-CIS Fraud Detection. Οι διαθέσιμες ετικέτες δεν χρησιμοποιούνται ως στόχος επιβλεπόμενης μάθησης για την εκπαίδευση των βασικών μοντέλων, αλλά αξιοποιούνται για την ποσοτική αξιολόγηση των αποτελεσμάτων και την επιλογή των τελικών πειραματικών διαμορφώσεων. Ειδικά στην τελική διαμόρφωση του Autoencoder, χρησιμοποιούνται επιπλέον για τον σχηματισμό συνόλου εκπαίδευσης που περιλαμβάνει μόνο κανονικές παρατηρήσεις, με αποτέλεσμα η συγκεκριμένη εφαρμογή να προσεγγίζει στρατηγική μίας κλάσης. Εξετάζονται οι τεχνικές συνδυασμού OR/Union, πλειοψηφικής ψηφοφορίας και σταθμισμένου συνδυασμού βαθμολογιών, ενώ η συνεισφορά των επιμέρους μοντέλων διερευνάται μέσω τιμών Shapley. Τα αποτελέσματα δείχνουν ότι το Isolation Forest παρουσιάζει την ισχυρότερη συνολική συμπεριφορά στο επιλεγμένο σημείο λειτουργίας, ενώ ο Autoencoder επιτυγχάνει την υψηλότερη AUPRC/AP. Οι τεχνικές συνδυασμού δεν υπερέχουν καθολικά των μεμονωμένων μοντέλων, αλλά διαμορφώνουν διαφορετικούς συμβιβασμούς μεταξύ Precision και Recall. Ως τελική διαμόρφωση επιλέγεται ο σταθμισμένος συνδυασμός IF-Dominant. Το τελικό πλαίσιο εφαρμόζεται σε 65.479 πραγματικές λογιστικές εγγραφές ERP χωρίς διαθέσιμες ετικέτες απάτης, με στόχο την παραγωγή βαθμολογιών κινδύνου και την ιεράρχηση εγγραφών για περαιτέρω έλεγχο. Η ανάλυση με τον Νόμο του Benford χρησιμοποιείται συμπληρωματικά και δεν υποδηλώνει από μόνη της ύπαρξη απάτης. Συνολικά, το προτεινόμενο πλαίσιο μπορεί να υποστηρίξει την ελεγκτική διαδικασία ως μηχανισμός εντοπισμού και προτεραιοποίησης ύποπτων εγγραφών, χωρίς να υποκαθιστά την ανθρώπινη αξιολόγηση.

Financial Fraud and Accounting Anomaly Detection

using Hybrid Unsupervised Learning Models

Kyriakos Papatheodoropoulos

Abstract

This thesis examines the detection of financial fraud and accounting anomalies using unsupervised learning models and model combination techniques. Four different anomaly detectors are evaluated: Isolation Forest, Local Outlier Factor, Histogram-Based Outlier Score, and Autoencoder, with the aim of investigating both their individual performance and their complementarity. The experimental evaluation is initially conducted on the IEEE-CIS Fraud Detection dataset. The available labels are not used as a supervised learning target for training the main models, but are used for the quantitative evaluation of the results and the selection of the final experimental configurations. Specifically, in the final Autoencoder configuration, the labels are additionally used to construct a training set containing only normal observations, making this implementation closer to a one-class strategy. The examined model combination techniques include OR/Union, majority voting, and weighted score-based combination, while the contribution of the individual models is investigated through Shapley Values. The results show that Isolation Forest presents the strongest overall behavior at the selected operating point, whereas the Autoencoder achieves the highest AUPRC/AP. The model combination techniques do not consistently outperform the individual models, but instead provide different trade-offs between Precision and Recall. The IF-Dominant weighted combination is selected as the final configuration. The final framework is applied to 65,479 real ERP accounting entries without available fraud labels, with the aim of producing risk scores and prioritizing entries for further investigation. Benford's Law is used as a complementary analytical mechanism and does not, by itself, indicate the existence of fraud. Overall, the proposed framework can support the auditing process as a mechanism for identifying and prioritizing suspicious accounting entries, without replacing human evaluation.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον κ. Παναγιώτη Αδαμίδη για την καθοδήγηση και τη βοήθειά του σε όλη τη διάρκεια της διπλωματικής εργασίας, καθώς και για τις γνώσεις και την εμπειρία που απέκτησα μέσα από την ερευνητική διαδικασία.

Περιεχόμενα

Πρόλογος.....	v
Περίληψη.....	vi
Abstract	vii
Ευχαριστίες	viii
Περιεχόμενα	ix
Κατάλογος Σχημάτων	xii
Κατάλογος Πινάκων	xiii
Συντομογραφίες.....	xiv
Γλωσσάρι.....	xv
Κεφάλαιο 1ο: Εισαγωγή	1
1.1 Εισαγωγή	1
1.2 Πρόβλημα και Κίνητρο της Εργασίας.....	1
1.3 Σκοπός και Στόχοι της Διπλωματικής Εργασίας.....	2
1.4 Ερευνητικά Ερωτήματα	3
1.5 Δομή της Εργασίας.....	4
Κεφάλαιο 2ο: Θεωρητικό Υπόβαθρο και Σχετικές Εργασίες.....	6
2.1 Ανίχνευση Απάτης και Ανωμαλιών.....	6
2.2 Μηχανική Μάθηση για Ανίχνευση Απάτης.....	9
2.3 Μοντέλα Ανίχνευσης Ανωμαλιών και Τεχνικές Συνδυασμού	11
2.3.1 Isolation Forest	11
2.3.2 Local Outlier Factor.....	13
2.3.3 Histogram-Based Outlier Score	16
2.3.4 Autoencoder.....	18
2.3.5 Τεχνικές Συνδυασμού	21
2.3.6 Μέθοδοι Ανάλυσης Συμπληρωματικότητας και Συνεισφοράς.....	25
2.4 Νόμος του Benford.....	26
2.4.1 Θεωρητικό Υπόβαθρο	27
2.4.2 Κατανομή Πρώτου Σημαντικού Ψηφίου	27
2.4.3 Εφαρμογές του Benford στην Ανάλυση Απάτης.....	28
2.4.4 Ανάλυση Benford σε Οικονομικά Δεδομένα.....	29
2.5 Ανασκόπηση Σχετικών Εργασιών	30

Κεφάλαιο 3ο: Δεδομένα και Μεθοδολογία.....	34
3.1 Περιγραφή των Δεδομένων	34
3.1.1 Σύνολο Δεδομένων IEEE-CIS Fraud Detection	35
3.1.2 Σύνολο Δεδομένων ERP.....	36
3.2 Προεπεξεργασία Δεδομένων	37
3.2.1 Καθαρισμός Δεδομένων.....	39
3.2.2 Διαχείριση Ελλειπών Τιμών	39
3.2.3 Κωδικοποίηση Κατηγορικών Μεταβλητών	39
3.2.4 Κανονικοποίηση / Κλιμάκωση Χαρακτηριστικών	40
3.3 Πειραματική Μεθοδολογία.....	40
3.3.1 Πειραματικό Πρωτόκολλο.....	41
3.3.2 Εκπαίδευση και Αξιολόγηση των Μοντέλων	42
3.3.3 Διαμόρφωση των Μοντέλων	43
3.3.4 Ανάλυση Συμπληρωματικότητας και Συνεισφοράς των Μοντέλων	45
3.4 Μετρικές Αξιολόγησης	45
3.4.1 Precision	46
3.4.2 Recall	46
3.4.3 F1-score	46
3.4.4 ROC-AUC και AUPRC.....	47
3.4.5 Πίνακας Σύγκρισης	48
Κεφάλαιο 4ο: Αξιολόγηση και Συγκριτική Ανάλυση των Μοντέλων.....	49
4.1 Πειραματικό Περιβάλλον.....	49
4.2 Isolation Forest	49
4.3 Local Outlier Factor.....	51
4.4 Histogram-Based Outlier Score	52
4.5 Autoencoder.....	53
4.6 Συγκριτική Αξιολόγηση Μοντέλων.....	54
4.7 Ανάλυση Συμπληρωματικότητας των Μοντέλων.....	56
4.7.1 Ανάλυση Επικάλυψης.....	56
4.7.2 Συζήτηση της Συμπληρωματικότητας	59
4.8 Συμπεράσματα Κεφαλαίου.....	61
Κεφάλαιο 5ο: Αξιολόγηση των Τεχνικών Συνδυασμού Μοντέλων	62

5.1	Συνδυασμός OR/Union.....	62
5.2	Συνδυασμός Πλειοψηφικής Ψηφοφορίας.....	63
5.3	Σταθμισμένος Συνδυασμός Βαθμολογιών.....	64
5.4	Συγκριτική Αξιολόγηση των Τεχνικών Συνδυασμού.....	66
5.5	Ανάλυση Συνεισφοράς των Μοντέλων με Θεωρία Παιγνίων.....	67
5.6	Επιλογή της Τελικής Τεχνικής Συνδυασμού.....	69
Κεφάλαιο 6ο: Εφαρμογή σε Πραγματικά Δεδομένα ERP και Ανάλυση Benford		71
6.1	Εφαρμογή και Ανάλυση των Αποτελεσμάτων στα Δεδομένα ERP.....	71
6.2	Ανάλυση των Ύποπτων Συναλλαγών.....	73
6.3	Ανάλυση με τον Νόμο του Benford	75
6.4	Συνδυαστική Ερμηνεία του Μοντέλου και της Ανάλυσης Benford.....	77
6.5	Συζήτηση Αποτελεσμάτων.....	79
Κεφάλαιο 7ο: Συζήτηση και Συμπεράσματα.....		81
7.1	Συνολική Αποτίμηση	81
7.2	Απαντήσεις στα Ερευνητικά Ερωτήματα.....	83
7.3	Περιορισμοί της Εργασίας.....	84
7.4	Μελλοντική Έρευνα.....	85
7.5	Συμπεράσματα	86
ΒΙΒΛΙΟΓΡΑΦΙΑ		88
Παράρτημα Α: Αποτελέσματα Παραμετροποίησης μοντέλων		91

Κατάλογος Σχημάτων

Σχήμα 4.1. Επίδραση των υπερπαραμέτρων του Isolation Forest στην AUPRC και στη σχέση Recall–Precision	52
Σχήμα 4.2. Επίδραση του αριθμού των γειτόνων και του contamination στην απόδοση του LOF	54
Σχήμα 4.3. Επίδραση του αριθμού των bins και του contamination στην απόδοση του HBOS	55
Σχήμα 4.4. Επίδραση της διάστασης του λανθάνοντος χώρου, των εποχών εκπαίδευσης και του contamination στην απόδοση του Autoencoder.....	56
Σχήμα 4.5. Επικάλυψη των πραγματικών περιπτώσεων απάτης που εντοπίστηκαν από τα επιμέρους ζεύγη μοντέλων.....	60
Σχήμα 4.6. Συσχέτιση Pearson μεταξύ των βαθμολογιών ανωμαλίας των τεσσάρων μοντέλων	62
Σχήμα 6.1. Κατανομή των τελικών βαθμολογιών κινδύνου στις λογιστικές εγγραφές ERP και λειτουργικό κατώφλι του 90ού εκατοστημορίου	79
Σχήμα 6.2. Παρατηρούμενη και αναμενόμενη κατανομή πρώτου σημαντικού ψηφίου στο συνολικό σύνολο ERP	82
Σχήμα 6.3. Παρατηρούμενη κατανομή πρώτου σημαντικού ψηφίου στις εγγραφές που επισημάνθηκαν από το συνδυαστικό μοντέλο και θεωρητική κατανομή Benford	83
Σχήμα 6.4. Σύγκριση της κατανομής πρώτου σημαντικού ψηφίου μεταξύ επισημασμένων και μη επισημασμένων εγγραφών και της θεωρητικής κατανομής Benford	85

Κατάλογος Πινάκων

Πίνακας 2.1. Απεικόνιση κατανομής πρώτου ψηφίου.....	28
Πίνακας 3.1. Βασικά χαρακτηριστικά του IEEE-CIS	35
Πίνακας 3.2. Βασικά χαρακτηριστικά του συνόλου ERP	37
Πίνακας 3.3. Χαρακτηριστικά του τελικού συνόλου ERP	39
Πίνακας 3.4. Τελικές διαμορφώσεις των μοντέλων	45
Πίνακας 4.1. Συγκριτική αξιολόγηση των τελικών διαμορφώσεων των μοντέλων ανίχνευσης ανωμαλιών	57
Πίνακας 4.2. Επικάλυψη πραγματικών θετικών ανιχνεύσεων και δείκτης Jaccard ανά ζεύγος μοντέλων	61
Πίνακας 4.3. Συντελεστής συσχέτισης Pearson μεταξύ των βαθμολογιών ανωμαλίας.....	62
Πίνακας 4.4. Αποτελέσματα του ελέγχου McNemar μετά από διόρθωση Holm στις πραγματικές περιπτώσεις απάτης	63
Πίνακας 5.1. Αποτελέσματα του OR/Union Ensemble στο κοινό δείγμα των 100.000 συναλλαγών...67	
Πίνακας 5.2. Συγκριτικά αποτελέσματα των δύο διαμορφώσεων του Majority Voting Ensemble.....69	
Πίνακας 5.3. Σει διαμορφώσεων του Weighted Score-Based Ensemble	70
Πίνακας 5.4. Συγκριτικά αποτελέσματα των τεσσάρων διαμορφώσεων του Weighted Score-Based Ensemble	70
Πίνακας 5.5. Συγκριτική αξιολόγηση των εξεταζόμενων τεχνικών συνδυασμού.....71	
Πίνακας 5.6. Απόδοση των δυνατών συνασπισμών των τεσσάρων μοντέλων με βάση την AUPRC/AP	73
Πίνακας 5.7. Shapley Values των τεσσάρων μοντέλων ως προς την AUPRC/AP	74
Πίνακας 6.1. Βασικά χαρακτηριστικά της κατανομής της βαθμολογίας κινδύνου	78
Πίνακας 6.2. Μέση πραγματική συνεισφορά των επιμέρους μοντέλων στις επισημασμένες εγγραφές	79
Πίνακας 6.3. Κατανομή των επισημασμένων εγγραφών ανά τύπο	80
Πίνακας 6.4. Κατανομή πρώτου σημαντικού ψηφίου	81
Πίνακας 6.5. Αποτελέσματα ανάλυσης Benford	83
Πίνακας 6.6. Ποσοστό επισήμανσης ανά πρώτο σημαντικό ψηφίο	84

Συντομογραφίες

ML	Machine Learning
ERP	Enterprise Resource Planning
EDA	Exploratory Data Analysis
SMOTE	Synthetic Minority Over-sampling Technique
GL	General Ledger
IF	Isolation Forest
LOF	Local Outlier Factor
HBOS	Histogram Based Outlier Score
AE	Autoencoder
AEs	Autoencoders
EIF	Extended Isolation Forest
GIF	Generalized Isolation Forest
kNN	k-Nearest Neighbor
LRD	Local Reachability Density
MSE	Mean Squared Error
CAE	Convolutional Autoencoders
DAE	Denoising Autoencoders
VAE	Variational Autoencoders
MAD	Mean Absolute Deviation

Γλωσσάρι

Bottleneck Στενωμένος/ενδιάμεσος χώρος αναπαράστασης – Το τμήμα του δικτύου με περιορισμένη διάσταση που αναγκάζει το μοντέλο να διατηρήσει τις σημαντικότερες πληροφορίες της εισόδου.

Reconstruction Error Σφάλμα ανακατασκευής – Η απόκλιση μεταξύ της αρχικής παρατήρησης x και της ανακατασκευασμένης παρατήρησης $\hat{x}$.

Mean Squared Error Μέσο Τετραγωνικό Σφάλμα – Συνάρτηση που υπολογίζει τη μέση τετραγωνική διαφορά μεταξύ πραγματικών και ανακατασκευασμένων τιμών.

Threshold Κατώφλι – Προκαθορισμένη τιμή βάσει της οποίας αποφασίζεται αν μια παρατήρηση θα χαρακτηριστεί ως πιθανή ανωμαλία.

Hidden Layer Κρυφό επίπεδο – Ενδιάμεσο επίπεδο ενός νευρωνικού δικτύου μεταξύ εισόδου και εξόδου.

Activation Function Συνάρτηση ενεργοποίησης – Συνάρτηση που εφαρμόζεται στους νευρώνες και επιτρέπει στο δίκτυο να μοντελοποιεί μη γραμμικές σχέσεις.

Conformity Συμμόρφωση – Ο βαθμός στον οποίο η παρατηρούμενη κατανομή των ψηφίων συμφωνεί με τη θεωρητική κατανομή του Benford.

Non-Conformity Μη συμμόρφωση / απόκλιση – Διαφορά μεταξύ της παρατηρούμενης και της αναμενόμενης κατανομής του Benford.

Scale Invariance Ανεξαρτησία από την κλίμακα – Ιδιότητα σύμφωνα με την οποία η συμμόρφωση με τον Benford δεν μεταβάλλεται από αλλαγή μονάδας μέτρησης.

Adversarial manipulation Σκόπιμη χειραγώγηση με στόχο να ξεγελαστεί ή να παρακαμφθεί ένας μηχανισμός ελέγχου ή ανίχνευσης

Κεφάλαιο 1ο: Εισαγωγή

1.1 Εισαγωγή

Η συνεχής ψηφιοποίηση των οικονομικών δραστηριοτήτων έχει μεταβάλει ουσιαστικά τον τρόπο με τον οποίο πραγματοποιούνται, καταγράφονται και ελέγχονται οι χρηματοοικονομικές συναλλαγές. Τραπεζικά συστήματα, ηλεκτρονικές πληρωμές, πλατφόρμες ηλεκτρονικού εμπορίου και πληροφοριακά συστήματα επιχειρησιακών πόρων (ERP) παράγουν καθημερινά μεγάλους όγκους συναλλακτικών δεδομένων. Η αυξημένη διαθεσιμότητα τέτοιων δεδομένων δημιουργεί σημαντικές δυνατότητες αυτοματοποίησης των διαδικασιών οικονομικού ελέγχου, παράλληλα όμως αυξάνει την πολυπλοκότητα της έγκαιρης αναγνώρισης ύποπτων ή ασυνήθιστων συναλλαγών. Η οικονομική απάτη αποτελεί ένα πολυδιάστατο φαινόμενο, το οποίο μπορεί να περιλαμβάνει παραποίηση οικονομικών στοιχείων, μη εξουσιοδοτημένες συναλλαγές, εσωτερική κατάχρηση, απάτη μέσω καρτών, νομιμοποίηση εσόδων από παράνομες δραστηριότητες και άλλες μορφές σκόπιμης οικονομικής παραπλάνησης. Η βιβλιογραφία υποδηλώνει ότι η αυξανόμενη πολυπλοκότητα των χρηματοοικονομικών συστημάτων έχει οδηγήσει σε αντίστοιχη ανάπτυξη υπολογιστικών τεχνικών για την υποστήριξη της ανίχνευσης τέτοιων συμπεριφορών [1], [6], [10]. Ιδιαίτερη σημασία αποκτά και η έννοια της οικονομικής απάτης που διευκολύνεται μέσω ψηφιακών τεχνολογιών (cyber-enabled fraud), όπου οι ψηφιακές υποδομές δεν αποτελούν απαραίτητα τον τελικό στόχο της επίθεσης, αλλά το μέσο μέσω του οποίου πραγματοποιείται ή διευκολύνεται η οικονομική εξαπάτηση. Η μεγάλη ταχύτητα εκτέλεσης συναλλαγών, ο αυτοματισμός των επιχειρησιακών διαδικασιών και η δυνατότητα πραγματοποίησης μεγάλου αριθμού ενεργειών σε μικρό χρονικό διάστημα αυξάνουν την ανάγκη για συστήματα που μπορούν να εντοπίζουν έγκαιρα αποκλίσεις από τη συνηθισμένη συμπεριφορά. Οι συμβατικοί μηχανισμοί ελέγχου, όπως στατικοί κανόνες, προκαθορισμένα κατώφλια ή χειροκίνητη δειγματοληπτική εξέταση συναλλαγών, εξακολουθούν να αποτελούν σημαντικά εργαλεία. Ωστόσο, σε περιβάλλοντα μεγάλου όγκου και υψηλής διαστατικότητας δεδομένων ενδέχεται να μην επαρκούν για την αναγνώριση πολύπλοκων ή προηγουμένως άγνωστων προτύπων. Για τον λόγο αυτό, η αξιοποίηση τεχνικών μηχανικής μάθησης έχει αποκτήσει αυξανόμενο ενδιαφέρον στην ανάλυση χρηματοοικονομικών δεδομένων [6], [9], [10]. Μία ιδιαίτερα σχετική κατηγορία τεχνικών είναι η ανίχνευση ανωμαλιών (anomaly detection). Αντί να επιχειρείται αποκλειστικά η αναγνώριση ενός ήδη γνωστού προτύπου απάτης, η ανίχνευση ανωμαλιών επικεντρώνεται στον εντοπισμό παρατηρήσεων που αποκλίνουν από την κυρίαρχη δομή των δεδομένων. Στο πλαίσιο οικονομικών συναλλαγών, μία ανωμαλία μπορεί να αφορά ασυνήθιστο ποσό, μη συνηθισμένο συνδυασμό χαρακτηριστικών, τοπική απόκλιση από παρόμοιες συναλλαγές ή γενικότερα συμπεριφορά που διαφοροποιείται από το αναμενόμενο πρότυπο. Η ύπαρξη ανωμαλίας, ωστόσο, δεν συνεπάγεται κατ' ανάγκη την ύπαρξη απάτης. Μία συναλλαγή μπορεί να είναι ασυνήθιστη αλλά νόμιμη, ενώ μία πραγματική περίπτωση απάτης μπορεί να έχει σχεδιαστεί ώστε να προσομοιάζει σε κανονικές συναλλαγές. Επομένως, τα συστήματα ανίχνευσης ανωμαλιών είναι περισσότερο κατάλληλο να αντιμετωπίζονται ως μηχανισμοί εντοπισμού και ιεράρχησης ύποπτων περιπτώσεων για περαιτέρω έλεγχο, παρά ως μηχανισμοί οριστικής απόδειξης απάτης [6], [20], [31]. Η παρούσα εργασία εξετάζει αυτήν ακριβώς την προσέγγιση. Αντί να βασίζεται αποκλειστικά σε ένα μοντέλο, διερευνά τη συμπεριφορά διαφορετικών μη επιβλεπόμενων μηχανισμών ανίχνευσης ανωμαλιών και τη δυνατότητα συνδυασμού τους μέσω τεχνικών συνδυασμού. Η βασική υπόθεση που διερευνάται είναι ότι διαφορετικοί αλγόριθμοι ενδέχεται να αποτυπώνουν διαφορετικές μορφές απόκλισης και, συνεπώς, η αξιοποίηση της μεταξύ τους συμπληρωματικότητας μπορεί να προσφέρει πληρέστερη εικόνα του συναλλακτικού κινδύνου.

1.2 Πρόβλημα και Κίνητρο της Εργασίας

Ένα από τα σημαντικότερα χαρακτηριστικά των προβλημάτων ανίχνευσης οικονομικής απάτης είναι η έντονη ανισορροπία κλάσεων (class imbalance). Στα περισσότερα πραγματικά συναλλακτικά περιβάλλοντα, οι νόμιμες συναλλαγές αποτελούν τη συντριπτική πλειονότητα, ενώ οι επιβεβαιωμένες περιπτώσεις απάτης αντιπροσωπεύουν ένα πολύ μικρό ποσοστό του συνολικού όγκου δεδομένων. Η συγκεκριμένη ανισορροπία μπορεί να επηρεάσει ουσιαστικά τόσο την εκπαίδευση όσο και την αξιολόγηση μοντέλων μηχανικής μάθησης [12]–[14].

Σε ένα τέτοιο περιβάλλον, μία φαινομενικά υψηλή συνολική ακρίβεια δεν αποτελεί απαραίτητα ένδειξη αποτελεσματικής ανίχνευσης της μειοψηφικής κλάσης. Ένα μοντέλο μπορεί να ταξινομεί σωστά μεγάλο αριθμό νόμιμων συναλλαγών και ταυτόχρονα να αποτυγχάνει να εντοπίσει σημαντικό μέρος των πραγματικών περιπτώσεων απάτης. Για τον λόγο αυτό, η αξιολόγηση απαιτεί μετρικές που εστιάζουν περισσότερο στη θετική κλάση, όπως Precision, Recall, F1-score και μετρικές που βασίζονται στη σχέση Precision–Recall [14], [52]–[54].

Παράλληλα, τα πραγματικά δεδομένα οικονομικής απάτης χαρακτηρίζονται συχνά από περιορισμένη ή καθυστερημένη διαθεσιμότητα αξιόπιστων ετικετών. Η επιβεβαίωση μίας ύποπτης συναλλαγής μπορεί να απαιτεί εσωτερικό έλεγχο, διερεύνηση από εξειδικευμένο προσωπικό ή ακόμη και νομική διαδικασία. Επιπλέον, τα πρότυπα απάτης δεν παραμένουν σταθερά στον χρόνο. Μεταβάλλονται καθώς μεταβάλλονται οι επιχειρησιακές διαδικασίες, οι τεχνολογίες και οι στρατηγικές των δραστών, γεγονός που μπορεί να περιορίσει την ικανότητα ενός μοντέλου που εκπαιδεύτηκε αποκλειστικά σε ιστορικές επισημασμένες περιπτώσεις να γενικεύει σε νέες μορφές συμπεριφοράς [12], [16], [17].

Οι επιβλεπόμενες προσεγγίσεις μπορούν να είναι ιδιαίτερα αποτελεσματικές όταν είναι διαθέσιμο επαρκές και αντιπροσωπευτικό επισημασμένο σύνολο δεδομένων. Ωστόσο, η εξάρτησή τους από γνωστές ετικέτες αποτελεί περιορισμό όταν το ζητούμενο είναι η αναγνώριση νέων ή σπάνιων μορφών απάτης. Επιπλέον, η μεγάλη αριθμητική υπεροχή της κανονικής κλάσης μπορεί να απαιτήσει ειδικούς μηχανισμούς αντιμετώπισης της ανισορροπίας [13], [14].

Η μη επιβλεπόμενη ανίχνευση ανωμαλιών αποτελεί μία εναλλακτική προσέγγιση, καθώς επιτρέπει την αναζήτηση αποκλινουσών παρατηρήσεων χωρίς να απαιτείται η χρήση της πραγματικής κλάσης ως στόχου κατά την εκπαίδευση. Τεχνικές όπως τα Isolation Forest, Local Outlier Factor, μεθόδους βασισμένες σε ιστογράμματα και Autoencoders βασίζονται σε διαφορετικές αρχές εντοπισμού ανωμαλιών και, επομένως, ενδέχεται να αποτυπώνουν διαφορετικές πτυχές της ίδιας συναλλακτικής δομής [20], [24], [31], [33].

Η διαφορετική αυτή συμπεριφορά δημιουργεί παράλληλα το κίνητρο για τη διερεύνηση τεχνικών συνδυασμού. Ένας μεμονωμένος ανιχνευτής μπορεί να εμφανίζει ισχυρή απόδοση σε συγκεκριμένο τύπο αποκλίσεων αλλά να μην ανταποκρίνεται εξίσου αποτελεσματικά σε διαφορετικές μορφές ανωμαλίας. Αντίθετα, η συνδυαστική αξιοποίηση πολλών ανιχνευτών μπορεί να επιτρέψει τη συνεκτίμηση διαφορετικών ενδείξεων [38]–[40]. Η χρησιμότητα ενός συνδυαστικού μοντέλου, ωστόσο, δεν πρέπει να θεωρείται δεδομένη. Εξαρτάται από τον βαθμό στον οποίο τα επιμέρους μοντέλα παρέχουν συμπληρωματική και όχι πλήρως πλεονάζουσα πληροφορία.

Το κίνητρο της παρούσας εργασίας προκύπτει επομένως από τρεις αλληλένδετες ανάγκες: την αντιμετώπιση της περιορισμένης διαθεσιμότητας ετικετών, την αποτελεσματική ανάλυση ενός έντονα ανισόρροπου προβλήματος και τη διερεύνηση του κατά πόσο διαφορετικοί μη επιβλεπόμενοι μηχανισμοί μπορούν να συνδυαστούν ώστε να υποστηρίξουν αποτελεσματικότερη ιεράρχηση ύποπτων συναλλαγών.

Η διερεύνηση δεν περιορίζεται σε ένα ελεγχόμενο σύνολο δεδομένων αναφοράς. Ιδιαίτερο ενδιαφέρον παρουσιάζει η εφαρμογή της ίδιας μεθοδολογικής λογικής σε πραγματικά λογιστικά δεδομένα ERP, όπου δεν υπάρχουν διαθέσιμες πραγματικές ετικέτες απάτης. Στην περίπτωση αυτή, η χρησιμότητα του συστήματος σχετίζεται περισσότερο με την παραγωγή βαθμολογιών κινδύνου και την ανάδειξη συναλλαγών που χρήζουν περαιτέρω ελέγχου παρά με τον άμεσο υπολογισμό κλασικών μετρικών αξιολόγησης που απαιτούν γνωστές ετικέτες.

1.3 Σκοπός και Στόχοι της Διπλωματικής Εργασίας

Κεντρικός σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη και πειραματική αξιολόγηση ενός πλαισίου ανίχνευσης οικονομικών και λογιστικών ανωμαλιών, στο οποίο διαφορετικοί μηχανισμοί ανίχνευσης εξετάζονται τόσο μεμονωμένα όσο και σε συνδυασμό. Η εργασία επιδιώκει να διερευνήσει κατά πόσο η διαφορετική συμπεριφορά των επιμέρους μοντέλων μπορεί να αξιοποιηθεί μέσω τεχνικών συνδυασμού και, στη συνέχεια, να μεταφερθεί σε πραγματικά συναλλακτικά δεδομένα ERP.

Οι βασικοί στόχοι της εργασίας είναι:

- η ανάπτυξη ενός ενιαίου πλαισίου προεπεξεργασίας, εκπαίδευσης και αξιολόγησης μοντέλων ανίχνευσης ανωμαλιών,

- η συγκριτική αξιολόγηση των Isolation Forest, Local Outlier Factor, Histogram-Based Outlier Score και Autoencoder,
- η διερεύνηση της επικάλυψης και της συμπληρωματικότητας των ανιχνεύσεων και των βαθμολογιών ανωμαλίας των επιμέρους μοντέλων,
- η ανάπτυξη και αξιολόγηση διαφορετικών στρατηγικών συνδυασμού, όπως OR/Union, πλειοψηφική ψηφοφορία και σταθμισμένος συνδυασμός βαθμολογιών,
- η διερεύνηση της συνεισφοράς των επιμέρους μοντέλων στον συνδυαστικό μηχανισμό,
- η επιλογή της καταλληλότερης στρατηγικής συνδυασμού με βάση τη συνολική συνεκτίμηση των μετρικών αξιολόγησης και του λειτουργικού προφίλ των εξεταζόμενων διαμορφώσεων,
- η εφαρμογή του επιλεγμένου πλαισίου σε πραγματικά λογιστικά δεδομένα ERP χωρίς διαθέσιμες ετικέτες απάτης,
- και η συμπληρωματική στατιστική διερεύνηση των αποτελεσμάτων μέσω του Νόμου του Benford, ο οποίος έχει χρησιμοποιηθεί στη βιβλιογραφία ως εργαλείο εξέτασης ασυνήθιστων αριθμητικών προτύπων σε οικονομικά και λογιστικά δεδομένα [44]–[49].

Η χρήση του Νόμου του Benford δεν αποσκοπεί στην ανεξάρτητη επιβεβαίωση ότι μία συναλλαγή αποτελεί απάτη. Αντίθετα, χρησιμοποιείται ως συμπληρωματικός στατιστικός μηχανισμός για την εξέταση της αριθμητικής συμπεριφοράς των συναλλαγών που επισημαίνονται από το προτεινόμενο πλαίσιο. Με τον τρόπο αυτό συνδυάζονται τεχνικές μηχανικής μάθησης και στατιστικής ανάλυσης, χωρίς η απόκλιση από μία αναμενόμενη κατανομή να εξισώνεται με αποδεδειγμένη δόλια δραστηριότητα.

Η πειραματική διαδικασία πραγματοποιείται αρχικά στο σύνολο δεδομένων IEEE-CIS Fraud Detection. Οι διαθέσιμες ετικέτες δεν χρησιμοποιούνται ως στόχος επιβλεπόμενης μάθησης για την εκπαίδευση των βασικών ανιχνευτών, αλλά αξιοποιούνται για την ποσοτική αξιολόγηση των αποτελεσμάτων και την επιλογή των τελικών πειραματικών διαμορφώσεων. Ειδικά στην τελική διαμόρφωση του Autoencoder, χρησιμοποιούνται επιπλέον για τον σχηματισμό συνόλου εκπαίδευσης που περιλαμβάνει μόνο κανονικές παρατηρήσεις, με αποτέλεσμα η συγκεκριμένη εφαρμογή να προσεγγίζει στρατηγική μίας κλάσης. Στη συνέχεια, η επιλεγμένη μεθοδολογία εφαρμόζεται σε πραγματικό σύνολο λογιστικών συναλλαγών ERP χωρίς γνωστές ετικέτες απάτης. Με αυτόν τον σχεδιασμό, το IEEE-CIS λειτουργεί ως ελεγχόμενο περιβάλλον αξιολόγησης, ενώ το σύνολο δεδομένων ERP χρησιμοποιείται για τη διερεύνηση της τεχνικής εφαρμοσιμότητας και της δυνατότητας ιεράρχησης ύποπτων λογιστικών εγγραφών.

1.4 Ερευνητικά Ερωτήματα

Με βάση τον παραπάνω σκοπό, η εργασία οργανώνεται γύρω από έξι βασικά ερευνητικά ερωτήματα.

RQ1 – Ποιο από τα εξεταζόμενα μοντέλα ανίχνευσης ανωμαλιών παρουσιάζει την ισχυρότερη συνολική συμπεριφορά για την ανίχνευση οικονομικής απάτης στο σύνολο δεδομένων αναφοράς;

Το ερώτημα εξετάζεται μέσω της συγκριτικής αξιολόγησης των Isolation Forest, LOF, HBOS και Autoencoder με βάση τις Precision, Recall, F1-score, ROC-AUC και AUPRC/AP.

RQ2 – Σε ποιο βαθμό τα διαφορετικά μοντέλα παρουσιάζουν συμπληρωματικότητα ως προς τις περιπτώσεις απάτης που εντοπίζουν και τις βαθμολογίες ανωμαλίας που παράγουν;

Η διερεύνηση δεν περιορίζεται στις συνολικές μετρικές απόδοσης, αλλά εξετάζει τόσο την επικάλυψη των πραγματικά θετικών ανιχνεύσεων όσο και τη συσχέτιση των συνεχών βαθμολογιών ανωμαλίας.

RQ3 – Μπορούν οι στρατηγικές συνδυασμού να βελτιώσουν επιμέρους χαρακτηριστικά της συνολικής συμπεριφοράς του συστήματος σε σχέση με τη χρήση μεμονωμένων ανιχνευτών;

Εξετάζονται διαφορετικοί μηχανισμοί συνδυασμού, ώστε να αποτιμηθεί κατά πόσο η αξιοποίηση πολλαπλών μοντέλων μπορεί να διαμορφώσει διαφορετικό ή περισσότερο ισορροπημένο προφίλ ανίχνευσης, χωρίς να προϋποτίθεται ότι κάποια τεχνική συνδυασμού υπερέχει καθολικά.

RQ4 – Πώς επηρεάζεται ο συμβιβασμός μεταξύ Recall και Precision από τον τρόπο συνδυασμού των μοντέλων και από το επιλεγμένο σημείο λειτουργίας;

Ιδιαίτερη έμφαση δίνεται στη σχέση μεταξύ της κάλυψης πραγματικών περιπτώσεων απάτης και του αντίστοιχου αριθμού ψευδώς θετικών επισημάνσεων (False Positives – FP).

RQ5 – Μπορεί ο Νόμος του Benford να λειτουργήσει συμπληρωματικά προς την ανίχνευση μέσω συνδυασμού μοντέλων ως μηχανισμός στατιστικής διερεύνησης των ύποπτων οικονομικών συναλλαγών;

Το ερώτημα εξετάζει κατά πόσο οι συναλλαγές που λαμβάνουν υψηλότερη βαθμολογία κινδύνου παρουσιάζουν διαφοροποιημένη συμπεριφορά ως προς την κατανομή του πρώτου σημαντικού ψηφίου, χωρίς η απόκλιση αυτή να ερμηνεύεται αυτομάτως ως απόδειξη απάτης.

RQ6 – Πώς συμπεριφέρεται το προτεινόμενο πλαίσιο όταν εφαρμόζεται σε πραγματικά δεδομένα ERP χωρίς διαθέσιμες ετικέτες απάτης;

Η εφαρμογή στο σύνολο δεδομένων ERP επικεντρώνεται στην παραγωγή βαθμολογιών κινδύνου, στην ιεράρχηση ύποπτων λογιστικών εγγραφών και στη διερεύνηση της τεχνικής εφαρμοσιμότητας της μεθοδολογίας σε πραγματικό λογιστικό περιβάλλον.

Τα ερευνητικά ερωτήματα ακολουθούν διαδοχική λογική. Αρχικά αξιολογούνται οι μεμονωμένοι ανιχνευτές και η μεταξύ τους συμπληρωματικότητα. Στη συνέχεια εξετάζεται η συμπεριφορά διαφορετικών στρατηγικών συνδυασμού και επιλέγεται η τελική διαμόρφωση για την εφαρμογή σε πραγματικά δεδομένα ERP. Τέλος, ο Νόμος του Benford χρησιμοποιείται ως συμπληρωματικός μηχανισμός στατιστικής διερεύνησης των αποτελεσμάτων.

1.5 Δομή της Εργασίας

Η παρούσα διπλωματική εργασία οργανώνεται σε επτά βασικά κεφάλαια, τα οποία ακολουθούν τη μετάβαση από το θεωρητικό υπόβαθρο και τον πειραματικό σχεδιασμό προς τη συγκριτική αξιολόγηση και την εφαρμογή σε πραγματικά δεδομένα.

Το Κεφάλαιο 1 παρουσιάζει το γενικό πλαίσιο της οικονομικής απάτης και της ανίχνευσης ανωμαλιών, το πρόβλημα που εξετάζεται, το κίνητρο της εργασίας, τους επιμέρους στόχους και τα βασικά ερευνητικά ερωτήματα.

Το Κεφάλαιο 2 αναπτύσσει το θεωρητικό υπόβαθρο της εργασίας. Παρουσιάζονται οι έννοιες της οικονομικής απάτης και της ανίχνευσης ανωμαλιών, η αξιοποίηση της μηχανικής μάθησης σε προβλήματα ανίχνευσης απάτης, οι αρχές λειτουργίας των Isolation Forest, LOF, HBOS και Autoencoder, οι βασικές τεχνικές συνδυασμού μοντέλων και ο Νόμος του Benford. Παράλληλα, παρουσιάζονται σχετικές ερευνητικές εργασίες και προσεγγίσεις της διεθνούς βιβλιογραφίας.

Το Κεφάλαιο 3 περιγράφει τη μεθοδολογία της πειραματικής διαδικασίας. Παρουσιάζονται τα σύνολα δεδομένων IEEE-CIS και ERP, τα στάδια προεπεξεργασίας, το πειραματικό πρωτόκολλο, η διαδικασία παραμετροποίησης των μοντέλων, οι υβριδικοί συνδυασμοί και οι μετρικές που χρησιμοποιούνται για την αξιολόγηση.

Στο Κεφάλαιο 4 πραγματοποιείται η αξιολόγηση των επιμέρους μοντέλων ανίχνευσης ανωμαλιών. Παρουσιάζονται τα αποτελέσματα παραμετροποίησης και η τελική επίδοση των Isolation Forest, LOF, HBOS και Autoencoder, ακολουθεί η συγκριτική αξιολόγησή τους και διερευνάται η συμπληρωματικότητά τους μέσω της ανάλυσης των κοινών και διαφορετικών ανιχνεύσεων και των αντίστοιχων βαθμολογιών ανωμαλίας.

Το Κεφάλαιο 5 επικεντρώνεται στην αξιολόγηση των τεχνικών συνδυασμού. Εξετάζονται οι τεχνικές OR/Union, πλειοψηφικής ψηφοφορίας και σταθμισμένου συνδυασμού βαθμολογιών, συγκρίνονται οι διαφορετικοί συμβιβασμοί μεταξύ Recall και Precision και διερευνάται η συνεισφορά των επιμέρους ανιχνευτών στον συνδυαστικό μηχανισμό. Στο τέλος του κεφαλαίου επιλέγεται η τεχνική συνδυασμού που χρησιμοποιείται στην εφαρμογή σε πραγματικά δεδομένα.

Στο Κεφάλαιο 6 η επιλεγμένη στρατηγική συνδυασμού εφαρμόζεται σε πραγματικές συναλλαγές ERP. Αναλύονται οι συναλλαγές που λαμβάνουν υψηλότερες βαθμολογίες κινδύνου και πραγματοποιείται συμπληρωματική ανάλυση μέσω του Νόμου του Benford. Με τον τρόπο αυτό εξετάζεται η μεταφορά της προτεινόμενης μεθοδολογίας από ένα ελεγχόμενο σύνολο δεδομένων αναφοράς με διαθέσιμες ετικέτες σε ένα πραγματικό λογιστικό περιβάλλον χωρίς επιβεβαιωμένες ετικέτες απάτης.

Τέλος, το Κεφάλαιο 7 συνθέτει και συζητά τα συνολικά ευρήματα της εργασίας. Παρουσιάζεται η συνολική αποτίμηση της πειραματικής διαδικασίας, δίνονται απαντήσεις στα ερευνητικά ερωτήματα που διατυπώθηκαν στο Κεφάλαιο 1, εξετάζονται οι βασικοί περιορισμοί της μεθοδολογίας και προτείνονται κατευθύνσεις για μελλοντική έρευνα. Το κεφάλαιο ολοκληρώνεται με τα κύρια συμπεράσματα σχετικά με τη δυνατότητα αξιοποίησης μη επιβλεπόμενων τεχνικών συνδυασμού ως μηχανισμοί ιεράρχησης ύποπτων οικονομικών και λογιστικών εγγραφών σε περιβάλλοντα ERP.

Η παραπάνω διάρθρωση επιτρέπει τη σταδιακή μετάβαση από τη θεωρητική θεμελίωση και τη μεθοδολογική ανάπτυξη στην πειραματική αξιολόγηση, στην εφαρμογή σε πραγματικά δεδομένα και, τελικά, στη συνολική ερμηνεία και αποτίμηση των αποτελεσμάτων.

Κεφάλαιο 2ο: Θεωρητικό Υπόβαθρο και Σχετικές Εργασίες

Το παρόν κεφάλαιο διαμορφώνει το θεωρητικό και ερευνητικό πλαίσιο της εργασίας. Αρχικά αποσαφηνίζονται οι έννοιες της οικονομικής απάτης, της λογιστικής ανωμαλίας και της ανίχνευσης ανωμαλιών. Στη συνέχεια παρουσιάζονται οι βασικές κατηγορίες μηχανικής μάθησης που χρησιμοποιούνται στην ανίχνευση απάτης και αιτιολογείται η επιλογή μη επιβλεπόμενων μεθόδων για περιβάλλοντα όπου οι αξιόπιστες ετικέτες είναι περιορισμένες ή καθυστερούν. Ακολουθεί η τεχνική ανάλυση των Isolation Forest, Local Outlier Factor, Histogram-Based Outlier Score και Autoencoder, καθώς και των σχημάτων συνδυασμού OR/Union, πλειοψηφικής ψηφοφορίας και σύντηξης βαθμολογιών. Τέλος, αναπτύσσεται ο νόμος του Benford ως συμπληρωματικό εργαλείο ψηφιακής ανάλυσης και συντίθεται η σχετική βιβλιογραφία, ώστε να προσδιοριστεί το ερευνητικό κενό στο οποίο τοποθετείται η προτεινόμενη μεθοδολογία.

2.1 Ανίχνευση Απάτης και Ανωμαλιών

Η οικονομική απάτη (financial fraud) αποτελεί ένα σύνθετο πρόβλημα που μπορεί να εμφανιστεί σε διαφορετικούς τομείς της οικονομικής δραστηριότητας, όπως στις τραπεζικές συναλλαγές, στις ασφαλιστικές απαιτήσεις, στις ηλεκτρονικές πληρωμές, στη φορολογία, στη νομιμοποίηση εσόδων από παράνομες δραστηριότητες και στη χρηματοοικονομική πληροφόρηση [1], [6], [10]. Βασικό χαρακτηριστικό της είναι η ύπαρξη σκόπιμης παραπλανητικής ή παράνομης ενέργειας που αποσκοπεί στην απόκτηση οικονομικού οφέλους ή στη διαστρέβλωση της πραγματικής οικονομικής κατάστασης ενός οργανισμού ή μιας συναλλαγής [1], [10]. Η ποικιλία των μορφών με τις οποίες μπορεί να εκδηλωθεί καθιστά την ανίχνευσή της ιδιαίτερα απαιτητική, καθώς διαφορετικά είδη απάτης παρουσιάζουν διαφορετικά χαρακτηριστικά και δεν μπορούν απαραίτητα να αντιμετωπιστούν με έναν ενιαίο μηχανισμό ανίχνευσης [6].

Μια βασική διάκριση αφορά την προέλευση της απάτης. Η εσωτερική απάτη προέρχεται από εργαζομένους, στελέχη ή άλλα άτομα που διαθέτουν νόμιμη πρόσβαση στους πόρους και στα πληροφοριακά συστήματα ενός οργανισμού, ενώ η εξωτερική απάτη πραγματοποιείται από άτομα που βρίσκονται εκτός του οργανισμού. Υπάρχουν επίσης περιπτώσεις στις οποίες εσωτερικοί και εξωτερικοί δράστες συνεργάζονται για την πραγματοποίηση της δόλιας ενέργειας [4]. Στις εξωτερικές μορφές συναντώνται συχνά απάτες που αφορούν πιστωτικές κάρτες, ηλεκτρονικές πληρωμές, ασφαλιστικές απαιτήσεις και δάνεια, ενώ στις εσωτερικές περιλαμβάνονται μεταξύ άλλων η υπεξαίρεση περιουσιακών στοιχείων, η διαφθορά, η παραποίηση οικονομικών καταστάσεων και άλλες μη εξουσιοδοτημένες οικονομικές ενέργειες [4], [10]. Η διαφοροποίηση αυτή είναι σημαντική, καθώς ο τρόπος εκδήλωσης μιας απάτης, τα διαθέσιμα δεδομένα και οι μηχανισμοί εντοπισμού της μπορούν να διαφέρουν σημαντικά ανά περίπτωση.

Ιδιαίτερη κατηγορία αποτελεί η απάτη στις οικονομικές καταστάσεις, η οποία συνδέεται με την εσκεμμένη παραποίηση ή παράλειψη οικονομικών πληροφοριών. Τέτοιες ενέργειες μπορεί να περιλαμβάνουν την υπερτίμηση εσόδων, κερδών ή περιουσιακών στοιχείων, την υποτίμηση εξόδων και υποχρεώσεων, την εσφαλμένη αποτίμηση στοιχείων του ενεργητικού ή την παροχή παραπλανητικών γνωστοποιήσεων [1], [5]. Μέσω αυτών των ενεργειών μπορεί να επιχειρείται η παρουσίαση μιας ευνοϊκότερης οικονομικής εικόνας, η επίτευξη συγκεκριμένων οικονομικών στόχων, η απόκρυψη αρνητικών πληροφοριών ή η άσκηση επιρροής στις αποφάσεις επενδυτών και πιστωτών [1], [5].

Η ύπαρξη όμως μιας ανακριβούς λογιστικής εγγραφής ή οικονομικής πληροφορίας δεν συνεπάγεται απαραίτητα την ύπαρξη απάτης. Οι ανακρίβειες στις λογιστικές πληροφορίες αποτελούν ευρύτερη κατηγορία, καθώς μπορεί να εμφανιστούν χωρίς να υπάρχει δόλια πρόθεση ή αποδεδειγμένη προσπάθεια χειραγώγησης [5], [8]. Ένα λογιστικό σφάλμα, μια εσφαλμένη εκτίμηση ή μια λανθασμένη καταχώριση μπορούν επομένως να δημιουργήσουν αποκλίσεις στα οικονομικά δεδομένα χωρίς η αιτία τους να είναι απαραίτητα δόλια δραστηριότητα. Η διάκριση μεταξύ σφάλματος, λογιστικής ανακρίβειας και σκόπιμης παραποίησης είναι ιδιαίτερα σημαντική, καθώς επηρεάζει άμεσα τον τρόπο με τον οποίο πρέπει να ερμηνεύονται τα αποτελέσματα ενός συστήματος αυτόματης ανίχνευσης [5], [8].

Σε αυτό το σημείο αποκτά ιδιαίτερη σημασία η έννοια της ανίχνευσης ανωμαλιών. Η ανίχνευση ανωμαλιών αφορά τον εντοπισμό παρατηρήσεων, συναλλαγών ή προτύπων τα οποία αποκλίνουν από την αναμενόμενη ή συνηθισμένη συμπεριφορά του συνόλου δεδομένων [6]. Μια ανωμαλία μπορεί να αποτελεί σημαντική ένδειξη ότι μια παρατήρηση έχει δημιουργηθεί από διαφορετικό μηχανισμό σε σχέση με το μεγαλύτερο μέρος των δεδομένων, χωρίς όμως αυτό να προσδιορίζει αυτόματα την πραγματική αιτία της απόκλισης [6]. Στο οικονομικό περιβάλλον, μια ασυνήθιστη συναλλαγή μπορεί να σχετίζεται με απάτη, αλλά μπορεί επίσης να προκύπτει από ανθρώπινο λάθος, πρόβλημα ποιότητας δεδομένων, τεχνική δυσλειτουργία ή μια νόμιμη αλλά σπάνια επιχειρησιακή δραστηριότητα [2], [6]. Για τον λόγο αυτό, οι έννοιες ανίχνευση απάτης και ανίχνευση ανωμαλιών είναι στενά συνδεδεμένες αλλά δεν είναι ταυτόσημες. Η ανίχνευση απάτης στοχεύει τελικά στην αναγνώριση πιθανών δόλιων ενεργειών, ενώ η ανίχνευση ανωμαλιών περιορίζεται στον εντοπισμό αποκλίσεων από ένα πρότυπο φυσιολογικής συμπεριφοράς. Μια ανωμαλία επομένως δεν αποτελεί από μόνη της απόδειξη απάτης, αλλά μπορεί να λειτουργήσει ως ένδειξη αυξημένου κινδύνου και ως αφετηρία για περαιτέρω διερεύνηση. Η προσέγγιση αυτή είναι ιδιαίτερα σημαντική στο πλαίσιο του οικονομικού ελέγχου, όπου μια ύποπτη συναλλαγή ή λογιστική εγγραφή πρέπει να εξεταστεί περαιτέρω πριν χαρακτηριστεί ως πραγματικό περιστατικό απάτης.

Η σύνδεση αυτή γίνεται ακόμη πιο εμφανής στις εγγραφές του γενικού καθολικού (GL). Τα δεδομένα του γενικού καθολικού αποτυπώνουν την οικονομική δραστηριότητα ενός οργανισμού και μπορεί να περιλαμβάνουν πληροφορίες όπως ο λογαριασμός καταχώρισης, ο τύπος της συναλλαγής, η ένδειξη χρέωσης ή πίστωσης, το ποσό, η ημερομηνία καταχώρισης και το πληροφοριακό σύστημα από το οποίο προέρχεται η συναλλαγή [2]. Σε πραγματικά λογιστικά και ERP περιβάλλοντα, οι εγγραφές αυτές μπορεί να παρουσιάζουν σημαντική ετερογένεια ως προς τη δομή τους, τον αριθμό των συναλλαγών και τον τρόπο με τον οποίο χρησιμοποιούνται οι λογαριασμοί [2].

Οι λογιστικές ανωμαλίες δεν περιορίζονται απαραίτητα σε εξαιρετικά υψηλά ποσά ή σε εμφανείς παραβιάσεις λογιστικών κανόνων. Μπορούν να εμφανιστούν μέσα από ασυνήθιστους συνδυασμούς λογαριασμών, μη αναμενόμενες σχέσεις μεταξύ συναλλαγών ή διαφορετικά πρότυπα χρήσης λογαριασμών και χαρακτηριστικών σε σχέση με την υπόλοιπη οικονομική δραστηριότητα [2]. Επομένως, η αναγνώριση λογιστικών ανωμαλιών μπορεί να λειτουργήσει ως ένα πρώτο φίλτρο ελέγχου για τον εντοπισμό εγγραφών που παρουσιάζουν αυξημένο ενδιαφέρον και απαιτούν αναλυτικότερη εξέταση.

Οι ανωμαλίες μπορούν να διακριθούν σε διαφορετικές κατηγορίες ανάλογα με τον τρόπο με τον οποίο εμφανίζονται στα δεδομένα. Οι σημειακές ανωμαλίες αφορούν μεμονωμένες παρατηρήσεις που διαφέρουν σημαντικά από το μεγαλύτερο μέρος του συνόλου δεδομένων. Για παράδειγμα, μια συναλλαγή ιδιαίτερα υψηλής αξίας σε σχέση με τη συνήθη συμπεριφορά ενός λογαριασμού μπορεί να αποτελέσει μια τέτοια περίπτωση. Οι ανωμαλίες βάσει πλαισίου εμφανίζονται όταν μια παρατήρηση θεωρείται ασυνήθιστη μόνο μέσα σε ένα συγκεκριμένο πλαίσιο. Μια συναλλαγή μπορεί για παράδειγμα να είναι φυσιολογική ως προς το ποσό της, αλλά να θεωρείται ασυνήθιστη λόγω της χρονικής στιγμής ή του επιχειρησιακού πλαισίου στο οποίο πραγματοποιήθηκε. Τέλος, οι συλλογικές ανωμαλίες αφορούν σύνολα παρατηρήσεων των οποίων η συνδυασμένη συμπεριφορά είναι ασυνήθιστη, ακόμη και όταν οι επιμέρους παρατηρήσεις δεν παρουσιάζουν κάποιο εμφανές πρόβλημα [6].

Η ύπαρξη αυτών των διαφορετικών μορφών δείχνει ότι η ανίχνευση οικονομικής απάτης δεν μπορεί να βασίζεται αποκλειστικά στον εντοπισμό ακραίων αριθμητικών τιμών. Σε αρκετές περιπτώσεις, η σημαντική πληροφορία μπορεί να βρίσκεται στις σχέσεις μεταξύ διαφορετικών χαρακτηριστικών ή στον συνδυασμό πολλών συναλλαγών. Η ταυτόχρονη αξιοποίηση λογιστικών, ελεγκτικών, χρηματιστηριακών και άλλων μεταβλητών μπορεί να αποκαλύψει σχέσεις που δεν είναι εμφανείς όταν κάθε χαρακτηριστικό εξετάζεται ανεξάρτητα [8]. Ορισμένες λογιστικές μεταβλητές μπορεί επομένως να έχουν περιορισμένη διακριτική ικανότητα όταν χρησιμοποιούνται μεμονωμένα, αλλά να αποκτούν μεγαλύτερη αξία όταν εξετάζονται σε συνδυασμό με πληροφορίες που σχετίζονται με τον έλεγχο, την αγορά ή άλλα χαρακτηριστικά της επιχείρησης [8].

Η διερευνητική ανάλυση δεδομένων (EDA) αποτελεί επίσης σημαντικό στάδιο για την κατανόηση αυτών των προτύπων. Στόχος της είναι η διερεύνηση της δομής των δεδομένων, η αναγνώριση σχέσεων μεταξύ μεταβλητών και ο εντοπισμός πιθανών ακραίων παρατηρήσεων και ανωμαλιών χωρίς να απαιτείται απαραίτητα μια πλήρως προκαθορισμένη υπόθεση [3]. Η οπτικοποίηση των κατανομών, η ανάλυση πολλών μεταβλητών και η αναζήτηση συσχετίσεων μπορούν να προσφέρουν αρχική εικόνα για τον τρόπο με τον οποίο οργανώνονται τα οικονομικά δεδομένα και για τις περιπτώσεις που

αποκλίνουν από τη συνηθισμένη συμπεριφορά [3]. Η λογική αυτή συνδέεται άμεσα με τις τεχνικές μη επιβλεπόμενης μάθησης, οι οποίες επιχειρούν να αναγνωρίσουν δομές και μοτίβα χωρίς να βασίζονται αποκλειστικά σε προκαθορισμένες ετικέτες.

Η ανίχνευση ανωμαλιών παρουσιάζει ωστόσο σημαντικές δυσκολίες. Ένα από τα βασικότερα προβλήματα είναι ο καθορισμός της φυσιολογικής συμπεριφοράς. Το όριο μεταξύ κανονικών και ασυνήθιστων παρατηρήσεων δεν είναι πάντοτε σαφές, με αποτέλεσμα ορισμένες φυσιολογικές παρατηρήσεις να βρίσκονται κοντά στις ανώμαλες και αντίστροφα [6]. Παράλληλα, τα πρότυπα της οικονομικής δραστηριότητας μεταβάλλονται με την πάροδο του χρόνου και η συμπεριφορά που θεωρείται φυσιολογική σε μια χρονική περίοδο δεν είναι απαραίτητο να παραμένει ίδια στο μέλλον [6]. Το ίδιο πρόβλημα εμφανίζεται και στις μορφές απάτης. Οι τεχνικές που χρησιμοποιούνται από τους δράστες μπορούν να εξελίσσονται ώστε να αποφεύγουν τους υπάρχοντες μηχανισμούς ελέγχου [6], [9]. Ένα μοντέλο που βασίζεται αποκλειστικά σε ιστορικά δεδομένα μπορεί επομένως να παρουσιάσει μειωμένη αποτελεσματικότητα όταν εμφανιστούν νέα πρότυπα απάτης ή όταν αλλάξει σημαντικά η κατανομή των δεδομένων [9]. Η υψηλή διαστατικότητα, η ύπαρξη θορύβου, η ετερογένεια και η αραιότητα ορισμένων χαρακτηριστικών δημιουργούν πρόσθετες δυσκολίες στην αναγνώριση των πραγματικά σημαντικών αποκλίσεων [2], [6].

Ιδιαίτερα σημαντική πρόκληση αποτελεί και η σπανιότητα των περιστατικών απάτης. Στα περισσότερα πραγματικά οικονομικά σύνολα δεδομένων, οι δόλιες περιπτώσεις αποτελούν πολύ μικρό ποσοστό σε σχέση με τις νόμιμες συναλλαγές, δημιουργώντας έντονη ανισορροπία μεταξύ των κατηγοριών [6], [7]. Η κατάσταση αυτή μπορεί να επηρεάσει σημαντικά την απόδοση των μοντέλων, καθώς μια μέθοδος μπορεί να εμφανίζει υψηλή συνολική ακρίβεια χωρίς να εντοπίζει αποτελεσματικά τη μικρή κατηγορία που παρουσιάζει το μεγαλύτερο ενδιαφέρον [7].

Παράλληλα, η απόκτηση αξιόπιστων ετικετών για πραγματικές περιπτώσεις απάτης είναι δύσκολη. Ο χαρακτηρισμός μιας συναλλαγής ως δόλια συχνά απαιτεί προηγούμενη διερεύνηση, ανθρώπινη αξιολόγηση ή επιβεβαίωση από αρμόδιους ελεγκτικούς φορείς [6]. Επιπλέον, δεν είναι πάντοτε δυνατό να θεωρηθεί ότι μια μη επισημασμένη συναλλαγή είναι με βεβαιότητα νόμιμη, καθώς ενδέχεται να υπάρχουν περιστατικά απάτης που δεν έχουν ακόμη εντοπιστεί [1]. Το ίδιο πρόβλημα εμφανίζεται στις οικονομικές καταστάσεις, όπου η απουσία επίσημου χαρακτηρισμού μιας εταιρείας ως δόλια δεν εγγυάται την πλήρη απουσία λανθασμένων ή παραποιημένων οικονομικών πληροφοριών [1].

Η διαφορά μεταξύ ανακριβιών και επιβεβαιωμένων παρατυπιών ενισχύει ακόμη περισσότερο αυτή την αβεβαιότητα. Μία ουσιώδης ανακρίβεια μπορεί να επηρεάζει σημαντικά την αξιοπιστία των οικονομικών πληροφοριών χωρίς να περιλαμβάνει απαραίτητα αποδεδειγμένη δόλια πρόθεση [8]. Έτσι, η αυστηρή κατηγοριοποίηση των πραγματικών οικονομικών δεδομένων σε δύο πλήρως καθορισμένες ομάδες, δόλιων και μη δόλιων, μπορεί να είναι δύσκολη σε πραγματικές συνθήκες.

Οι παραδοσιακοί μηχανισμοί ανίχνευσης βασίζονται συχνά σε χειροκίνητους ελέγχους, προκαθορισμένους κανόνες και στατιστικές μεθόδους. Τα συστήματα βασισμένα σε κανόνες μπορούν να εντοπίζουν γρήγορα γνωστά πρότυπα απάτης μέσω συγκεκριμένων ορίων ή συνθηκών, ενώ παράλληλα χαρακτηρίζονται από σχετικά υψηλή ερμηνευσιμότητα [9]. Η αποτελεσματικότητά τους όμως εξαρτάται σημαντικά από την εμπειρία των ειδικών που διαμορφώνουν τους κανόνες. Επιπλέον, η συνεχής εμφάνιση νέων προτύπων απάτης απαιτεί την τακτική τροποποίηση και συντήρηση αυτών των κανόνων, περιορίζοντας την ευελιξία τους [9].

Οι καθαρά χειροκίνητες διαδικασίες αντιμετωπίζουν επίσης δυσκολίες όταν ο όγκος των διαθέσιμων δεδομένων αυξάνεται σημαντικά. Η εξέταση μεγάλου αριθμού οικονομικών καταστάσεων ή λογιστικών εγγραφών μπορεί να είναι χρονοβόρα και να απαιτεί σημαντικούς ανθρώπινους πόρους [1]. Παρόμοιοι περιορισμοί εμφανίζονται όταν χειροκίνητοι έλεγχοι και στατικοί κανόνες εφαρμόζονται σε σύνθετα δεδομένα γενικού καθολικού, στα οποία οι πιθανές αποκλίσεις δεν περιορίζονται πάντοτε σε απλές και εκ των προτέρων γνωστές συνθήκες [2].

Η αξιοποίηση της μηχανικής μάθησης και της ανάλυσης δεδομένων προσφέρει τη δυνατότητα επεξεργασίας μεγάλου όγκου οικονομικών πληροφοριών και αναζήτησης περισσότερο σύνθετων προτύπων [1], [8]. Τα μοντέλα μπορούν να αξιοποιήσουν μεγάλο αριθμό χαρακτηριστικών και να εντοπίσουν σχέσεις ή αλληλεπιδράσεις που θα ήταν δύσκολο να προσδιοριστούν εξ αρχής μέσω συμβατικών κανόνων [8]. Η δυνατότητα αυτή είναι ιδιαίτερα χρήσιμη σε περιβάλλοντα οικονομικών συναλλαγών όπου τα δεδομένα είναι μεγάλης κλίμακας, πολυδιάστατα και μεταβάλλονται συνεχώς [9]. Η τεχνολογία ωστόσο δεν λειτουργεί ανεξάρτητα από τους υπόλοιπους μηχανισμούς ελέγχου. Η πρόληψη και η ανίχνευση απάτης μπορούν να συνδυάζουν τεχνολογικά εργαλεία, εσωτερικό και

εξωτερικό έλεγχο, εταιρική διακυβέρνηση, πολιτικές και διαδικασίες ελέγχου [4]. Τα αυτοματοποιημένα συστήματα μπορούν να λειτουργήσουν ως μηχανισμοί υποστήριξης, επισημαίνοντας ύποπτες περιπτώσεις και βοηθώντας τους ελεγκτές να επικεντρώσουν την προσοχή τους σε συναλλαγές ή οικονομικές αναφορές που εμφανίζουν αυξημένες ενδείξεις κινδύνου [5]. Η τελική αξιολόγηση μιας ύποπτης περίπτωσης εξακολουθεί επομένως να απαιτεί ελεγκτική διερεύνηση και επαγγελματική κρίση.

Η εφαρμογή τεχνικών μηχανικής μάθησης στην ανίχνευση οικονομικής απάτης έχει επικεντρωθεί σε μεγάλο βαθμό σε επιβλεπόμενες προσεγγίσεις, οι οποίες αξιοποιούν ήδη χαρακτηρισμένες δόλιες και νόμιμες περιπτώσεις για την εκπαίδευση του μοντέλου [1], [10]. Ωστόσο, η περιορισμένη διαθεσιμότητα αξιόπιστων ετικετών, η έντονη ανισορροπία μεταξύ των κατηγοριών και η συνεχής εμφάνιση νέων μορφών απάτης δημιουργούν σημαντικούς περιορισμούς για την αποκλειστική χρήση αυτής της προσέγγισης [6], [7]. Για τον λόγο αυτό, οι μη επιβλεπόμενες και ημι-επιβλεπόμενες τεχνικές αποκτούν ιδιαίτερο ενδιαφέρον σε περιπτώσεις όπου οι πραγματικές περιπτώσεις απάτης είναι άγνωστες, περιορισμένες ή δύσκολο να επιβεβαιωθούν [1], [6], [7], [10].

Συνολικά, η ανίχνευση ανωμαλιών προσφέρει ένα διαφορετικό πλαίσιο αντιμετώπισης του προβλήματος, καθώς δεν απαιτεί απαραίτητα πλήρη γνώση όλων των πιθανών μορφών απάτης. Η αναζήτηση αποκλίσεων από τη συνήθη συμπεριφορά μπορεί να συμβάλει στον εντοπισμό προηγουμένως άγνωστων ή σπάνιων προτύπων, τα οποία στη συνέχεια μπορούν να αξιολογηθούν από τον ελεγκτή. Στο πλαίσιο των οικονομικών και λογιστικών δεδομένων, η προσέγγιση αυτή είναι ιδιαίτερα κατάλληλη όταν τα επιβεβαιωμένα περιστατικά απάτης είναι περιορισμένα και όταν ο βασικός στόχος είναι αρχικά ο εντοπισμός συναλλαγών και λογιστικών εγγραφών που παρουσιάζουν ασυνήθιστη συμπεριφορά. Η τελική διάκριση μεταξύ πραγματικής απάτης, λογιστικού σφάλματος και νόμιμης αλλά ασυνήθιστης δραστηριότητας απαιτεί περαιτέρω διερεύνηση, γεγονός που καθιστά την ανίχνευση ανωμαλιών περισσότερο έναν μηχανισμό εντοπισμού πιθανών κινδύνων παρά έναν μηχανισμό οριστικής ταξινόμησης μιας συναλλαγής ως δόλια.

2.2 Μηχανική Μάθηση για Ανίχνευση Απάτης

Η συνεχής αύξηση του όγκου και της πολυπλοκότητας των οικονομικών συναλλαγών έχει ενισχύσει τη χρήση της μηχανικής μάθησης για την ανίχνευση απάτης. Σε αντίθεση με τις προσεγγίσεις που βασίζονται αποκλειστικά σε προκαθορισμένους κανόνες, τα μοντέλα μηχανικής μάθησης μπορούν να αξιοποιούν μεγάλο αριθμό χαρακτηριστικών και ιστορικών παρατηρήσεων, ώστε να αναγνωρίζουν σύνθετες σχέσεις και πρότυπα συμπεριφοράς μέσα στα δεδομένα [11]. Η δυνατότητα αυτή αποκτά ιδιαίτερη σημασία σε οικονομικά περιβάλλοντα όπου οι συναλλαγές παράγονται με υψηλή συχνότητα και οι μέθοδοι που χρησιμοποιούνται για την πραγματοποίηση απάτης μεταβάλλονται διαρκώς.

Οι προσεγγίσεις μηχανικής μάθησης που εφαρμόζονται στην ανίχνευση απάτης μπορούν να διακριθούν κυρίως σε επιβλεπόμενη, μη επιβλεπόμενη και ημι-επιβλεπόμενη μάθηση. Η καταλληλότητα κάθε προσέγγισης επηρεάζεται από τη διαθεσιμότητα και την αξιοπιστία των ετικετών, την αναλογία μεταξύ νόμιμων και δόλιων συναλλαγών, καθώς και από το κατά πόσο τα χαρακτηριστικά των δεδομένων μεταβάλλονται με την πάροδο του χρόνου [11], [16]. Οι τρεις αυτές κατηγορίες ακολουθούν διαφορετική λογική εκπαίδευσης και αντιμετωπίζουν με διαφορετικό τρόπο τις ιδιαίτερες δυσκολίες των οικονομικών δεδομένων.

Η επιβλεπόμενη μάθηση βασίζεται σε δεδομένα για τα οποία είναι ήδη γνωστή η κατηγορία κάθε παρατήρησης, όπως για παράδειγμα αν μια συναλλαγή είναι νόμιμη ή δόλια. Κατά την εκπαίδευση, το μοντέλο επιχειρεί να μάθει τη σχέση μεταξύ των χαρακτηριστικών των συναλλαγών και των αντίστοιχων ετικετών, ώστε στη συνέχεια να μπορεί να ταξινομεί νέες παρατηρήσεις. Στην ανίχνευση οικονομικής απάτης έχουν χρησιμοποιηθεί μεταξύ άλλων η Λογιστική Παλινδρόμηση, τα Δέντρα Απόφασης, τα Τυχαία Δάση και οι Μηχανές Διανυσμάτων Υποστήριξης, καθώς και πιο σύνθετες τεχνικές μηχανικής και βαθιάς μάθησης [11]. Όταν υπάρχουν επαρκή, αξιόπιστα και αντιπροσωπευτικά επισημασμένα δεδομένα, η επιβλεπόμενη μάθηση μπορεί να αξιοποιήσει άμεσα τη γνώση που παρέχουν προηγουμένως επιβεβαιωμένες περιπτώσεις απάτης.

Η εφαρμογή της, ωστόσο, περιορίζεται από ορισμένα χαρακτηριστικά των πραγματικών οικονομικών δεδομένων. Οι επιβεβαιωμένες περιπτώσεις απάτης αποτελούν συνήθως πολύ μικρό ποσοστό του συνόλου των συναλλαγών, με αποτέλεσμα να δημιουργούνται έντονα ανισόρροπα σύνολα δεδομένων [11], [12]. Η κατηγορία των νόμιμων συναλλαγών κυριαρχεί αριθμητικά, ενώ οι δόλιες συναλλαγές

αποτελούν μια σπάνια κατηγορία. Υπό αυτές τις συνθήκες, ένας συμβατικός αλγόριθμος ταξινόμησης μπορεί να προσαρμοστεί περισσότερο στα χαρακτηριστικά της πλειοψηφικής κατηγορίας και να παρουσιάσει δυσκολία στην εκμάθηση των προτύπων της σπανιότερης κατηγορίας [13].

Για την αντιμετώπιση της ανισορροπίας έχουν αναπτυχθεί διαφορετικές τεχνικές επαναδειγματοληψίας. Η υπερδειγματοληψία αυξάνει την παρουσία της μειοψηφικής κατηγορίας στα δεδομένα εκπαίδευσης, ενώ η υποδειγματοληψία περιορίζει τον αριθμό των παρατηρήσεων της πλειοψηφικής κατηγορίας. Επιπλέον, τεχνικές όπως το SMOTE δημιουργούν συνθετικές παρατηρήσεις με βάση ήδη υπάρχοντα δείγματα της μειοψηφικής κατηγορίας [13]. Ωστόσο, η τεχνητή μεταβολή της αρχικής κατανομής των δεδομένων δεν εγγυάται σε κάθε περίπτωση καλύτερη λειτουργία του μοντέλου. Η υποδειγματοληψία μπορεί να οδηγήσει σε απώλεια χρήσιμης πληροφορίας, ενώ η υπερδειγματοληψία μεταβάλλει την κατανομή του συνόλου εκπαίδευσης σε σχέση με εκείνη που συναντάται στην πραγματική λειτουργία ενός συστήματος [13], [14]. Για τον λόγο αυτό, τέτοιες τεχνικές δεν αποτελούν καθολική λύση αλλά χρειάζεται να εξετάζονται σε συνάρτηση με τα χαρακτηριστικά του εκάστοτε προβλήματος.

Μια δεύτερη σημαντική δυσκολία σχετίζεται με τη μεταβολή των προτύπων των δεδομένων με την πάροδο του χρόνου. Η συμπεριφορά των χρηστών δεν παραμένει σταθερή, ενώ παράλληλα οι μέθοδοι απάτης προσαρμόζονται ώστε να αποφεύγουν τους υπάρχοντες μηχανισμούς ελέγχου. Το φαινόμενο αυτό περιγράφεται συχνά ως *concept drift* και μπορεί σταδιακά να μειώσει την αντιπροσωπευτικότητα ενός μοντέλου που έχει εκπαιδευτεί αποκλειστικά σε παλαιότερα δεδομένα [11], [12]. Για την αντιμετώπιση τέτοιων μεταβολών μπορούν να χρησιμοποιηθούν προσεγγίσεις σταδιακής ή συνεχούς μάθησης, στις οποίες το μοντέλο ενημερώνεται καθώς εμφανίζονται νεότερα δεδομένα [12].

Η ανάγκη προσαρμογής γίνεται ακόμη σημαντικότερη όταν η ανίχνευση πραγματοποιείται σε πραγματικό χρόνο. Οι χρηματοοικονομικές συναλλαγές μπορούν να δημιουργούν συνεχή ροή δεδομένων και επομένως οι αποφάσεις πρέπει συχνά να λαμβάνονται με μικρή χρονική καθυστέρηση [11], [17]. Παράλληλα, η πραγματική κατάσταση μιας συναλλαγής δεν είναι πάντοτε γνωστή τη στιγμή που πραγματοποιείται. Η επιβεβαίωση μιας απάτης μπορεί να απαιτεί διερεύνηση και να γίνει διαθέσιμη αρκετό χρονικό διάστημα αργότερα [11]. Η καθυστέρηση στην απόκτηση αξιόπιστων ετικετών δυσκολεύει την άμεση ενημέρωση των επιβλεπόμενων μοντέλων και περιορίζει τη δυνατότητα συνεχούς εκπαίδευσής τους με πλήρως επιβεβαιωμένα δεδομένα.

Οι περιορισμοί αυτοί ενισχύουν το ενδιαφέρον για τη μη επιβλεπόμενη μάθηση, στην οποία δεν απαιτείται να είναι γνωστή εκ των προτέρων η κατηγορία κάθε παρατήρησης. Αντί για την εκμάθηση ενός άμεσου διαχωρισμού μεταξύ νόμιμων και δόλιων συναλλαγών, το μοντέλο εξετάζει τη δομή και την κατανομή των δεδομένων, αναζητώντας αποκλίσεις, ασυνήθιστα πρότυπα ή ομάδες παρατηρήσεων που διαφέρουν από τη συνήθη συμπεριφορά [16], [17]. Η προσέγγιση αυτή είναι ιδιαίτερα χρήσιμη όταν οι επιβεβαιωμένες περιπτώσεις απάτης είναι περιορισμένες ή όταν υπάρχει ενδιαφέρον για τον εντοπισμό νέων μορφών ύποπτης δραστηριότητας που δεν περιλαμβάνονται στα διαθέσιμα δεδομένα εκπαίδευσης.

Στο πλαίσιο αυτό, η ανίχνευση απάτης μπορεί να αντιμετωπιστεί ως πρόβλημα ανίχνευσης ανωμαλιών, όπου η βασική επιδίωξη είναι η αναγνώριση συναλλαγών που αποκλίνουν σημαντικά από τα πρότυπα της συνηθισμένης δραστηριότητας. Οι παρατηρήσεις αυτές δεν θεωρούνται αυτομάτως αποδεδειγμένες περιπτώσεις απάτης, αλλά ύποπτες συναλλαγές που μπορούν να προωθηθούν για περαιτέρω έλεγχο. Η διάκριση αυτή είναι σημαντική, καθώς μια ασυνήθιστη οικονομική συμπεριφορά μπορεί να οφείλεται σε νόμιμους παράγοντες και να μην συνδέεται απαραίτητα με δόλια πρόθεση [16].

Συναφής με τη μη επιβλεπόμενη ανίχνευση είναι η ταξινόμηση μίας κλάσης. Η συγκεκριμένη προσέγγιση μπορεί να χρησιμοποιηθεί όταν υπάρχει επαρκής αριθμός παρατηρήσεων για μία κατηγορία, συνήθως τη φυσιολογική, ενώ οι παρατηρήσεις της άλλης κατηγορίας είναι ελάχιστες ή δεν είναι διαθέσιμες [15]. Το μοντέλο επιχειρεί να περιγράψει τη γνωστή συμπεριφορά και να δημιουργήσει ένα όριο γύρω από αυτή, ώστε νέες παρατηρήσεις που αποκλίνουν σημαντικά να θεωρούνται πιθανές ανωμαλίες. Η λογική αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον στην οικονομική απάτη, καθώς οι νόμιμες συναλλαγές είναι συνήθως πολύ περισσότερες από τις επιβεβαιωμένες περιπτώσεις απάτης [15].

Η μη επιβλεπόμενη μάθηση μειώνει την εξάρτηση από τη χειροκίνητη επισήμανση μεγάλου αριθμού συναλλαγών και μπορεί να συμβάλει στην αναγνώριση συμπεριφορών που δεν είχαν εμφανιστεί στα προηγούμενα δεδομένα [16]. Παράλληλα, όμως, η απουσία άμεσης πληροφορίας σχετικά με το ποια παρατήρηση είναι πραγματικά δόλια καθιστά δυσκολότερο τον διαχωρισμό μιας πραγματικής απάτης από μια νόμιμη αλλά ασυνήθιστη συναλλαγή. Η επιλογή του ορίου που χρησιμοποιείται για τον

χαρακτηρισμό μιας παρατήρησης ως ανωμαλίας αποτελεί συνεπώς σημαντική παράμετρο τέτοιων προσεγγίσεων [16].

Εκτός από τα χαρακτηριστικά των μεμονωμένων συναλλαγών, σε πραγματικά οικονομικά συστήματα υπάρχουν και σύνθετες σχέσεις μεταξύ χρηστών, λογαριασμών, εμπορών και συναλλαγών. Η συμπεριφορά μιας οντότητας δεν είναι πάντοτε ανεξάρτητη από τη συμπεριφορά άλλων οντοτήτων, ενώ οργανωμένες μορφές απάτης μπορούν να δημιουργούν συνδεδεμένες ομάδες λογαριασμών ή επαναλαμβανόμενα συναλλακτικά πρότυπα [17]. Η αναπαράσταση των δεδομένων μέσω γράφων επιτρέπει να εξετάζονται ταυτόχρονα τα χαρακτηριστικά των οντοτήτων και οι μεταξύ τους σχέσεις. Με αυτόν τον τρόπο μπορούν να εντοπιστούν ύποπτες δομές ή ομάδες που δεν είναι εύκολα ορατές όταν κάθε συναλλαγή εξετάζεται μεμονωμένα [17].

Μεταξύ της επιβλεπόμενης και της μη επιβλεπόμενης προσέγγισης βρίσκεται η ημι-επιβλεπόμενη μάθηση. Σε αυτή την περίπτωση χρησιμοποιείται ένας περιορισμένος αριθμός επισημασμένων παρατηρήσεων μαζί με μεγαλύτερο πλήθος μη επισημασμένων δεδομένων [18]. Η προσέγγιση αυτή είναι κατάλληλη για περιπτώσεις όπου η απόκτηση αξιόπιστων ετικετών είναι δαπανηρή ή χρονοβόρα, καθώς επιτρέπει την αξιοποίηση της πληροφορίας που περιέχεται στα μη επισημασμένα δεδομένα χωρίς να εγκαταλείπεται πλήρως η καθοδήγηση από γνωστές περιπτώσεις [18]. Στην ανίχνευση απάτης, όπου η επιβεβαίωση ύποπτων περιπτώσεων συχνά απαιτεί ανθρώπινη διερεύνηση, η δυνατότητα αυτή μπορεί να περιορίσει την εξάρτηση από μεγάλα πλήρως επισημασμένα σύνολα δεδομένων.

Η επιλογή μεταξύ επιβλεπόμενης, μη επιβλεπόμενης και ημι-επιβλεπόμενης μάθησης δεν μπορεί επομένως να βασιστεί στην υπόθεση ότι μία από τις τρεις προσεγγίσεις είναι πάντοτε ανώτερη. Επηρεάζεται από τη φύση των διαθέσιμων δεδομένων, την έκταση της ανισορροπίας μεταξύ των κατηγοριών, την ποιότητα και την καθυστέρηση των ετικετών, τη μεταβολή της συμπεριφοράς με την πάροδο του χρόνου και τις λειτουργικές απαιτήσεις του συστήματος [11], [16]–[18]. Όταν διατίθεται μεγάλο και αξιόπιστο ιστορικό επισημασμένων συναλλαγών, η επιβλεπόμενη μάθηση μπορεί να αξιοποιήσει άμεσα αυτή την πληροφορία. Αντίθετα, όταν οι ετικέτες είναι περιορισμένες, καθυστερημένες ή απουσιάζουν, οι μη επιβλεπόμενες και ημι-επιβλεπόμενες τεχνικές αποκτούν μεγαλύτερη πρακτική σημασία.

Στην παρούσα εργασία δίνεται ιδιαίτερη έμφαση στη μη επιβλεπόμενη ανίχνευση ανωμαλιών, καθώς επιτρέπει την εξέταση οικονομικών συναλλαγών χωρίς να προϋποθέτει πλήρη και αξιόπιστη επισήμανσή τους. Επιπλέον, διαφορετικοί αλγόριθμοι μπορούν να εξετάζουν διαφορετικές μορφές απόκλισης μέσα στα δεδομένα, γεγονός που δημιουργεί τη βάση για τη συνδυασμένη αξιοποίησή τους. Τα επιμέρους μοντέλα ανίχνευσης ανωμαλιών και οι τεχνικές συνδυασμού τους παρουσιάζονται αναλυτικά στην επόμενη ενότητα.

2.3 Μοντέλα Ανίχνευσης Ανωμαλιών και Τεχνικές Συνδυασμού

Η μη επιβλεπόμενη ανίχνευση ανωμαλιών μπορεί να υλοποιηθεί με διαφορετικές αλγοριθμικές προσεγγίσεις, οι οποίες εξετάζουν διαφορετικές ιδιότητες της κατανομής και της δομής των δεδομένων. Ανάλογα με τη μέθοδο, μια παρατήρηση μπορεί να θεωρηθεί ασυνήθιστη λόγω της ευκολίας με την οποία απομονώνεται από το υπόλοιπο σύνολο, της τοπικής πυκνότητας που παρουσιάζει σε σχέση με τις γειτονικές παρατηρήσεις, της συχνότητας εμφάνισης των χαρακτηριστικών της ή της αδυναμίας ενός μοντέλου να την ανακατασκευάσει με ακρίβεια [19]–[21]. Η διαφορετική αυτή λογική καθιστά τα μοντέλα δυνητικά συμπληρωματικά, καθώς δεν είναι απαραίτητο να εντοπίζουν τις ίδιες μορφές ανωμαλιών. Στην παρούσα ενότητα παρουσιάζονται τα Isolation Forest, Local Outlier Factor, Histogram-Based Outlier Score και Autoencoders, ενώ στη συνέχεια εξετάζονται τεχνικές συνδυασμού των επιμέρους αποτελεσμάτων τους.

2.3.1 Isolation Forest

Το IF αποτελεί μέθοδο μη επιβλεπόμενης ανίχνευσης ανωμαλιών, η οποία βασίζεται σε διαφορετική λογική από προσεγγίσεις που απαιτούν άμεσο υπολογισμό αποστάσεων ή εκτίμηση της πυκνότητας των δεδομένων. Βασική του αρχή είναι ότι οι ανωμαλίες αποτελούν συνήθως μικρό μέρος του συνόλου και παρουσιάζουν χαρακτηριστικά που τις διαφοροποιούν από την πλειονότητα των παρατηρήσεων. Κατά συνέπεια, αναμένεται να μπορούν να απομονωθούν από τα υπόλοιπα δεδομένα με μικρότερο αριθμό διαδοχικών διαχωρισμών [19], [21], [22].

Η διαδικασία απομόνωσης πραγματοποιείται μέσω ενός συνόλου δυαδικών δέντρων απομόνωσης (Isolation Trees – iTrees). Για την κατασκευή κάθε δέντρου χρησιμοποιείται ένα τυχαίο υποσύνολο των διαθέσιμων δεδομένων. Σε κάθε κόμβο επιλέγεται τυχαία ένα χαρακτηριστικό και στη συνέχεια μία τιμή διαχωρισμού μεταξύ της ελάχιστης και της μέγιστης τιμής που εμφανίζει το συγκεκριμένο χαρακτηριστικό στις παρατηρήσεις του κόμβου. Οι παρατηρήσεις διαχωρίζονται σε δύο επιμέρους ομάδες ανάλογα με τη θέση τους ως προς την επιλεγμένη τιμή και η διαδικασία επαναλαμβάνεται αναδρομικά μέχρι να απομονωθεί μια παρατήρηση ή να επιτευχθεί το προκαθορισμένο μέγιστο βάθος του δέντρου [21], [22].

Η βασική ποσότητα που χρησιμοποιείται για την εκτίμηση του βαθμού ανωμαλίας μιας παρατήρησης x είναι το μήκος διαδρομής $h(x)$, δηλαδή ο αριθμός των διαχωρισμών που απαιτούνται από τη ρίζα ενός δέντρου μέχρι το σημείο στο οποίο η συγκεκριμένη παρατήρηση απομονώνεται. Παρατηρήσεις που ανήκουν σε πυκνές περιοχές και παρουσιάζουν παρόμοια χαρακτηριστικά με μεγάλο αριθμό άλλων παρατηρήσεων απαιτούν συνήθως περισσότερους διαχωρισμούς. Αντίθετα, μια σπάνια ή έντονα διαφοροποιημένη παρατήρηση τείνει να απομονώνεται σε μικρότερο βάθος [19], [21], [22].

Επειδή η κατασκευή κάθε δέντρου περιλαμβάνει τυχαίες επιλογές, ένα μεμονωμένο δέντρο δεν επαρκεί για τη σταθερή εκτίμηση της ανωμαλίας. Για τον λόγο αυτό, το IF αποτελείται από ένα σύνολο διαφορετικών δέντρων απομόνωσης. Για κάθε παρατήρηση υπολογίζεται το μέσο μήκος διαδρομής της σε όλα τα δέντρα και η πληροφορία αυτή μετατρέπεται σε βαθμολογία ανωμαλίας (anomaly score) [21], [22].

Η βαθμολογία ανωμαλίας μιας παρατήρησης x μπορεί να εκφραστεί ως:

$$s(x) = 2^{-\frac{E[h(x)]}{c(\psi)}} \quad (2.1)$$

όπου $E[h(x)]$ είναι το μέσο μήκος διαδρομής της παρατήρησης x στο σύνολο των δέντρων, ψ είναι το μέγεθος του τυχαίου υποδείγματος που χρησιμοποιείται για την κατασκευή κάθε δέντρου και $c(\psi)$ αποτελεί παράγοντα κανονικοποίησης που αντιστοιχεί στο αναμενόμενο μήκος διαδρομής σε ένα δυαδικό δέντρο [21], [22].

Ο παράγοντας κανονικοποίησης υπολογίζεται ως:

$$c(\psi) = 2H(\psi - 1) - \frac{2(\psi - 1)}{\psi} \quad (2.2)$$

όπου $H(\psi - 1)$ είναι ο αρμονικός αριθμός που αντιστοιχεί στο μέγεθος του υποδείγματος [22]. Η κανονικοποίηση αυτή επιτρέπει τη σύγκριση των μηκών διαδρομής και τη μετατροπή τους σε κοινή βαθμολογία ανωμαλίας.

Η ερμηνεία της βαθμολογίας συνδέεται άμεσα με την ευκολία απομόνωσης μιας παρατήρησης. Όταν το $E[h(x)]$ είναι μικρό σε σχέση με το $c(\psi)$, η τιμή του $s(x)$ προσεγγίζει το 1 και η παρατήρηση θεωρείται περισσότερο πιθανό να αποτελεί ανωμαλία. Αντίθετα, μεγαλύτερο μέσο μήκος διαδρομής οδηγεί σε μικρότερη βαθμολογία, υποδηλώνοντας μεγαλύτερη ομοιότητα με τη συνηθισμένη δομή των δεδομένων [21], [22]. Επομένως, η βαθμολογία δεν εκφράζει άμεσα την πιθανότητα ύπαρξης απάτης, αλλά τον βαθμό στον οποίο μια παρατήρηση απομονώνεται ευκολότερα από το υπόλοιπο σύνολο.

Ένα ακόμη βασικό χαρακτηριστικό του IF είναι η χρήση τυχαίας υποδειγματοληψίας. Δεν απαιτείται κάθε δέντρο να κατασκευάζεται χρησιμοποιώντας ολόκληρο το διαθέσιμο σύνολο δεδομένων. Μικρότερα τυχαία υποσύνολα μπορούν να χρησιμοποιηθούν για την ανάπτυξη των επιμέρους δέντρων χωρίς να είναι απαραίτητη η πλήρης μοντελοποίηση της κατανομής του συνόλου [19], [22]. Η προσέγγιση αυτή μειώνει τις υπολογιστικές απαιτήσεις και ενισχύει την καταλληλότητα του αλγορίθμου για μεγάλα σύνολα δεδομένων.

Η υπολογιστική αποδοτικότητα αποτελεί ένα από τα σημαντικότερα πλεονεκτήματα της μεθόδου. Σε αντίθεση με αλγορίθμους που χρειάζονται να υπολογίσουν αποστάσεις ή τοπικές πυκνότητες μεταξύ μεγάλου αριθμού παρατηρήσεων, το IF βασίζεται κυρίως στη δημιουργία τυχαίων διαχωρισμών. Η προσέγγιση αυτή χαρακτηρίζεται από καλή κλιμάκωση ως προς το μέγεθος των δεδομένων και έχει χρησιμοποιηθεί σε εφαρμογές όπου ο μεγάλος αριθμός παρατηρήσεων καθιστά υπολογιστικά

απαιτητικές άλλες μεθόδους ανίχνευσης ανωμαλιών [19], [21]. Παράλληλα, η μέθοδος απαιτεί σχετικά περιορισμένο αριθμό βασικών υπερπαραμέτρων [23].

Η εφαρμογή του IF παρουσιάζει ιδιαίτερο ενδιαφέρον σε οικονομικά και λογιστικά δεδομένα. Σε δεδομένα γενικού καθολικού, όπου μπορεί να υπάρχουν χιλιάδες ή εκατομμύρια λογιστικές εγγραφές και οι πραγματικές ανωμαλίες δεν είναι πάντοτε γνωστές εκ των προτέρων, η μέθοδος μπορεί να χρησιμοποιηθεί για την αναζήτηση εγγραφών που αποκλίνουν από τη συνηθισμένη λογιστική δραστηριότητα [20]. Οι εγγραφές με υψηλότερο βαθμό ανωμαλίας μπορούν στη συνέχεια να αποτελέσουν αντικείμενο πιο στοχευμένου ελέγχου, περιορίζοντας την ανάγκη αποκλειστικής εξάρτησης από τυχαία δειγματοληψία ή προκαθορισμένους κανόνες.

Η μη επιβλεπόμενη φύση του αλγορίθμου είναι επίσης σημαντική σε τέτοια περιβάλλοντα, καθώς δεν απαιτείται η χρήση προκαθορισμένων ετικετών κατά την εκπαίδευση. Αυτό επιτρέπει την αναζήτηση αποκλίσεων ακόμη και όταν οι διαθέσιμες επιβεβαιωμένες περιπτώσεις απάτης είναι περιορισμένες ή όταν υπάρχουν άγνωστα πρότυπα ανωμαλιών μέσα στα δεδομένα [20]. Ωστόσο, μια λογιστική εγγραφή με υψηλή βαθμολογία ανωμαλίας δεν πρέπει να θεωρείται αυτομάτως δόλια. Η έξοδος του μοντέλου λειτουργεί περισσότερο ως ένδειξη ότι η συγκεκριμένη εγγραφή παρουσιάζει ασυνήθιστη συμπεριφορά και ενδέχεται να απαιτεί περαιτέρω διερεύνηση.

Παρά τα πλεονεκτήματά του, το αρχικό IF παρουσιάζει ορισμένους περιορισμούς που σχετίζονται κυρίως με τον τρόπο κατασκευής των διαχωρισμών. Σε κάθε κόμβο χρησιμοποιείται ένα μόνο χαρακτηριστικό και οι διαχωρισμοί πραγματοποιούνται παράλληλα προς τους άξονες του χώρου χαρακτηριστικών. Η συγκεκριμένη διαδικασία μπορεί να δυσκολεύει την απομόνωση ανωμαλιών των οποίων η διαφοροποίηση εξαρτάται από σύνθετους συνδυασμούς πολλών χαρακτηριστικών ή από μη γραμμικές σχέσεις μεταξύ τους [21], [22]. Για τον λόγο αυτό έχουν αναπτυχθεί επεκτάσεις του αρχικού μοντέλου που χρησιμοποιούν διαφορετικούς μηχανισμούς διαχωρισμού ή νέες αναπαραστάσεις των δεδομένων.

Επιπλέον, το IF βασίζεται κυρίως σε μια συνολική εκτίμηση της δυνατότητας απομόνωσης, γεγονός που μπορεί να περιορίσει την ικανότητά του να εντοπίζει ορισμένες τοπικές ανωμαλίες. Μια παρατήρηση μπορεί να μην είναι ιδιαίτερα ασυνήθιστη σε σχέση με ολόκληρο το σύνολο δεδομένων, αλλά να αποκλίνει σημαντικά από τις παρατηρήσεις της άμεσης περιοχής στην οποία βρίσκεται. Δυσκολίες έχουν επίσης αναφερθεί σε περιπτώσεις δεδομένων με πολλές διαφορετικές ομάδες, σε υψηλές διαστάσεις με μεγάλο αριθμό μη σχετικών χαρακτηριστικών και σε κατανομές όπου οι ανωμαλίες καλύπτονται από τον άξονα-παράλληλο μηχανισμό διαχωρισμού [19].

Οι περιορισμοί αυτοί έχουν οδηγήσει στην ανάπτυξη παραλλαγών όπως το EIF, το GIF και πιο πρόσφατες επεκτάσεις που ενσωματώνουν διαφορετικούς μηχανισμούς αναπαράστασης και διαχωρισμού [21], [22]. Οι μέθοδοι αυτές διατηρούν τη βασική αρχή της απομόνωσης, αλλά επιχειρούν να μειώσουν την επίδραση των περιορισμών που προκύπτουν από τους αποκλειστικά άξονοπαράλληλους διαχωρισμούς.

Συνολικά, το IF αποτελεί μια αποδοτική μέθοδο μη επιβλεπόμενης ανίχνευσης ανωμαλιών, στην οποία η ασυνήθιστη συμπεριφορά μιας παρατήρησης εκτιμάται με βάση την ευκολία απομόνωσής της. Η χρήση πολλαπλών τυχαίων δέντρων, μικρών υποδειγμάτων και βαθμολογίας βασισμένης στο μέσο μήκος διαδρομής προσφέρει έναν σχετικά απλό και κλιμακούμενο μηχανισμό για την αναγνώριση αποκλίσεων [19], [21], [22]. Παράλληλα, οι περιορισμοί του ως προς την τοπική δομή των δεδομένων δημιουργούν ενδιαφέρον για τη σύγκρισή του με μεθόδους που χρησιμοποιούν διαφορετική λογική ανίχνευσης, όπως το LOF, το οποίο εξετάζεται στην επόμενη υποενότητα.

2.3.2 Local Outlier Factor

Το LOF αποτελεί μη επιβλεπόμενη μέθοδο ανίχνευσης ανωμαλιών που βασίζεται στην τοπική πυκνότητα των δεδομένων. Σε αντίθεση με προσεγγίσεις που αξιολογούν μια παρατήρηση σε σχέση με το σύνολο της κατανομής, το LOF εξετάζει τη συμπεριφορά της σε σχέση με τις γειτονικές της παρατηρήσεις. Η βασική αρχή είναι ότι μια παρατήρηση μπορεί να θεωρηθεί ανώμαλη όταν βρίσκεται σε περιοχή σημαντικά μικρότερης πυκνότητας από τις περιοχές στις οποίες βρίσκονται οι πλησιέστεροι γείτονές της [24], [26]. Με αυτόν τον τρόπο, ο αλγόριθμος είναι σχεδιασμένος κυρίως για τον εντοπισμό τοπικών ανωμαλιών, οι οποίες ενδέχεται να μην εμφανίζονται ως ασυνήθιστες όταν εξετάζεται αποκλειστικά η συνολική κατανομή των δεδομένων.

Η έννοια της τοπικότητας διαφοροποιεί τον LOF από μεθόδους καθολικής ανίχνευσης. Μια παρατήρηση μπορεί να βρίσκεται σχετικά κοντά σε ένα σύνολο φυσιολογικών παρατηρήσεων και επομένως να μην εμφανίζεται ως απομακρυσμένη από το σύνολο δεδομένων. Εάν όμως η πυκνότητα γύρω από την ίδια είναι αισθητά χαμηλότερη από την πυκνότητα των γειτονικών περιοχών, μπορεί να χαρακτηριστεί ως πιθανή ανωμαλία [24]. Η ιδιότητα αυτή είναι ιδιαίτερα χρήσιμη σε δεδομένα που περιλαμβάνουν ομάδες διαφορετικής πυκνότητας, καθώς η έννοια της ανωμαλίας προσδιορίζεται σε σχέση με το τοπικό περιβάλλον κάθε παρατήρησης και όχι με ένα ενιαίο καθολικό πρότυπο.

Η λειτουργία του LOF στηρίζεται στην έννοια των k πλησιέστερων γειτόνων (kNN). Για κάθε παρατήρηση p , προσδιορίζονται οι k πλησιέστερες παρατηρήσεις σύμφωνα με μια επιλεγμένη συνάρτηση απόστασης. Στην τυπική διατύπωση χρησιμοποιείται η Ευκλείδεια απόσταση, η οποία για δύο παρατηρήσεις p και o σε χώρο n χαρακτηριστικών εκφράζεται ως [24]:

$$d(p, o) = \sum_{i=1}^n (p_i - o_i)^2 \quad (2.3)$$

Η απόσταση της παρατήρησης p από τον k -οστό πλησιέστερο γείτονά της ονομάζεται k -distance, ενώ το σύνολο των παρατηρήσεων των οποίων η απόσταση από την p δεν υπερβαίνει αυτή την τιμή σχηματίζει την τοπική γειτονιά $N_k(p)$ [24]. Η επιλογή του k καθορίζει επομένως το μέγεθος της περιοχής που λαμβάνεται υπόψη για την εκτίμηση της τοπικής συμπεριφοράς.

Για να περιοριστεί η επίδραση πολύ μικρών αποστάσεων μεταξύ κοντινών παρατηρήσεων, ο LOF χρησιμοποιεί την έννοια της απόστασης προσβασιμότητας (reachability distance). Για δύο παρατηρήσεις p και o , η απόσταση αυτή ορίζεται ως [24]:

$$reach - dist_k(p, o) = \max\{k - distance(o), d(p, o)\} \quad (2.4)$$

Η συγκεκριμένη σχέση σημαίνει ότι η απόσταση προσβασιμότητας δεν μπορεί να είναι μικρότερη από την k -distance του γειτονικού σημείου o . Εάν η πραγματική απόσταση $d(p, o)$ είναι μεγαλύτερη, χρησιμοποιείται η πραγματική απόσταση. Αντίθετα, όταν η παρατήρηση p βρίσκεται πολύ κοντά στην o , χρησιμοποιείται η k -distance της o . Με αυτόν τον τρόπο περιορίζονται έντονες διακυμάνσεις που θα μπορούσαν να προκύψουν από πολύ μικρές διαφορές μεταξύ γειτονικών παρατηρήσεων [24].

Με βάση τις αποστάσεις προσβασιμότητας υπολογίζεται στη συνέχεια η τοπική πυκνότητα προσβασιμότητας (LRD) μιας παρατήρησης. Η ποσότητα αυτή ορίζεται ως το αντίστροφο της μέσης απόστασης προσβασιμότητας της παρατήρησης p από τους γείτονές της:

$$lrd_k(p) = \left(\frac{\sum_{o \in N_k(p)} reach - dist_k(p, o)}{|N_k(p)|} \right)^{-1} \quad (2.5)$$

Όσο μικρότερη είναι η μέση απόσταση της παρατήρησης από το τοπικό της περιβάλλον, τόσο μεγαλύτερη είναι η τοπική πυκνότητα προσβασιμότητας. Αντίθετα, μια παρατήρηση που βρίσκεται πιο απομονωμένη από τους γείτονές της παρουσιάζει μικρότερη τιμή τοπικής πυκνότητας [24]. Η συγκεκριμένη ποσότητα δεν χρησιμοποιείται ανεξάρτητα για τον χαρακτηρισμό μιας ανωμαλίας, αλλά συγκρίνεται με τις αντίστοιχες πυκνότητες των γειτονικών παρατηρήσεων.

Η τελική βαθμολογία LOF μιας παρατήρησης p προκύπτει από τη σύγκριση της τοπικής πυκνότητάς της με τις τοπικές πυκνότητες των k πλησιέστερων γειτόνων της και μπορεί να εκφραστεί ως [24]:

$$LOF_k(p) = \frac{1}{|N_k(p)|} \sum_{o \in N_k(p)} \frac{lrd_k(o)}{lrd_k(p)} \quad (2.6)$$

Η σχέση αυτή αποτελεί τον βασικό μαθηματικό ορισμό του μοντέλου. Ο LOF επομένως δεν εξετάζει απλώς εάν μια παρατήρηση βρίσκεται σε περιοχή μικρής πυκνότητας, αλλά κατά πόσο η πυκνότητά της είναι μικρότερη σε σχέση με την πυκνότητα της άμεσης γειτονιάς της [24], [25]. Η σχετική αυτή σύγκριση επιτρέπει στον αλγόριθμο να διακρίνει παρατηρήσεις που αποκλίνουν από ένα τοπικό πρότυπο, ακόμη και όταν δεν βρίσκονται σε ακραία θέση ως προς ολόκληρο το σύνολο δεδομένων.

Η ερμηνεία της βαθμολογίας LOF βασίζεται στην τιμή της σε σχέση με τη μονάδα. Όταν η βαθμολογία μιας παρατήρησης βρίσκεται κοντά στο 1, η τοπική της πυκνότητα είναι παρόμοια με εκείνη των γειτόνων της και επομένως η παρατήρηση θεωρείται συμβατή με το τοπικό της περιβάλλον. Τιμές σημαντικά μεγαλύτερες από 1 υποδηλώνουν ότι η παρατήρηση βρίσκεται σε περιοχή μικρότερης πυκνότητας από τους γείτονές της και συνεπώς παρουσιάζει αυξημένο βαθμό τοπικής ανωμαλίας [24], [27]. Αντίθετα, τιμές μικρότερες από 1 μπορούν να εμφανιστούν όταν η παρατήρηση βρίσκεται σε περιοχή μεγαλύτερης πυκνότητας από τη γειτονιά της. Η βαθμολογία αυτή εκφράζει βαθμό απόκλισης και όχι άμεση πιθανότητα ότι μια συναλλαγή αποτελεί απάτη.

Ένα από τα σημαντικότερα πλεονεκτήματα του LOF είναι η δυνατότητα εντοπισμού ανωμαλιών μέσα σε δεδομένα που παρουσιάζουν διαφορετικές τοπικές δομές. Οι μέθοδοι που χρησιμοποιούν ένα καθολικό κριτήριο μπορεί να χαρακτηρίσουν ως ασυνήθιστες φυσιολογικές παρατηρήσεις που ανήκουν σε μια αραιή περιοχή ή να μην αναγνωρίσουν μια ανωμαλία που βρίσκεται κοντά σε μια πυκνή ομάδα δεδομένων. Ο LOF επιχειρεί να περιορίσει αυτό το πρόβλημα συγκρίνοντας κάθε παρατήρηση με το άμεσο περιβάλλον της [24], [26]. Για τον λόγο αυτό, η μέθοδος είναι ιδιαίτερα χρήσιμη όταν η έννοια της φυσιολογικής συμπεριφοράς δεν είναι ομοιόμορφη σε ολόκληρο το σύνολο δεδομένων.

Η συγκεκριμένη ιδιότητα παρουσιάζει ενδιαφέρον και στην ανίχνευση οικονομικής απάτης. Οι νόμιμες οικονομικές συναλλαγές μπορούν να εμφανίζουν διαφορετικές συμπεριφορές ανάλογα με τον χρήστη, το ποσό, τον χρόνο, τον έμπορο ή άλλα χαρακτηριστικά. Μια ύποπτη συναλλαγή δεν είναι επομένως απαραίτητο να είναι ακραία σε σχέση με όλες τις διαθέσιμες συναλλαγές, αλλά μπορεί να εμφανίζει σημαντική απόκλιση από ένα συγκεκριμένο τοπικό πρότυπο. Η εφαρμογή του LOF σε τραπεζικές συναλλαγές επιτρέπει τον εντοπισμό παρατηρήσεων που παρουσιάζουν αισθητά χαμηλότερη τοπική πυκνότητα από παρόμοιες συναλλαγές, χωρίς να απαιτείται η ύπαρξη ετικετών κατά την εφαρμογή της μη επιβλεπόμενης διαδικασίας [27]. Όπως και στις υπόλοιπες μεθόδους ανίχνευσης ανωμαλιών, μια υψηλή βαθμολογία LOF υποδηλώνει ύποπτη απόκλιση και όχι οριστική απόδειξη οικονομικής απάτης. Παρά τα πλεονεκτήματά του, ο LOF εξαρτάται σημαντικά από την επιλογή της παραμέτρου k , η οποία καθορίζει τον αριθμό των γειτονικών παρατηρήσεων που χρησιμοποιούνται για τον υπολογισμό της τοπικής πυκνότητας [24], [28]. Μια πολύ μικρή τιμή του k δημιουργεί στενά ορισμένη γειτονιά και μπορεί να καταστήσει τη βαθμολογία περισσότερο ευαίσθητη σε θόρυβο ή μεμονωμένες παρατηρήσεις. Αντίθετα, μια πολύ μεγάλη τιμή περιλαμβάνει μεγαλύτερο τμήμα του συνόλου δεδομένων και μπορεί να περιορίσει τον τοπικό χαρακτήρα του αλγορίθμου. Επομένως, η κατάλληλη επιλογή του αριθμού των γειτόνων μπορεί να επηρεάσει ουσιαστικά τις ανωμαλίες που τελικά αναδεικνύονται [24], [28].

Επιπλέον, επειδή ο LOF στηρίζεται στην έννοια της απόστασης και στους πλησιέστερους γείτονες, η απόδοσή του μπορεί να επηρεαστεί από την υψηλή διαστατικότητα των δεδομένων. Καθώς αυξάνεται ο αριθμός των χαρακτηριστικών, οι αποστάσεις μεταξύ των παρατηρήσεων τείνουν να γίνονται λιγότερο διακριτές, δυσκολεύοντας τον αξιόπιστο προσδιορισμό της τοπικής γειτονιάς και της αντίστοιχης πυκνότητας [25], [28]. Πειραματική διερεύνηση σε πολυδιάστατα δεδομένα έχει δείξει ότι η συμπεριφορά του LOF μπορεί να μεταβάλλεται αισθητά καθώς αυξάνεται ο αριθμός των μεταβλητών, γεγονός που καθιστά σημαντική την επιλογή κατάλληλων χαρακτηριστικών ή, όπου απαιτείται, την εφαρμογή τεχνικών μείωσης διαστατικότητας [25].

Η ανάγκη υπολογισμού των πλησιέστερων γειτόνων και των τοπικών πυκνοτήτων δημιουργεί επίσης μεγαλύτερες υπολογιστικές απαιτήσεις σε μεγάλα σύνολα δεδομένων. Η αρχική μορφή του LOF απαιτεί επαναλαμβανόμενες αναζητήσεις γειτόνων, γεγονός που μπορεί να περιορίσει την αποδοτικότητά του όταν ο αριθμός των παρατηρήσεων ή των χαρακτηριστικών αυξάνεται σημαντικά [24]. Οι περιορισμοί αυτοί έχουν οδηγήσει στην ανάπτυξη διαφορετικών παραλλαγών του LOF και συναφών μεθόδων, οι οποίες επιχειρούν να βελτιώσουν τον προσδιορισμό της τοπικής πυκνότητας, την επιλογή της γειτονιάς και την αξιοποίηση ευρύτερης πληροφορίας από τη δομή των δεδομένων [26].

Ένας επιπλέον περιορισμός προκύπτει από το γεγονός ότι ο LOF βασίζει την απόφασή του κυρίως στην τοπική πληροφορία. Η σχετική πυκνότητα μιας παρατήρησης ως προς τους γείτονές της μπορεί να μην αποτυπώνει πάντοτε τη συνολική δομή του συνόλου δεδομένων. Σε ορισμένες περιπτώσεις, μια φυσιολογική παρατήρηση σε αραιή περιοχή μπορεί να λάβει αυξημένη βαθμολογία, ενώ μια πραγματική ανωμαλία σε περιοχή σχετικά υψηλής πυκνότητας μπορεί να είναι δυσκολότερο να διακριθεί [26]. Για τον λόγο αυτό έχουν αναπτυχθεί επεκτάσεις που επιχειρούν να συνδυάσουν την τοπική πληροφορία με ευρύτερα χαρακτηριστικά της κατανομής.

Συνολικά, ο LOF αποτελεί μια διαφορετική προσέγγιση από το IF ως προς τον τρόπο με τον οποίο ορίζεται μια ανωμαλία. Το IF βασίζεται στην ευκολία απομόνωσης μιας παρατήρησης μέσω τυχαίων

διαχωρισμών, ενώ ο LOF αξιολογεί τη σχετική τοπική πυκνότητα μιας παρατήρησης σε σύγκριση με τη γειτονιά της [25]. Η διαφορετική αυτή λογική σημαίνει ότι τα δύο μοντέλα μπορούν να ανταποκρίνονται διαφορετικά σε συγκεκριμένες δομές δεδομένων και να αναγνωρίζουν διαφορετικούς τύπους αποκλίσεων. Η συμπληρωματικότητα αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον στο πλαίσιο της παρούσας εργασίας, καθώς δημιουργεί τη βάση για τη μεταγενέστερη συνδυασμένη αξιοποίηση διαφορετικών μοντέλων ανίχνευσης ανωμαλιών.

2.3.3 Histogram-Based Outlier Score

Το HBOS αποτελεί μη επιβλεπόμενη μέθοδο ανίχνευσης ανωμαλιών που χρησιμοποιεί ιστογράμματα για την εκτίμηση της κατανομής των τιμών των επιμέρους χαρακτηριστικών. Σε αντίθεση με το IF, το οποίο βασίζεται στην ευκολία απομόνωσης μιας παρατήρησης, και το LOF, το οποίο εξετάζει τη σχετική πυκνότητα της τοπικής της γειτονιάς, το HBOS αξιολογεί πόσο συχνές ή σπάνιες είναι οι τιμές μιας παρατήρησης ως προς τις μονοδιάστατες κατανομές των χαρακτηριστικών της [29], [30]. Η βασική λογική είναι ότι τιμές οι οποίες εμφανίζονται σε περιοχές χαμηλής συχνότητας είναι πιθανότερο να σχετίζονται με ασυνήθιστες παρατηρήσεις.

Η κατασκευή του μοντέλου πραγματοποιείται ξεχωριστά για κάθε χαρακτηριστικό του συνόλου δεδομένων. Έστω ότι μια παρατήρηση p περιγράφεται από d χαρακτηριστικά. Για κάθε χαρακτηριστικό δημιουργείται ένα μονοδιάστατο ιστόγραμμα, το οποίο διαιρεί το εύρος των παρατηρούμενων τιμών σε επιμέρους διαστήματα, γνωστά ως bins. Ο αριθμός των παρατηρήσεων που εμπίπτουν σε κάθε διάστημα χρησιμοποιείται για την εκτίμηση της συχνότητας ή της πυκνότητας της συγκεκριμένης περιοχής [29], [32]. Με αυτόν τον τρόπο, περιοχές του ιστογράμματος με μεγάλο αριθμό παρατηρήσεων αντιστοιχούν σε συνηθέστερη συμπεριφορά, ενώ περιοχές με μικρή συχνότητα υποδηλώνουν περισσότερο ασυνήθιστες τιμές.

Η διαμόρφωση των ιστογραμμάτων μπορεί να πραγματοποιηθεί με διαφορετικούς τρόπους. Στην περίπτωση των ισομεγέθων διαστημάτων (static bin-width), το εύρος τιμών ενός χαρακτηριστικού χωρίζεται σε k διαστήματα ίδιου πλάτους και η πυκνότητα κάθε διαστήματος υπολογίζεται με βάση τον αριθμό των παρατηρήσεων που περιλαμβάνει. Μια εναλλακτική προσέγγιση είναι η χρήση μεταβλητού πλάτους διαστημάτων (dynamic bin-width). Στην περίπτωση αυτή οι τιμές ταξινομούνται και τα διαστήματα διαμορφώνονται έτσι ώστε να περιέχουν περίπου ίσο αριθμό παρατηρήσεων, με αποτέλεσμα το πλάτος τους να προσαρμόζεται στην κατανομή των δεδομένων [29]. Η επιλογή της μεθόδου διαχωρισμού επηρεάζει επομένως τον τρόπο με τον οποίο αποτυπώνεται η πυκνότητα των επιμέρους χαρακτηριστικών.

Αφού δημιουργηθούν τα ιστογράμματα, κάθε τιμή μιας παρατήρησης αντιστοιχίζεται στο διάστημα στο οποίο ανήκει. Εάν η τιμή βρίσκεται σε διάστημα υψηλής συχνότητας, η αντίστοιχη συνεισφορά στη βαθμολογία ανωμαλίας είναι σχετικά μικρή. Αντίθετα, εάν βρίσκεται σε διάστημα με χαμηλή συχνότητα, η συνεισφορά της στη συνολική βαθμολογία αυξάνεται. Οι επιμέρους βαθμολογίες συνδυάζονται ώστε να προκύψει ένας συνολικός βαθμός ανωμαλίας για την παρατήρηση [29], [30].

Για μια παρατήρηση p , η βασική μαθηματική διατύπωση της βαθμολογίας HBOS μπορεί να εκφραστεί ως [29]:

$$HBOS(p) = \sum_{i=1}^d \log \frac{1}{hist_i(p)} \quad (2.7)$$

όπου d είναι ο αριθμός των χαρακτηριστικών και $hist_i(p)$ αντιπροσωπεύει το ύψος ή την εκτιμώμενη πυκνότητα του διαστήματος του ιστογράμματος στο οποίο βρίσκεται η τιμή της παρατήρησης p για το χαρακτηριστικό i [29].

Η χρήση του αντιστρόφου της πυκνότητας σημαίνει ότι όσο μικρότερη είναι η συχνότητα εμφάνισης μιας τιμής, τόσο μεγαλύτερη είναι η συνεισφορά της στη βαθμολογία ανωμαλίας. Η λογαριθμική μετατροπή επιτρέπει στη συνέχεια τον συνδυασμό των επιμέρους συνεισφορών με άθροιση. Επομένως, μια παρατήρηση που παρουσιάζει συνηθισμένες τιμές στα περισσότερα χαρακτηριστικά αναμένεται να λάβει σχετικά χαμηλή βαθμολογία, ενώ μια παρατήρηση που εμφανίζει μία ή περισσότερες τιμές σε αραιά τμήματα των αντίστοιχων ιστογραμμάτων μπορεί να λάβει υψηλότερο HBOS.

Χαρακτηριστικό της συγκεκριμένης προσέγγισης είναι ότι τα επιμέρους χαρακτηριστικά αναλύονται ξεχωριστά. Δεν απαιτείται υπολογισμός αποστάσεων μεταξύ όλων των παρατηρήσεων ούτε δημιουργία

τοπικών γειτονιών, όπως συμβαίνει στον LOF. Αντίστοιχα, δεν απαιτείται η κατασκευή μεγάλου αριθμού δέντρων όπως στο IF. Η βαθμολογία προκύπτει άμεσα από τη θέση των τιμών μιας παρατήρησης στα ήδη κατασκευασμένα ιστογράμματα [29], [30]. Η σχεδίαση αυτή συμβάλλει στην απλότητα του μοντέλου και στην περιορισμένη υπολογιστική επιβάρυνσή του.

Η υπολογιστική αποδοτικότητα αποτελεί ένα από τα σημαντικότερα πλεονεκτήματα του HBOS. Η κατασκευή μονοδιάστατων ιστογραμμάτων και η αντιστοίχιση των τιμών στα αντίστοιχα διαστήματα είναι σημαντικά απλούστερες διαδικασίες από επαναλαμβανόμενους υπολογισμούς αποστάσεων ή την εκπαίδευση σύνθετων μοντέλων. Για τον λόγο αυτό, το HBOS έχει εξεταστεί ως σχετικά ελαφριά μη επιβλεπόμενη προσέγγιση για περιβάλλοντα όπου ο όγκος των δεδομένων είναι μεγάλος και η υπολογιστική αποδοτικότητα αποτελεί σημαντικό κριτήριο [32], [33].

Επιπλέον, η μέθοδος παρουσιάζει υψηλό βαθμό ερμηνευσιμότητας. Εφόσον κάθε χαρακτηριστικό αντιστοιχεί σε ένα συγκεκριμένο ιστόγραμμα, είναι δυνατό να εξεταστεί εάν μια υψηλή βαθμολογία ανωμαλίας συνδέεται με την παρουσία τιμών σε περιοχές μικρής συχνότητας. Η απλότητα και η άμεση σχέση μεταξύ της κατανομής ενός χαρακτηριστικού και της αντίστοιχης βαθμολογίας καθιστούν το HBOS ευκολότερο στην κατανόηση σε σύγκριση με περισσότερα σύνθετες μη επιβλεπόμενες προσεγγίσεις [30].

Η συγκεκριμένη ιδιότητα είναι σημαντική σε περιβάλλοντα οικονομικών συναλλαγών, όπου μια ασυνήθιστη συναλλαγή μπορεί να παρουσιάζει χαρακτηριστικές τιμές που εμφανίζονται σπάνια στο ιστορικό σύνολο δεδομένων. Για παράδειγμα, μια ασυνήθιστη τιμή σε κάποιο χαρακτηριστικό μιας συναλλαγής μπορεί να τοποθετηθεί σε περιοχή χαμηλής συχνότητας του αντίστοιχου ιστογράμματος και επομένως να αυξήσει τη συνολική βαθμολογία ανωμαλίας. Το HBOS έχει συμπεριληφθεί σε συστήματα ανίχνευσης ασυνήθιστης συμπεριφοράς που επεξεργάζονται οικονομικές συναλλαγές μαζί με άλλες τεχνικές μηχανικής μάθησης [33].

Όπως και στα υπόλοιπα μη επιβλεπόμενα μοντέλα, μια υψηλή τιμή HBOS δεν αποτελεί από μόνη της απόδειξη ότι μια συναλλαγή είναι δόλια. Υποδηλώνει ότι η συγκεκριμένη παρατήρηση περιλαμβάνει τιμές που εμφανίζονται με σχετικά μικρή συχνότητα σε σχέση με τις κατανομές που έχουν αποτυπωθεί από τα ιστογράμματα. Συνεπώς, η βαθμολογία μπορεί να χρησιμοποιηθεί για την ιεράρχηση συναλλαγών ή λογιστικών εγγραφών που ενδέχεται να απαιτούν περαιτέρω διερεύνηση.

Παρά την απλότητά του, η απόδοση του HBOS εξαρτάται σημαντικά από τον τρόπο κατασκευής των ιστογραμμάτων και ειδικότερα από την επιλογή του αριθμού και του πλάτους των bins. Η χρήση πολύ μικρού αριθμού διαστημάτων μπορεί να οδηγήσει σε υπερβολικά γενική αναπαράσταση της κατανομής, κατά την οποία διαφορετικές τιμές συγκεντρώνονται στην ίδια περιοχή και πιθανές ανωμαλίες δεν διακρίνονται επαρκώς. Αντίθετα, υπερβολικά μεγάλος αριθμός διαστημάτων μπορεί να δημιουργήσει πολύ αραιά ιστογράμματα και να οδηγήσει σε χαρακτηρισμό φυσιολογικών παρατηρήσεων ως ασυνήθιστων [29]. Επομένως, η επιλογή των διαστημάτων αποτελεί ουσιαστική παράμετρο της συμπεριφοράς του μοντέλου.

Η σημασία της συγκεκριμένης παραμέτρου έχει οδηγήσει και στην εξέταση περισσότερων συστηματικών τεχνικών επιλογής του αριθμού των bins. Σε ευρείας κλίμακας συγκριτική αξιολόγηση μη επιβλεπόμενων μοντέλων, ο αριθμός των διαστημάτων του HBOS καθορίστηκε μέσω ειδικού κριτηρίου επιλογής μεγέθους ιστογράμματος, γεγονός που υπογραμμίζει ότι η διακριτοποίηση της κατανομής αποτελεί σημαντικό τμήμα της εφαρμογής του αλγορίθμου [31].

Ένας ακόμη σημαντικός περιορισμός προκύπτει από το γεγονός ότι το HBOS δημιουργεί ξεχωριστή μονοδιάστατη αναπαράσταση για κάθε χαρακτηριστικό. Η προσέγγιση αυτή είναι ιδιαίτερα αποτελεσματική όταν η ανωμαλία είναι εμφανής στις επιμέρους κατανομές, αλλά δεν μοντελοποιεί άμεσα τις σύνθετες σχέσεις μεταξύ διαφορετικών χαρακτηριστικών. Έτσι, μια παρατήρηση μπορεί να περιλαμβάνει τιμές που ξεχωριστά εμφανίζονται συχνά, αλλά ο συγκεκριμένος συνδυασμός των τιμών τους να είναι ασυνήθιστος. Σε μια τέτοια περίπτωση, η πληροφορία που περιέχεται αποκλειστικά στα μονοδιάστατα ιστογράμματα μπορεί να μην επαρκεί για την πλήρη αποτύπωση της ανωμαλίας. Η ιδιότητα αυτή αποτελεί άμεση συνέπεια του τρόπου με τον οποίο το HBOS εκτιμά ξεχωριστά την πυκνότητα κάθε διάστασης [29], [30].

Αντίστοιχος περιορισμός εμφανίζεται στην ανίχνευση τοπικών ανωμαλιών. Επειδή το HBOS αξιολογεί τη σπανιότητα των τιμών με βάση τις συνολικές μονοδιάστατες κατανομές, είναι περισσότερο προσανατολισμένο σε ανωμαλίες που εμφανίζονται σε περιοχές συνολικά χαμηλής συχνότητας. Αντιθέτως, μια παρατήρηση που φαίνεται συνηθισμένη σε συνολικό επίπεδο αλλά είναι ασυνήθιστη σε σχέση με μια συγκεκριμένη τοπική ομάδα μπορεί να είναι δυσκολότερο να αναγνωριστεί [29], [30],

[33]. Η αδυναμία αυτή διαφοροποιεί σημαντικά το HBOS από τον LOF, ο οποίος κατασκευάστηκε ειδικά γύρω από την έννοια της τοπικής απόκλισης.

Οι συγκριτικές αξιολογήσεις υποδηλώνουν επίσης ότι η αποτελεσματικότητα του HBOS δεν είναι σταθερή σε όλα τα σύνολα δεδομένων. Σε ευρεία αξιολόγηση πολλών μη επιβλεπόμενων αλγορίθμων σε πραγματικά πολυμεταβλητά σύνολα δεδομένων παρατηρήθηκαν σημαντικές διαφοροποιήσεις μεταξύ προβλημάτων τοπικών και καθολικών ανωμαλιών [31]. Αντίστοιχα, σε εφαρμογή που περιλάμβανε οικονομικές συναλλαγές, η απόδοση του HBOS διέφερε από εκείνη άλλων αλγορίθμων, γεγονός που καταδεικνύει ότι η καταλληλότητα ενός μοντέλου εξαρτάται από τη δομή και τις ιδιότητες του εκάστοτε συνόλου δεδομένων [33].

Παράλληλα, το HBOS έχει παρουσιάσει ανταγωνιστική συμπεριφορά σε άλλου τύπου εφαρμογές. Σε πλαίσιο μη επιβλεπόμενης ανίχνευσης ανωμαλιών σε βιομηχανικά δεδομένα συμπεριλήφθηκε μεταξύ των βασικών μοντέλων που εξετάστηκαν και παρουσίασε ισχυρή συμπεριφορά σε μέρος των εξεταζόμενων συνόλων δεδομένων [32]. Το αποτέλεσμα αυτό ενισχύει την άποψη ότι η απλότητα της μεθόδου δεν συνεπάγεται απαραίτητα χαμηλή αποτελεσματικότητα, αλλά η καταλληλότητά της εξαρτάται από τον τύπο των ανωμαλιών και τη δομή των διαθέσιμων χαρακτηριστικών.

Συνολικά, το HBOS προσφέρει μια απλή και υπολογιστικά αποδοτική προσέγγιση μη επιβλεπόμενης ανίχνευσης ανωμαλιών, στην οποία η ασυνήθιστη συμπεριφορά εκτιμάται μέσω της σπανιότητας των τιμών των επιμέρους χαρακτηριστικών. Η χρήση ιστογραμμάτων επιτρέπει την άμεση μοντελοποίηση των μονοδιάστατων κατανομών χωρίς την ανάγκη υπολογισμού γειτονιών ή κατασκευής σύνθετων μοντέλων [29], [30]. Ωστόσο, η ευαισθησία στην επιλογή των bins και η περιορισμένη δυνατότητα αποτύπωσης τοπικών ή σύνθετων πολυμεταβλητών σχέσεων πρέπει να λαμβάνονται υπόψη κατά την εφαρμογή του. Η διαφορετική αρχή λειτουργίας του σε σχέση με το IF και το LOF αποτελεί ταυτόχρονα σημαντικό στοιχείο για τη μεταγενέστερη διερεύνηση συνδυαστικών προσεγγίσεων, καθώς τα τρία μοντέλα αξιολογούν την ανωμαλία χρησιμοποιώντας διαφορετικές ιδιότητες της δομής των δεδομένων.

2.3.4 Autoencoder

Οι Autoencoders (AEs) αποτελούν νευρωνικά δίκτυα τα οποία μπορούν να χρησιμοποιηθούν για μη επιβλεπόμενη εκμάθηση αναπαραστάσεων και ανίχνευση ανωμαλιών. Σε αντίθεση με τα IF, LOF και HBOS, τα οποία βασίζονται αντίστοιχα στην απομόνωση, στην τοπική πυκνότητα και στη συχνότητα εμφάνισης των χαρακτηριστικών, οι AEs επιχειρούν να μάθουν μια συμπίεσμένη αναπαράσταση της δομής των δεδομένων και στη συνέχεια να ανακατασκευάσουν την αρχική είσοδο [34]–[36]. Η απόκλιση μεταξύ της αρχικής παρατήρησης και της ανακατασκευασμένης εκδοχής της μπορεί να αξιοποιηθεί για την εκτίμηση του βαθμού ανωμαλίας.

Η βασική αρχιτεκτονική ενός AE αποτελείται από δύο κύρια τμήματα: τον κωδικοποιητή (encoder) και τον αποκωδικοποιητή (decoder). Ο κωδικοποιητής δέχεται την αρχική παρατήρηση και τη μετασχηματίζει σε μια εσωτερική αναπαράσταση, η οποία συνήθως έχει μικρότερη διάσταση από τον αρχικό χώρο χαρακτηριστικών. Η ενδιάμεση αυτή αναπαράσταση αναφέρεται ως λανθάνων χώρος (latent space) ή κωδικοποιημένη αναπαράσταση. Στη συνέχεια, ο αποκωδικοποιητής χρησιμοποιεί την πληροφορία του λανθάνοντος χώρου για να δημιουργήσει μια ανακατασκευή της αρχικής εισόδου [34], [35], [37].

Έστω μια παρατήρηση $x \in R^d$. Η διαδικασία κωδικοποίησης μπορεί να εκφραστεί ως:

$$z = f_{\theta}(x) \quad (2.8)$$

όπου f_{θ} είναι η συνάρτηση του κωδικοποιητή με παραμέτρους θ και z η λανθάνουσα αναπαράσταση της αρχικής παρατήρησης. Στη συνέχεια, ο αποκωδικοποιητής επιχειρεί να επαναφέρει την πληροφορία στον αρχικό χώρο:

$$\hat{x} = g_{\phi}(z) = g_{\phi}(f_{\theta}(x)) \quad (2.9)$$

όπου g_{ϕ} είναι η συνάρτηση του αποκωδικοποιητή με παραμέτρους ϕ και $\hat{x}$ η ανακατασκευασμένη παρατήρηση [34], [36], [37].

Η ύπαρξη του λανθάνοντος χώρου δημιουργεί ένα είδος στενώματος πληροφορίας (bottleneck). Όταν η διάσταση του z είναι μικρότερη από τη διάσταση της αρχικής εισόδου, το δίκτυο δεν μπορεί απλώς να αντιγράψει κάθε χαρακτηριστικό ανεξάρτητα. Αντίθετα, καλείται να διατηρήσει εκείνη την πληροφορία που είναι περισσότερο χρήσιμη για την ανακατασκευή των παρατηρήσεων. Με αυτόν τον τρόπο μπορεί να μάθει συμπιεσμένες αναπαραστάσεις που αποτυπώνουν σχέσεις και πρότυπα τα οποία υπάρχουν στα δεδομένα [34], [36]. Η δυνατότητα αυτή διαφοροποιεί τους Autoencoders από τις προηγούμενες μεθόδους, καθώς η αναπαράσταση δεν καθορίζεται από μια προκαθορισμένη έννοια απόστασης, πυκνότητας ή μονοδιάστατης συχνότητας.

Η εκπαίδευση του Autoencoder πραγματοποιείται με στόχο την ελαχιστοποίηση της διαφοράς μεταξύ της πραγματικής εισόδου και της ανακατασκευασμένης εξόδου. Μια συνηθισμένη συνάρτηση απώλειας για αριθμητικά δεδομένα είναι το Μέσο Τετραγωνικό Σφάλμα (MSE). Για ένα σύνολο N παρατηρήσεων μπορεί να εκφραστεί ως [34], [36], [37]:

$$L(\theta, \varphi) = \frac{1}{N} \sum_{i=1}^N \|x_i - \hat{x}_i\|^2 \quad (2.10)$$

όπου x_i είναι η αρχική παρατήρηση και $\hat{x}_i = g_\varphi(f_\theta(x_i))$ η αντίστοιχη ανακατασκευή της. Κατά την εκπαίδευση προσαρμόζονται οι παράμετροι του κωδικοποιητή και του αποκωδικοποιητή ώστε να ελαχιστοποιηθεί η παραπάνω συνάρτηση απώλειας [34], [36].

Η αξιοποίηση του Autoencoder για ανίχνευση ανωμαλιών στηρίζεται στην υπόθεση ότι το μοντέλο μαθαίνει κυρίως τα επαναλαμβανόμενα και κυρίαρχα πρότυπα του συνόλου εκπαίδευσης. Παρατηρήσεις που ακολουθούν αυτά τα πρότυπα αναμένεται να μπορούν να ανακατασκευαστούν με σχετικά μικρό σφάλμα. Αντίθετα, όταν μια νέα παρατήρηση διαφέρει σημαντικά από τη δομή που έχει μάθει το δίκτυο, η ανακατασκευή της μπορεί να είναι λιγότερο ακριβής, δημιουργώντας μεγαλύτερη απόκλιση μεταξύ x και $\hat{x}$ [34]–[36].

Στην απλούστερη μορφή της προσέγγισης, το σφάλμα ανακατασκευής (reconstruction error) χρησιμοποιείται επομένως ως βαθμολογία ανωμαλίας. Για μία παρατήρηση x μπορεί να οριστεί ως:

$$s(x) = \frac{1}{d} \sum_{j=1}^d (x_j - \hat{x}_j)^2 \quad (2.11)$$

όπου d είναι ο αριθμός των χαρακτηριστικών της παρατήρησης. Υψηλότερη τιμή του $s(x)$ υποδηλώνει μεγαλύτερη διαφορά μεταξύ της αρχικής και της ανακατασκευασμένης παρατήρησης και, συνεπώς, μεγαλύτερο βαθμό απόκλισης από τα πρότυπα που έχει μάθει το μοντέλο [35], [36].

Για τη μετατροπή της συνεχούς βαθμολογίας σε απόφαση ανωμαλίας χρησιμοποιείται συνήθως ένα κατώφλι (threshold) τ . Η λογική μπορεί να εκφραστεί ως:

$$x = \begin{cases} \text{ανωμαλία} & s(x) > \tau \\ \text{φυσιολογική παρατήρηση} & s(x) \leq \tau \end{cases} \quad (2.12)$$

Με αυτόν τον τρόπο, οι παρατηρήσεις των οποίων το σφάλμα ανακατασκευής υπερβαίνει την καθορισμένη τιμή εξετάζονται ως πιθανές ανωμαλίες [34], [35]. Η επιλογή του κατάλληλου κατωφλίου είναι ιδιαίτερα σημαντική, επειδή μια υπερβολικά χαμηλή τιμή μπορεί να αυξήσει τον αριθμό των φυσιολογικών παρατηρήσεων που επισημαίνονται ως ανώμαλες, ενώ μια υπερβολικά υψηλή τιμή μπορεί να οδηγήσει σε μη εντοπισμό πραγματικών αποκλίσεων [35].

Η μορφή της εκπαίδευσης μπορεί να διαφέρει ανάλογα με το διαθέσιμο σύνολο δεδομένων. Σε περιπτώσεις όπου υπάρχουν διαθέσιμες παρατηρήσεις που είναι γνωστό ότι αντιπροσωπεύουν φυσιολογική συμπεριφορά, ο Autoencoder μπορεί να εκπαιδευτεί κυρίως ή αποκλειστικά σε αυτές, ώστε να μάθει μια αναπαράσταση του κανονικού προτύπου [34], [35]. Σε αυστηρά μη επιβλεπόμενο περιβάλλον, αντίθετα, οι ετικέτες δεν χρησιμοποιούνται κατά την εκπαίδευση και το σύνολο μπορεί να περιλαμβάνει τόσο φυσιολογικές όσο και ανώμαλες παρατηρήσεις. Στην περίπτωση αυτή βασική παραδοχή είναι συνήθως ότι οι φυσιολογικές παρατηρήσεις κυριαρχούν αριθμητικά και επομένως επηρεάζουν περισσότερο τη δομή που μαθαίνει το δίκτυο [36].

Ένα σημαντικό πλεονέκτημα των AEs είναι η δυνατότητά τους να μαθαίνουν μη γραμμικές σχέσεις μεταξύ των χαρακτηριστικών. Τα διαδοχικά επίπεδα και οι μη γραμμικές συναρτήσεις ενεργοποίησης επιτρέπουν την κατασκευή περισσότερο σύνθετων εσωτερικών αναπαραστάσεων σε σχέση με γραμμικές τεχνικές μείωσης διαστατικότητας [35], [36]. Αυτό καθιστά τους AEs ιδιαίτερα ενδιαφέροντες σε πολυδιάστατα δεδομένα, όπου μια ανωμαλία μπορεί να μην οφείλεται σε μία ακραία τιμή, αλλά σε έναν ασυνήθιστο συνδυασμό πολλών χαρακτηριστικών.

Η ιδιότητα αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον στην ανίχνευση οικονομικής απάτης, καθώς οι οικονομικές συναλλαγές μπορούν να περιγράφονται από μεγάλο αριθμό αλληλεξαρτώμενων χαρακτηριστικών. Η χρήση AEs επιτρέπει την εκμάθηση συμπίεσμένων αναπαραστάσεων της συναλλακτικής συμπεριφοράς και την αναζήτηση περιπτώσεων που δεν ανακατασκευάζονται επαρκώς από το μοντέλο [10], [37]. Οι AEs συγκαταλέγονται στις μη επιβλεπόμενες τεχνικές που έχουν χρησιμοποιηθεί σε εφαρμογές ανίχνευσης οικονομικής απάτης, ενώ έχουν εξεταστεί ειδικότερα και σε δεδομένα πιστωτικών καρτών [10], [37].

Σε ένα τέτοιο πλαίσιο, μια νόμιμη συναλλαγή που ακολουθεί συνηθισμένα πρότυπα αναμένεται να μπορεί να αναπαρασταθεί και να ανακατασκευαστεί με σχετικά μικρή απόκλιση. Αντίθετα, μια συναλλαγή που παρουσιάζει ασυνήθιστο συνδυασμό χαρακτηριστικών μπορεί να οδηγήσει σε υψηλότερο σφάλμα ανακατασκευής και να λάβει μεγαλύτερη βαθμολογία ανωμαλίας. Όπως και στα υπόλοιπα μοντέλα της παρούσας εργασίας, όμως, η υψηλή βαθμολογία δεν αποτελεί από μόνη της απόδειξη απάτης. Αποτελεί ένδειξη ότι η συναλλαγή αποκλίνει από τη δομή που έχει μάθει το μοντέλο και επομένως μπορεί να απαιτεί περαιτέρω διερεύνηση.

Οι AEs προσφέρουν επίσης ευελιξία ως προς την αρχιτεκτονική τους. Ανάλογα με τη φύση των δεδομένων μπορούν να χρησιμοποιηθούν διαφορετικές παραλλαγές, όπως CAE, DAE και VAE [34], [37]. Οι παραλλαγές αυτές μεταβάλλουν τον τρόπο με τον οποίο πραγματοποιείται η κωδικοποίηση, η ανακατασκευή ή η κανονικοποίηση του λανθάνοντος χώρου. Ωστόσο, η βασική αρχή της κλασικής προσέγγισης παραμένει η ίδια: εκμάθηση κατάλληλης αναπαραστάσης και αξιοποίηση της απόκλισης από αυτή για την αναγνώριση ασυνήθιστων παρατηρήσεων.

Παρά τα πλεονεκτήματά τους, η χρήση AEs για ανίχνευση ανωμαλιών παρουσιάζει σημαντικούς περιορισμούς. Ο βασικότερος αφορά την ίδια την υπόθεση του σφάλματος ανακατασκευής. Ένα νευρωνικό δίκτυο με αρκετή αναπαραστατική ικανότητα μπορεί σε ορισμένες περιπτώσεις να μάθει να ανακατασκευάζει ικανοποιητικά όχι μόνο τις φυσιολογικές παρατηρήσεις αλλά και τις ανωμαλίες. Ως αποτέλεσμα, η διαφορά μεταξύ των σφαλμάτων ανακατασκευής των δύο κατηγοριών μπορεί να μειωθεί, δυσκολεύοντας τον σαφή διαχωρισμό τους [36].

Το πρόβλημα μπορεί να γίνει εντονότερο σε πλήρως μη επιβλεπόμενη εκπαίδευση όταν το σύνολο δεδομένων είναι μολυσμένο από ανωμαλίες (anomaly-contaminated). Εφόσον οι ανώμαλες παρατηρήσεις συμμετέχουν και αυτές στη διαδικασία βελτιστοποίησης, το δίκτυο μπορεί σταδιακά να μάθει μέρος της δομής τους και να μειώσει το σφάλμα ανακατασκευής τους. Συνεπώς, η ελαχιστοποίηση της συνολικής απώλειας ανακατασκευής δεν συνεπάγεται απαραίτητα αντίστοιχη βελτίωση της δυνατότητας διάκρισης μεταξύ φυσιολογικών και ανώμαλων παρατηρήσεων [36]. Ο περιορισμός αυτός έχει οδηγήσει στην ανάπτυξη τροποποιημένων αρχιτεκτονικών και συναρτήσεων απώλειας που επιχειρούν να δημιουργήσουν περισσότερο διακριτικές αναπαραστάσεις στον λανθάνων χώρο.

Ένας επιπλέον παράγοντας είναι η επιλογή του κατωφλίου ανωμαλίας. Το σφάλμα ανακατασκευής αποτελεί συνεχή βαθμολογία και επομένως πρέπει να καθοριστεί το σημείο πάνω από το οποίο μια παρατήρηση θα θεωρείται ύποπτη. Η συμπεριφορά του μοντέλου μπορεί να αλλάξει σημαντικά ανάλογα με αυτή την επιλογή, ενώ δεν υπάρχει μία ενιαία τιμή που να είναι κατάλληλη για κάθε σύνολο δεδομένων [35]. Αυτό καθιστά την επιλογή της μεθόδου διαχωρισμού κατωφλίου σημαντικό μέρος της πρακτικής εφαρμογής του AE και όχι εγγενώς λυμένο στοιχείο της αρχιτεκτονικής του.

Παράλληλα, η απόδοση του μοντέλου επηρεάζεται από επιλογές όπως η διάσταση του λανθάνοντα χώρου, ο αριθμός και το μέγεθος των κρυφών επιπέδων, οι συναρτήσεις ενεργοποίησης και η διαδικασία εκπαίδευσης. Μια υπερβολικά περιορισμένη αναπαραστάση μπορεί να μην καταφέρει να αποτυπώσει επαρκώς τη φυσιολογική δομή των δεδομένων, ενώ ένα υπερβολικά ισχυρό δίκτυο μπορεί να ανακατασκευάζει με μεγάλη ακρίβεια ακόμη και παρατηρήσεις που αποκλίνουν από αυτή [34], [36]. Επομένως, η αρχιτεκτονική του AE πρέπει να εξισορροπεί την ικανότητα αναπαραστάσης με τη διατήρηση της δυνατότητας διάκρισης των ανωμαλιών.

Συνολικά, οι AEs προσεγγίζουν την ανίχνευση ανωμαλιών με διαφορετικό τρόπο από τα προηγούμενα μοντέλα της παρούσας ενότητας. Το IF χρησιμοποιεί την ευκολία απομόνωσης, το LOF τη σχετική τοπική πυκνότητα, το HBOS τη σπανιότητα των επιμέρους χαρακτηριστικών, ενώ ο AE αξιοποιεί την ικανότητα ανακατασκευής μιας παρατήρησης μέσω μιας μαθημένης εσωτερικής αναπαράστασης. Η δυνατότητα εκμάθησης μη γραμμικών και πολυμεταβλητών σχέσεων αποτελεί σημαντικό πλεονέκτημα, ενώ η ευαισθησία στο σφάλμα ανακατασκευής, στο κατώφλι και στην παρουσία ανωμαλιών κατά την εκπαίδευση αποτελούν βασικούς περιορισμούς [34]–[36]. Η διαφορετική αρχή λειτουργίας του ενισχύει παράλληλα τη δυνατότητα συμπληρωματικής αξιοποίησής του με τα υπόλοιπα μοντέλα στο πλαίσιο τεχνικών συνδυασμού.

2.3.5 Τεχνικές Συνδυασμού

Οι τεχνικές συνδυασμού (Ensemble Techniques) βασίζονται στην αξιοποίηση περισσότερων του ενός μοντέλων για την παραγωγή μιας ενιαίας τελικής εκτίμησης. Η βασική λογική είναι ότι διαφορετικοί αλγόριθμοι μπορούν να εξετάζουν διαφορετικές ιδιότητες των δεδομένων και, κατά συνέπεια, να μην εντοπίζουν πάντοτε τις ίδιες περιπτώσεις ως ανώμαλες. Η συνδυασμένη αξιοποίησή τους μπορεί επομένως να μειώσει την εξάρτηση της τελικής απόφασης από τις ιδιαίτερες παραδοχές και τους περιορισμούς ενός μεμονωμένου μοντέλου [38]–[40].

Η συγκεκριμένη ιδιότητα είναι ιδιαίτερα σημαντική στην ανίχνευση ανωμαλιών. Οι διαφορετικές οικογένειες αλγορίθμων χρησιμοποιούν διαφορετικά κριτήρια για να προσδιορίσουν κατά πόσο μια παρατήρηση αποκλίνει από τη συνηθισμένη συμπεριφορά. Μια μέθοδος μπορεί να βασίζεται στην ευκολία απομόνωσης της παρατήρησης, μια δεύτερη στην τοπική πυκνότητα, μια τρίτη στη συχνότητα εμφάνισης των χαρακτηριστικών και μια τέταρτη στην ικανότητα ανακατασκευής της παρατήρησης. Ως αποτέλεσμα, διαφορετικοί ανιχνευτές μπορούν να εντοπίζουν διαφορετικά υποσύνολα ανωμαλιών [40], [43]. Η ύπαρξη περιορισμένης επικάλυψης μεταξύ των προβλέψεων διαφορετικών ανιχνευτών υποστηρίζει την αξιοποίηση συμπληρωματικών μηχανισμών στο πλαίσιο ενός συνδυασμού [43].

Στα μη επιβλεπόμενα περιβάλλοντα, η δημιουργία ενός υβριδικού μοντέλου παρουσιάζει πρόσθετες δυσκολίες λόγω της απουσίας αξιόπιστων ετικετών κατά τη διαδικασία εκπαίδευσης και ανίχνευσης. Δεν είναι επομένως πάντοτε δυνατό να εκτιμηθεί εκ των προτέρων ποιος ανιχνευτής είναι περισσότερο αξιόπιστος για κάθε είδος ανωμαλίας. Οι τεχνικές συνδυασμού μπορούν να λειτουργήσουν είτε στο επίπεδο των δυαδικών αποφάσεων των επιμέρους μοντέλων είτε στο επίπεδο των συνεχών βαθμολογιών ανωμαλίας που αυτά παράγουν [40], [41].

Η διάκριση αυτή οδηγεί σε δύο βασικές κατηγορίες. Στην πρώτη περίπτωση, κάθε ανιχνευτής αποφασίζει εάν μια παρατήρηση είναι φυσιολογική ή ανώμαλη και οι επιμέρους αποφάσεις συνδυάζονται μέσω κάποιου κανόνα ψηφοφορίας. Στη δεύτερη, διατηρείται η πληροφορία που περιέχεται στις βαθμολογίες ανωμαλίας και οι βαθμολογίες συνδυάζονται ώστε να δημιουργηθεί ένας νέος συνολικός δείκτης [40], [41].

Η εφαρμογή υβριδικών προσεγγίσεων παρουσιάζει ιδιαίτερο ενδιαφέρον και στην ανίχνευση οικονομικής απάτης, όπου η πολυπλοκότητα των συναλλαγών και η ετερογένεια των πιθανών μορφών απάτης μπορούν να περιορίσουν την αποτελεσματικότητα ενός μοναδικού μοντέλου. Συνδυαστικές προσεγγίσεις έχουν εφαρμοστεί σε δεδομένα πιστωτικών καρτών και οικονομικών συναλλαγών με τη χρήση ψηφοφορίας, σταθμισμένης ψηφοφορίας και άλλων μηχανισμών συγκέντρωσης των αποτελεσμάτων διαφορετικών μοντέλων [38], [39], [42].

Στην παρούσα εργασία εξετάζονται τρεις τεχνικές συνδυασμού: OR/Union, πλειοψηφικής ψηφοφορίας και σταθμισμένης ψηφοφορίας. Οι δύο πρώτες χρησιμοποιούν τις τελικές δυαδικές αποφάσεις των επιμέρους ανιχνευτών, ενώ η τρίτη αξιοποιεί τις συνεχείς βαθμολογίες ανωμαλίας και επιτρέπει διαφορετικό βαθμό συνεισφοράς κάθε μοντέλου.

Συνδυασμός OR/Union

Το OR/Union αποτελεί τεχνική συνδυασμού κατά την οποία μια παρατήρηση χαρακτηρίζεται ως ανωμαλία όταν τουλάχιστον ένας από τους συμμετέχοντες ανιχνευτές την έχει επισημάνει ως τέτοια [43]. Σε αντίθεση με περισσότερους αυστηρούς κανόνες ψηφοφορίας, δεν απαιτείται συμφωνία μεταξύ πολλών μοντέλων. Αρκεί η ύπαρξη μίας θετικής ένδειξης ώστε η παρατήρηση να συμπεριληφθεί στο

τελικό σύνολο των πιθανών ανωμαλιών. Η συγκεκριμένη λογική έχει οριστεί ως ψηφοφορία OR σε συστήματα τεχνικών συνδυασμού πολλαπλών ανιχνευτών [43].

Έστω ότι χρησιμοποιούνται M μοντέλα και ότι $y_m(x) \in \{0,1\}$ είναι η δυαδική απόφαση του ανιχνευτή m για την παρατήρηση x , όπου 1 δηλώνει ανωμαλία και 0 φυσιολογική παρατήρηση. Η τελική απόφαση του τεχνικού συνδυασμού OR μπορεί να εκφραστεί ως:

$$y_{OR}(x) = \begin{cases} 1, \sum_{m=1}^M y_m(x) \geq 1 \\ 0, \sum_{m=1}^M y_m(x) < 1 \end{cases} \quad (2.13)$$

Η ίδια λογική μπορεί να εκφραστεί μέσω της θεωρίας συνόλων. Εάν A_m είναι το σύνολο των παρατηρήσεων που χαρακτηρίζονται ως ανωμαλίες από τον ανιχνευτή m , τότε:

$$A_{OR} = \cup_{m=1}^M A_m \quad (2.14)$$

Το τελικό σύνολο περιέχει επομένως όλες τις παρατηρήσεις που εντοπίστηκαν από τουλάχιστον ένα από τα συμμετέχοντα μοντέλα.

Το κύριο χαρακτηριστικό του OR/Union είναι ότι δίνει έμφαση στην κάλυψη των πιθανών ανωμαλιών. Εφόσον δεν απαιτείται συμφωνία μεταξύ των ανιχνευτών, διατηρούνται ακόμη και περιπτώσεις που αναγνωρίζονται αποκλειστικά από ένα μοντέλο. Η συγκεκριμένη συμπεριφορά μπορεί να είναι χρήσιμη όταν οι ανιχνευτές είναι ετερογενείς και έχουν σχεδιαστεί ώστε να εντοπίζουν διαφορετικές μορφές απόκλισης. Η ύπαρξη παρατηρήσεων που εντοπίζονται μόνο από έναν από πολλούς ανιχνευτές έχει παρατηρηθεί και σε εφαρμογές συνδυαστικής ανίχνευσης, γεγονός που αναδεικνύει τη διαφορετικότητα των επιμέρους μηχανισμών [43].

Η αυξημένη κάλυψη συνοδεύεται όμως από έναν σημαντικό συμβιβασμό. Επειδή αρκεί μία θετική απόφαση, ένας ανιχνευτής που χαρακτηρίζει μεγάλο αριθμό παρατηρήσεων ως ανώμαλες μπορεί να αυξήσει σημαντικά το μέγεθος του τελικού συνόλου. Συνεπώς, η συγκεκριμένη τεχνική τείνει να ευνοεί την ευαισθησία της ανίχνευσης, αλλά μπορεί ταυτόχρονα να αυξήσει τις περιπτώσεις που απαιτούν περαιτέρω διερεύνηση. Η συμπεριφορά αυτή αποτυπώνεται στη διάκριση μεταξύ Recall και Precision, καθώς η ψηφοφορία OR μπορεί να αποτελεί κατάλληλη επιλογή όταν δίνεται μεγαλύτερη έμφαση στον περιορισμό των μη εντοπισμένων θετικών περιπτώσεων [43].

Στην περίπτωση της παρούσας εργασίας, η τεχνική μπορεί να εκφραστεί ως:

$$A_{OR} = A_{IF} \cup A_{LOF} \cup A_{HBOS} \cup A_{AE} \quad (2.15)$$

Έτσι, μια συναλλαγή συμπεριλαμβάνεται στο τελικό σύνολο πιθανών ανωμαλιών όταν έχει χαρακτηριστεί ως ανώμαλη από τουλάχιστον ένα από τα τέσσερα μοντέλα. Η τελική ένδειξη δεν ισοδυναμεί με επιβεβαιωμένη οικονομική απάτη, αλλά με επιλογή της συναλλαγής για περαιτέρω αξιολόγηση.

Συνδυασμός Πλειοψηφικής Ψηφοφορίας

Ο συνδυασμός πλειοψηφικής ψηφοφορίας (Majority Voting Ensemble) χρησιμοποιεί επίσης τις δυαδικές αποφάσεις των επιμέρους μοντέλων, αλλά εφαρμόζει αυστηρότερο κριτήριο από το OR/Union. Μια παρατήρηση χαρακτηρίζεται ως ανωμαλία όταν η πλειοψηφία των ανιχνευτών συμφωνήσει με τον συγκεκριμένο χαρακτηρισμό [42], [43]. Η ψηφοφορία λειτουργεί επομένως ως μηχανισμός συναίνεσης μεταξύ διαφορετικών μοντέλων.

Για M ανιχνευτές, ο κανόνας μπορεί να εκφραστεί ως:

$$y_{MV}(x) = \begin{cases} 1, \sum_{m=1}^M y_m(x) > \frac{M}{2} \\ 0, \sum_{m=1}^M y_m(x) \leq \frac{M}{2} \end{cases} \quad (2.16)$$

Για παράδειγμα, όταν το μοντέλο τεχνικού συνδυασμού αποτελείται από τέσσερα μοντέλα, η αυστηρή πλειοψηφία απαιτεί τουλάχιστον τρεις θετικές αποφάσεις για τον χαρακτηρισμό μιας παρατήρησης ως ανωμαλίας.

Η λογική αυτή περιορίζει την επίδραση μιας μεμονωμένης απόφασης. Εάν ένας ανιχνευτής χαρακτηρίσει μια παρατήρηση ως ανώμαλη αλλά τα περισσότερα από τα υπόλοιπα μοντέλα δεν συμφωνήσουν, η παρατήρηση δεν εντάσσεται στο τελικό σύνολο. Με αυτόν τον τρόπο μπορεί να μειωθεί η επίδραση ενός ανιχνευτή που είναι περισσότερο ευαίσθητος σε θόρυβο ή σε συγκεκριμένες δομές των δεδομένων.

Η χρήση πλειοψηφικής ψηφοφορίας έχει εφαρμοστεί και στην οικονομική ανίχνευση απάτης ως μηχανισμός συγκέντρωσης των προβλέψεων διαφορετικών μοντέλων [38], [42]. Παράλληλα, η άμεση σύγκριση διαφορετικών κανόνων ψηφοφορίας δείχνει ότι το επίπεδο συμφωνίας που απαιτείται επηρεάζει τη συμπεριφορά του τελικού υβριδικού μοντέλου [43].

Η πιο συντηρητική φύση της πλειοψηφικής ψηφοφορίας ενδέχεται, ωστόσο, να οδηγήσει στην απόρριψη περιπτώσεων που ανιχνεύονται μόνο από έναν ή δύο εξειδικευμένους ανιχνευτές. Αυτό έχει ιδιαίτερη σημασία όταν χρησιμοποιούνται μοντέλα διαφορετικής αρχής λειτουργίας. Για παράδειγμα, μια τοπική ανωμαλία μπορεί να αναγνωριστεί από το LOF χωρίς να εμφανίζει αντίστοιχα ισχυρή απόκλιση για ένα μοντέλο που στηρίζεται σε διαφορετικό μηχανισμό.

Συνεπώς, το OR/Union και το Majority Voting αντιπροσωπεύουν δύο διαφορετικά επίπεδα συναίνεσης. Το πρώτο απαιτεί:

$$\sum y_m(x) \geq 1 \quad (2.17)$$

ενώ το δεύτερο απαιτεί:

$$\sum y_m(x) > \frac{M}{2} \quad (2.18)$$

Το OR/Union διατηρεί επομένως μεγαλύτερο αριθμό υποψήφιων ανωμαλιών, ενώ το Majority Voting επικεντρώνεται σε περιπτώσεις στις οποίες υπάρχει ισχυρότερη συμφωνία μεταξύ των μοντέλων.

Ένας βασικός περιορισμός και των δύο προσεγγίσεων είναι ότι χρησιμοποιούν μόνο τη δυαδική απόφαση κάθε μοντέλου. Ένας ανιχνευτής που θεωρεί μια παρατήρηση οριακά ανώμαλη και ένας ανιχνευτής που της αποδίδει εξαιρετικά υψηλό βαθμό ανωμαλίας παρέχουν στην τελική ψηφοφορία ακριβώς την ίδια ψήφο. Με αυτόν τον τρόπο χάνεται μέρος της πληροφορίας που περιέχεται στις συνεχείς βαθμολογίες των μοντέλων.

Συνδυασμός Σταθμισμένης Ψηφοφορίας

Οι συνδυασμοί σταθμισμένης ψηφοφορίας (Weighted Score-Based Ensembles) επιχειρούν να διατηρήσουν αυτή την πρόσθετη πληροφορία, χρησιμοποιώντας τις συνεχείς βαθμολογίες ανωμαλίας αντί αποκλειστικά των τελικών δυαδικών αποφάσεων. Κάθε ανιχνευτής παρέχει έναν αριθμητικό βαθμό που εκφράζει πόσο ασυνήθιστη θεωρεί μια παρατήρηση και οι επιμέρους βαθμολογίες συνδυάζονται σε έναν συνολικό δείκτη [40], [41].

Η άμεση πρόσθεση των αρχικών βαθμολογιών δεν είναι γενικά κατάλληλη, επειδή διαφορετικοί ανιχνευτές μπορούν να παράγουν αποτελέσματα σε διαφορετικές αριθμητικές κλίμακες και με διαφορετικές κατανομές. Ένα μοντέλο μπορεί, για παράδειγμα, να παράγει βαθμολογίες μεταξύ 0 και 1, ενώ ένα άλλο να εμφανίζει πολύ διαφορετικό αριθμητικό εύρος. Χωρίς προηγούμενη προσαρμογή, ο ανιχνευτής με τις αριθμητικά μεγαλύτερες τιμές θα μπορούσε να αποκτήσει δυσανάλογη επίδραση στο τελικό αποτέλεσμα [40], [41].

Για τον λόγο αυτό, προηγείται συνήθως διαδικασία κανονικοποίησης ή βαθμονόμησης των βαθμολογιών σε κοινή κλίμακα. Για τον σκοπό αυτό μπορούν να χρησιμοποιηθούν τεχνικές όπως η κανονικοποίηση Min–Max, η κανονικοποίηση βάσει z-score, οι μετασχηματισμοί βάσει κατάταξης και η βαθμονόμηση βάσει εκατοστημορίων. Η τελευταία αποτελεί προσέγγιση μέσω της οποίας οι

ετερογενείς βαθμολογίες των επιμέρους μοντέλων μετατρέπονται σε συγκρίσιμη κλίμακα πριν από τον τελικό συνδυασμό τους [41].

Έστω ότι $\xi_m(x)$ είναι η κανονικοποιημένη βαθμολογία της παρατήρησης x από τον ανιχνευτή m . Μια σταθμισμένη βαθμολογία τεχνικής συνδυασμού μπορεί να οριστεί ως:

$$S_w(x) = \sum_{m=1}^M w_m \xi_m(x) \quad (2.19)$$

όπου:

$$w_m \geq 0 \quad (2.20)$$

και συνήθως:

$$\sum_{m=1}^M w_m = 1 \quad (2.21)$$

Το w_m εκφράζει τη σχετική συνεισφορά του ανιχνευτή m στην τελική βαθμολογία. Μεγαλύτερη τιμή του βάρους αυξάνει την επίδραση του αντίστοιχου μοντέλου. Όταν όλα τα βάρη είναι ίσα,

$$w_m = \frac{1}{M} \quad (2.22)$$

η τελική συνδυασμένη βαθμολογία προκύπτει από τον απλό μέσο όρο των κανονικοποιημένων βαθμολογιών των επιμέρους μοντέλων. Η συγκεκριμένη προσέγγιση αποτελεί μορφή συνδυασμού σε επίπεδο βαθμολογιών, καθώς αξιοποιούνται οι συνεχείς εξοδοί των μοντέλων και όχι μόνο οι τελικές δυαδικές αποφάσεις τους [41].

Η τελική συνεχής βαθμολογία μπορεί να χρησιμοποιηθεί είτε για κατάταξη των παρατηρήσεων είτε να μετατραπεί σε δυαδική απόφαση μέσω ενός κατωφλίου τ :

$$y_w(x) = \begin{cases} 1, & S_w(x) > \tau \\ 0, & S_w(x) \leq \tau \end{cases} \quad (2.23)$$

Η δυνατότητα κατάταξης αποτελεί σημαντικό πλεονέκτημα σε εφαρμογές οικονομικού ελέγχου. Αντί όλες οι επισημασμένες συναλλαγές να αντιμετωπίζονται με τον ίδιο τρόπο, μπορούν να ταξινομηθούν από την περισσότερο προς τη λιγότερο ασυνήθιστη, επιτρέποντας την ιεράρχηση των περιπτώσεων που θα εξεταστούν πρώτες.

Η επιλογή των βαρών αποτελεί ωστόσο κρίσιμο ζήτημα. Σε επιβλεπόμενο περιβάλλον μπορούν να χρησιμοποιηθούν διαθέσιμες ετικέτες για να εκτιμηθεί η αποτελεσματικότητα των επιμέρους μοντέλων και να δοθεί μεγαλύτερη βαρύτητα στα ισχυρότερα. Σε πλήρως μη επιβλεπόμενο περιβάλλον η διαδικασία είναι δυσκολότερη, καθώς η πραγματική ποιότητα κάθε ανιχνευτή δεν είναι άμεσα γνωστή [40]. Για τον λόγο αυτό, η στάθμιση πρέπει να βασίζεται σε σαφώς καθορισμένο κριτήριο και να μην επιλέγεται αυθαίρετα.

Η λογική συνδυασμού βαθμολογιών είναι περισσότερο ευέλικτη από τη δυαδική ψηφοφορία, καθώς λαμβάνει υπόψη όχι μόνο εάν ένα μοντέλο θεωρεί μια παρατήρηση ανώμαλη, αλλά και σε ποιο βαθμό. Η δυνατότητα αυτή είναι ιδιαίτερα σημαντική όταν οι επιμέρους ανιχνευτές έχουν διαφορετικές αρχές λειτουργίας και επομένως προσφέρουν διαφορετικό είδος πληροφορίας για την ίδια παρατήρηση.

Συνολικά, οι τρεις τεχνικές συνδυασμού εκφράζουν διαφορετικές στρατηγικές αξιοποίησης των επιμέρους μοντέλων. Ο OR/Union συνδυασμός επισημαίνει μια παρατήρηση όταν τουλάχιστον ένας ανιχνευτής την αναγνωρίσει ως ανωμαλία και επομένως δίνει μεγαλύτερη έμφαση στην κάλυψη [43].

Ο συνδυασμός πλειοψηφικής ψηφοφορίας (Majority Voting) απαιτεί συμφωνία της πλειοψηφίας και εφαρμόζει περισσότερο συντηρητικό κανόνα συναίνεσης [42], [43]. Ο συνδυασμός σταθμισμένης ψηφοφορίας (Weighted Score-Based) διατηρεί τις συνεχείς βαθμολογίες ανωμαλίας, τις μεταφέρει σε συγκρίσιμη κλίμακα και επιτρέπει διαφορετική συνεισφορά κάθε ανιχνευτή [40], [41]. Η συγκριτική εξέταση αυτών των προσεγγίσεων επιτρέπει να διερευνηθεί κατά πόσο η συμπληρωματικότητα των Isolation Forest, LOF, HBOS και Autoencoder μπορεί να αξιοποιηθεί αποτελεσματικότερα από ένα ενιαίο σύστημα ανίχνευσης.

2.3.6 Μέθοδοι Ανάλυσης Συμπληρωματικότητας και Συνεισφοράς

Η αξιολόγηση ενός συνόλου διαφορετικών μοντέλων ανίχνευσης ανωμαλιών δεν περιορίζεται στη σύγκριση των συνολικών μετρικών απόδοσης. Δύο μοντέλα μπορεί να παρουσιάζουν παρόμοια συνολική επίδοση και ταυτόχρονα να εντοπίζουν διαφορετικές παρατηρήσεις ή να αποδίδουν διαφορετικές βαθμολογίες ανωμαλίας στις ίδιες συναλλαγές. Η διερεύνηση της συμπληρωματικότητας είναι επομένως σημαντική όταν εξετάζεται η δυνατότητα συνδυασμού διαφορετικών ανιχνευτών σε μια τεχνική συνδυασμού. Για την ανάλυση αυτή μπορούν να χρησιμοποιηθούν μέτρα που εξετάζουν τόσο τις συνεχείς βαθμολογίες των μοντέλων όσο και την επικάλυψη ή τη διαφοροποίηση των δυαδικών τους αποφάσεων.

Μία προσέγγιση για τη σύγκριση των συνεχών εξόδων των μοντέλων είναι ο δείκτης συσχέτισης Pearson. Ο συγκεκριμένος συντελεστής αποτυπώνει τον βαθμό γραμμικής συσχέτισης μεταξύ δύο αριθμητικών μεταβλητών και ορίζεται ως:

$$r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.24)$$

Η τιμή του βρίσκεται στο διάστημα $[-1, 1]$. Τιμές κοντά στο 1 υποδηλώνουν ισχυρή θετική γραμμική σχέση, τιμές κοντά στο -1 ισχυρή αρνητική γραμμική σχέση, ενώ τιμές κοντά στο 0 υποδηλώνουν περιορισμένη γραμμική συσχέτιση. Η τελευταία περίπτωση δεν συνεπάγεται απαραίτητα πλήρη απουσία οποιασδήποτε σχέσης, καθώς ο Pearson εξετάζει ειδικότερα τη γραμμική συνιστώσα της συσχέτισης. Η χρήση πληροφορίας βασισμένη στη συσχέτιση έχει αξιοποιηθεί και σε εφαρμογές ανίχνευσης ανωμαλιών για τη διερεύνηση σχέσεων που εμφανίζονται μέσα στα δεδομένα [54].

Στην περίπτωση πολλαπλών ανιχνευτών ανωμαλιών, ο συντελεστής συσχέτισης Pearson μπορεί να εφαρμοστεί μεταξύ των συνεχών βαθμολογιών ανωμαλίας δύο μοντέλων. Υψηλή θετική συσχέτιση υποδηλώνει ότι τα μοντέλα τείνουν να αποδίδουν παρόμοια σχετική βαθμολογία στις ίδιες παρατηρήσεις, ενώ χαμηλότερη συσχέτιση μπορεί να αποτελεί ένδειξη διαφορετικής συμπεριφοράς των ανιχνευτών. Η χαμηλή συσχέτιση, ωστόσο, δεν συνεπάγεται από μόνη της ότι ο συνδυασμός των δύο μοντέλων θα είναι περισσότερο αποτελεσματικός.

Ένα δεύτερο μέτρο αφορά την επικάλυψη μεταξύ δύο συνόλων και είναι ο δείκτης ομοιότητας Jaccard. Για δύο σύνολα A και B, ο δείκτης ορίζεται ως:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (2.25)$$

και λαμβάνει τιμές από 0 έως 1. Τιμή ίση με 0 αντιστοιχεί σε πλήρη απουσία κοινών στοιχείων, ενώ τιμή ίση με 1 αντιστοιχεί σε πλήρη ταύτιση των δύο συνόλων. Ο Jaccard έχει χρησιμοποιηθεί και σε εφαρμογές ανίχνευσης απάτης ως μέτρο ομοιότητας μεταξύ συνόλων που περιγράφουν διαφορετικές δομές ή γειτονιές των δεδομένων [55].

Στο πλαίσιο σύγκρισης διαφορετικών ανιχνευτών, τα σύνολα μπορούν να αντιστοιχούν στις παρατηρήσεις που αναγνωρίζονται από κάθε μοντέλο ως περιπτώσεις ενδιαφέροντος. Με τον τρόπο αυτό, ο δείκτης επιτρέπει την ποσοτικοποίηση της επικάλυψης των ανιχνεύσεων. Υψηλή τιμή υποδηλώνει ότι δύο μοντέλα εντοπίζουν μεγάλο ποσοστό κοινών περιπτώσεων, ενώ χαμηλότερη τιμή υποδηλώνει περισσότερο διαφοροποιημένη συμπεριφορά.

Η περιγραφική σύγκριση των προβλέψεων μπορεί να συμπληρωθεί από στατιστική διερεύνηση της διαφοροποίησης των δυαδικών αποφάσεων. Ο έλεγχος McNemar αποτελεί έλεγχο κατάλληλο για ζευγαρωμένα δυαδικά αποτελέσματα, δηλαδή για περιπτώσεις στις οποίες δύο μέθοδοι αξιολογούνται πάνω στις ίδιες ακριβώς παρατηρήσεις. Ο έλεγχος επικεντρώνεται στις ασύμφωνες περιπτώσεις των δύο μεθόδων. Εάν b είναι ο αριθμός των παρατηρήσεων στις οποίες η πρώτη μέθοδος δίνει την επιθυμητή απόφαση ενώ η δεύτερη όχι και c ο αντίστροφος αριθμός, η εκδοχή του McNemar με διόρθωση συνέχειας μπορεί να εκφραστεί ως:

$$X^2 = \frac{(|b-c|-1)^2}{b+c} \quad (2.26)$$

με έναν βαθμό ελευθερίας. Η βασική μηδενική υπόθεση είναι ότι οι δύο ασύμφωνες κατηγορίες εμφανίζονται με την ίδια πιθανότητα. Επομένως, ένα μικρό p-value παρέχει ένδειξη ασυμμετρίας μεταξύ των δύο μεθόδων και όχι απλώς ύπαρξης διαφορετικών προβλέψεων. Ο McNemar έχει μελετηθεί ως στατιστικός έλεγχος για τη σύγκριση αλγορίθμων μάθησης πάνω σε κοινές παρατηρήσεις [56].

Όταν πραγματοποιούνται πολλαπλοί στατιστικοί έλεγχοι, αυξάνεται η πιθανότητα να προκύψει τουλάχιστον ένα στατιστικά σημαντικό αποτέλεσμα λόγω τυχαίας διακύμανσης. Για τον περιορισμό αυτού του προβλήματος μπορούν να εφαρμοστούν διορθώσεις για πολλαπλές συγκρίσεις. Η διόρθωση Holm αποτελεί διαδοχική διαδικασία ελέγχου πολλαπλών υποθέσεων, με στόχο τον έλεγχο του συνολικού σφάλματος Τύπου I [57]. Τα p-values ταξινομούνται από το μικρότερο προς το μεγαλύτερο και εξετάζονται διαδοχικά με προσαρμοσμένα επίπεδα σημαντικότητας. Η μέθοδος προσφέρει μία συστηματική αντιμετώπιση του προβλήματος που προκύπτει όταν πραγματοποιούνται ταυτόχρονα πολλές ζευγαρωτές συγκρίσεις.

Εκτός από τη διερεύνηση της ομοιότητας και της διαφοροποίησης μεταξύ των μοντέλων, είναι χρήσιμο να εξετάζεται και η σχετική συνεισφορά κάθε ανιχνευτή στη συμπεριφορά ενός συνδυαστικού συστήματος. Για τον σκοπό αυτό μπορεί να αξιοποιηθεί το πλαίσιο της συνεργατικής Θεωρίας Παιγνίων. Σε ένα συνεργατικό παίγνιο θεωρείται ένα σύνολο παικτών N , οι οποίοι μπορούν να σχηματίζουν διαφορετικούς συνασπισμούς $S \subseteq N$. Σε κάθε συνασπισμό αντιστοιχίζεται μία τιμή $v(S)$, η οποία εκφράζει τη χρησιμότητα ή την απόδοση του συγκεκριμένου συνδυασμού.

Οι τιμές Shapley, οι οποίες εισήχθησαν από τον Shapley, προσφέρουν έναν τρόπο κατανομής της συνολικής αξίας ενός συνεργατικού παιγνίου στους επιμέρους παίκτες με βάση τη μέση οριακή συνεισφορά τους σε όλους τους δυνατούς συνασπισμούς [58]. Για έναν παίκτη i , η τιμή Shapley ορίζεται ως:

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} [u(S \cup \{i\}) - u(S)] \quad (2.27)$$

όπου N είναι το σύνολο όλων των παικτών και S ένας συνασπισμός που δεν περιλαμβάνει τον παίκτη i . Ο όρος $u(S \cup \{i\}) - u(S)$ εκφράζει την οριακή μεταβολή της αξίας του συνασπισμού όταν προστεθεί ο συγκεκριμένος παίκτης. Επειδή εξετάζονται όλοι οι δυνατοί συνασπισμοί, η Shapley Value δεν εξαρτάται μόνο από την ανεξάρτητη απόδοση ενός παίκτη αλλά από τη συνεισφορά του σε διαφορετικά συνεργατικά πλαίσια.

Η λογική των τιμών Shapley έχει αξιοποιηθεί ευρύτερα στη μηχανική μάθηση για την ερμηνεία σύνθετων και μοντέλων συνδυασμού [59]. Αντίστοιχα, προσεγγίσεις βασισμένες στις τιμές Shapley έχουν εφαρμοστεί και στο πεδίο της οικονομικής ανίχνευσης απάτης για την αποτίμηση της οριακής συνεισφοράς διαφορετικών χαρακτηριστικών ή ομάδων χαρακτηριστικών σε μία τελική πρόβλεψη [60]. Στην παρούσα εργασία, η βασική αρχή της τιμής Shapley μεταφέρεται από το επίπεδο των χαρακτηριστικών στο επίπεδο των ίδιων των ανιχνευτών ανωμαλιών. Τα επιμέρους μοντέλα αντιμετωπίζονται ως παίκτες και η συνεισφορά τους εξετάζεται σε διαφορετικούς συνασπισμούς μοντέλων. Η ακριβής διαμόρφωση της διαδικασίας και της συνάρτησης αξίας που χρησιμοποιείται στην πειραματική ανάλυση παρουσιάζεται στο Κεφάλαιο 3.

2.4 Νόμος του Benford

Ο Νόμος του Benford αποτελεί στατιστική προσέγγιση ψηφιακής ανάλυσης, η οποία περιγράφει την αναμενόμενη συχνότητα εμφάνισης των σημαντικών ψηφίων σε πολλά σύνολα αριθμητικών δεδομένων. Σε αντίθεση με την υπόθεση ότι τα ψηφία εμφανίζονται με ίση πιθανότητα, ο νόμος προβλέπει μια συγκεκριμένη λογαριθμική κατανομή, σύμφωνα με την οποία τα μικρότερα ψηφία εμφανίζονται συχνότερα στις πρώτες θέσεις ενός αριθμού [44], [47]. Η ιδιότητα αυτή έχει αξιοποιηθεί σε οικονομικά και λογιστικά δεδομένα, στον εσωτερικό έλεγχο, στην ανίχνευση πιθανών παραποιήσεων, σε εμπορικές συναλλαγές και σε εφαρμογές οικονομικής απάτης [44], [46]–[49].

Η χρήση του Benford διαφοροποιείται από τα μοντέλα μη επιβλεπόμενης μάθησης που παρουσιάστηκαν προηγουμένως. Δεν επιχειρεί να μάθει τη δομή των δεδομένων μέσω εκπαίδευσης ενός αλγορίθμου, αλλά συγκρίνει την παρατηρούμενη κατανομή των ψηφίων με μια θεωρητικά

αναμενόμενη κατανομή. Για τον λόγο αυτό μπορεί να λειτουργήσει ως συμπληρωματικό στατιστικό εργαλείο για την αναζήτηση αποκλίσεων οι οποίες απαιτούν περαιτέρω διερεύνηση.

2.4.1 Θεωρητικό Υπόβαθρο

Η αρχική παρατήρηση που οδήγησε στη διαμόρφωση του Νόμου του Benford βασίστηκε στο γεγονός ότι οι πρώτες σελίδες των παλαιών πινάκων λογαρίθμων παρουσίαζαν μεγαλύτερη φθορά από τις τελευταίες. Η συμπεριφορά αυτή υποδήλωνε ότι αριθμοί που ξεκινούσαν από μικρά ψηφία, όπως το 1 και το 2, χρησιμοποιούνταν συχνότερα από αριθμούς των οποίων το πρώτο ψηφίο ήταν μεγαλύτερο. Η συγκεκριμένη παρατήρηση διατυπώθηκε αρχικά το 1881 και διερευνήθηκε εκ νέου το 1938 μέσα από την εξέταση περισσότερων από 20.000 αριθμητικών παρατηρήσεων διαφορετικής προέλευσης [44], [45].

Η θεμελιώδης ιδέα του νόμου είναι ότι τα σημαντικά ψηφία (significant digits) πολλών φυσικά παραγόμενων αριθμητικών συνόλων δεν ακολουθούν ομοιόμορφη κατανομή. Ειδικότερα, η πιθανότητα εμφάνισης ενός ψηφίου στην πρώτη σημαντική θέση μειώνεται λογαριθμικά καθώς το ψηφίο αυξάνεται [46], [47]. Έτσι, το ψηφίο 1 εμφανίζεται ως πρώτο ψηφίο πολύ συχνότερα από το 9, ενώ η διαφορά μεταξύ των αναμενόμενων συχνοτήτων μειώνεται στις επόμενες ψηφιακές θέσεις.

Για έναν μη μηδενικό πραγματικό αριθμό x , η σημαντική του τιμή μπορεί να αναπαρασταθεί μέσω της συνάρτησης σημαντικού μέρους:

$$S(x) = 10^{\{\log_{10}|x|\}} \quad (2.28)$$

όπου $\{\log_{10}|x|\}$ αντιπροσωπεύει το κλασματικό μέρος του δεκαδικού λογαρίθμου του x . Η ποσότητα $S(x)$ βρίσκεται στο διάστημα $[1,10)$ και επιτρέπει τον προσδιορισμό των σημαντικών ψηφίων ανεξάρτητα από την τάξη μεγέθους του αρχικού αριθμού [47], [48].

Μια σημαντική ιδιότητα του Benford είναι η ανεξαρτησία από την κλίμακα μέτρησης (scale invariance). Η αλλαγή μονάδας μέτρησης δεν μεταβάλλει θεωρητικά τη συμμόρφωση ενός κατάλληλου συνόλου δεδομένων με τον νόμο. Για παράδειγμα, η μετατροπή χρηματικών ποσών από ένα νόμισμα σε άλλο μέσω πολλαπλασιασμού με σταθερό συντελεστή δεν αναμένεται από μόνη της να αλλάξει την ψηφιακή συμπεριφορά ενός συνόλου που ακολουθεί τον Benford [44], [46]. Η ιδιότητα αυτή ενισχύει τη χρησιμότητά του σε οικονομικά δεδομένα που μπορούν να εκφράζονται σε διαφορετικά νομίσματα ή μονάδες.

Ο Benford, ωστόσο, δεν αποτελεί καθολικό νόμο που πρέπει να ακολουθεί κάθε αριθμητικό σύνολο δεδομένων. Η καταλληλότητά του εξαρτάται από τη διαδικασία παραγωγής των δεδομένων και από τα χαρακτηριστικά του εξεταζόμενου συνόλου. Αποκλίσεις από την αναμενόμενη κατανομή μπορούν να προκύψουν ακόμη και χωρίς απάτη, λόγω της φυσικής δομής των δεδομένων, του μεγέθους του δείγματος, περιορισμών στην περιοχή των τιμών, στρογγυλοποιήσεων ή άλλων χαρακτηριστικών της διαδικασίας παραγωγής [45], [48]. Για τον λόγο αυτό, η μη συμμόρφωση δεν πρέπει να αντιμετωπίζεται αυτομάτως ως απόδειξη χειραγώγησης.

2.4.2 Κατανομή Πρώτου Σημαντικού Ψηφίου

Η πιο γνωστή μορφή του Νόμου του Benford αφορά την κατανομή του πρώτου σημαντικού ψηφίου (First-Digit Distribution). Έστω D_1 το πρώτο μη μηδενικό ψηφίο ενός αριθμού. Για $d \in \{1,2,\dots,9\}$, η πιθανότητα εμφάνισης του ψηφίου d στην πρώτη θέση δίνεται από [46]–[49]:

$$P(D_1 = d) = \log_{10} 1 + \frac{1}{d} \quad (2.29)$$

Η σχέση αυτή οδηγεί σε έντονα μη ομοιόμορφη κατανομή. Οι θεωρητικές πιθανότητες για τα εννέα δυνατά πρώτα ψηφία είναι:

Πίνακας 2.1. Απεικόνιση κατανομής πρώτου ψηφίου

Πρώτο ψηφίο d	Αναμενόμενη συχνότητα
1	30,10%
2	17,61%
3	12,49%
4	9,69%
5	7,92%
6	6,69%
7	5,80%
8	5,12%
9	4,58%

Επομένως, περίπου τρεις στις δέκα παρατηρήσεις ενός κατάλληλου συνόλου που ακολουθεί την κατανομή Benford αναμένεται να ξεκινούν από 1, ενώ λιγότερες από μία στις είκοσι ξεκινούν από 9 [44], [45], [49]. Η διαφορά αυτή αποτελεί το χαρακτηριστικότερο στοιχείο της κατανομής πρώτου σημαντικού ψηφίου.

Η λογική του νόμου μπορεί να επεκταθεί και σε περισσότερα σημαντικά ψηφία. Για μία ακολουθία k ψηφίων $d_1, d_2, \dots, d_k$, η αντίστοιχη πιθανότητα μπορεί να εκφραστεί ως:

$$P(D_1 = d_1, \dots, D_k = d_k) = \log_{10} \left(1 + \frac{1}{\sum_{i=1}^k d_i 10^{k-i}} \right) \quad (2.30)$$

[47],[49].Καθώς εξετάζονται επόμενες θέσεις, η κατανομή γίνεται σταδιακά περισσότερο ομοιόμορφη. Στη δεύτερη θέση επιτρέπεται πλέον και το ψηφίο 0, ενώ οι διαφορές μεταξύ των αναμενόμενων πιθανοτήτων είναι αισθητά μικρότερες. Για παράδειγμα, η αναμενόμενη συχνότητα του 0 ως δεύτερου ψηφίου είναι περίπου 12%, ενώ του 9 περίπου 8,5% [44], [45]. Στις ακόμη μεταγενέστερες θέσεις η κατανομή προσεγγίζει περισσότερο την ομοιομορφία.

Στην πρακτική ανάλυση, οι πραγματικές συχνότητες συγκρίνονται με τις θεωρητικές συχνότητες του Benford. Η απόκλιση μπορεί να αξιολογηθεί με στατιστικές τεχνικές όπως ο έλεγχος χ^2 Pearson, ο έλεγχος Z (Z-test) και η μέση απόλυτη απόκλιση (MAD) [44], [45], [49]. Ωστόσο, η επιλογή του ελέγχου πρέπει να γίνεται με προσοχή. Σε πολύ μεγάλα δείγματα, ορισμένοι κλασικοί έλεγχοι στατιστικής σημαντικότητας μπορεί να χαρακτηρίζουν ακόμη και μικρές και πρακτικά ασήμαντες αποκλίσεις ως στατιστικά σημαντικές [45].

Για τον λόγο αυτό, η συμμόρφωση με τον Benford δεν πρέπει να αντιμετωπίζεται αποκλειστικά ως δυαδική κατάσταση. Διαφορετικά σύνολα μπορούν να παρουσιάζουν διαφορετικό βαθμό συμμόρφωσης, ο οποίος επηρεάζεται από τα ιδιαίτερα χαρακτηριστικά της διαδικασίας που δημιούργησε τα δεδομένα [45]. Η αξιολόγηση της απόκλισης πρέπει επομένως να συνδυάζεται με γνώση του εξεταζόμενου οικονομικού ή λογιστικού περιβάλλοντος.

2.4.3 Εφαρμογές του Benford στην Ανάλυση Απάτης

Η χρήση του Benford στην ανάλυση απάτης βασίζεται στην υπόθεση ότι, όταν ένα κατάλληλο σύνολο οικονομικών δεδομένων παράγεται μέσω μιας φυσικής διαδικασίας, οι συχνότητες των σημαντικών ψηφίων ενδέχεται να προσεγγίζουν τη θεωρητική κατανομή. Αντίθετα, η ανθρώπινη παρέμβαση και η τεχνητή δημιουργία αριθμών μπορούν να μεταβάλουν την ψηφιακή δομή και να δημιουργήσουν αποκλίσεις από το αναμενόμενο πρότυπο [44], [46], [47].

Η ψηφιακή ανάλυση μπορεί επομένως να χρησιμοποιηθεί ως μηχανισμός αρχικού ελέγχου. Ένα σύνολο συναλλαγών που παρουσιάζει σημαντικές αποκλίσεις μπορεί να επισημανθεί για βαθύτερη εξέταση, επιτρέποντας τη συγκέντρωση των διαθέσιμων ελεγκτικών πόρων σε περισσότερο ύποπτες περιοχές [46], [49]. Η προσέγγιση αυτή είναι ιδιαίτερα χρήσιμη σε περιβάλλοντα με μεγάλο αριθμό οικονομικών εγγραφών, όπου η αναλυτική εξέταση κάθε συναλλαγής ξεχωριστά μπορεί να είναι πρακτικά αδύνατη. Η απόκλιση από τον Benford, ωστόσο, δεν ισοδυναμεί με απόδειξη απάτης. Μπορεί να αποτελεί ένδειξη ασυνήθιστης συμπεριφοράς, σφάλματος, διαφορετικού μηχανισμού παραγωγής των δεδομένων ή ενδεχόμενης χειραγώγησης και επομένως να δικαιολογεί περαιτέρω έλεγχο [45], [46], [49]. Η διάκριση αυτή είναι σημαντική, καθώς ακόμη και φυσικά παραγόμενα οικονομικά δεδομένα μπορούν να εμφανίζουν κάποιο βαθμό απόκλισης.

Χαρακτηριστική εφαρμογή αποτελεί η χρησιμοποίηση του Benford σε πραγματική υπόθεση νομιμοποίησης εσόδων από παράνομες δραστηριότητες. Οι ψηφιακές ιδιότητες των λογιστικών συναλλαγών χρησιμοποιήθηκαν για τη δημιουργία χαρακτηριστικών, τα οποία στη συνέχεια συνδυάστηκαν με μοντέλα Μηχανικής Μάθησης. Με τον τρόπο αυτό αναδείχθηκαν πρόσθετες εταιρείες που μπορούσαν να αποτελέσουν αντικείμενο περαιτέρω διερεύνησης [46]. Η εφαρμογή αυτή είναι ιδιαίτερα σημαντική επειδή παρουσιάζει τον Benford όχι ως απομονωμένο μηχανισμό τελικής απόφασης, αλλά ως πηγή πληροφορίας που μπορεί να συνδυαστεί με άλλες αναλυτικές τεχνικές.

Η σύγχρονη χρήση του νόμου αντιμετωπίζει και ένα περισσότερο σύνθετο πρόβλημα. Ένας δράστης που γνωρίζει τον Benford μπορεί θεωρητικά να διαμορφώσει τεχνητά δεδομένα με τρόπο ώστε η κατανομή του πρώτου ψηφίου να φαίνεται φυσιολογική. Σε μια τέτοια περίπτωση, ένας απλός έλεγχος πρώτου ψηφίου μπορεί να αποτύχει να αναγνωρίσει τη χειραγώγηση [48]. Για την αντιμετώπιση τέτοιων Benford-aware ή adversarial manipulations, έχουν προταθεί στατιστικές τεχνικές οι οποίες εξετάζουν πρόσθετη πληροφορία πέρα από την απλή συχνότητα του πρώτου ψηφίου και συνδυάζουν διαφορετικούς ελέγχους συμμόρφωσης.

Η εξέλιξη αυτή αναδεικνύει ότι ο Benford δεν πρέπει να θεωρείται στατικός και αλάνθαστος μηχανισμός ανίχνευσης απάτης. Η αποτελεσματικότητά του εξαρτάται από τη φύση των δεδομένων, τον τρόπο χειραγώγησης και τις αναλυτικές τεχνικές που χρησιμοποιούνται για την αξιολόγηση της συμμόρφωσης. Η χρήση του ως συμπληρωματικού ελέγχου μαζί με άλλες στατιστικές ή υπολογιστικές τεχνικές μπορεί επομένως να είναι περισσότερο κατάλληλη από την αποκλειστική χρήση της κατανομής του πρώτου σημαντικού ψηφίου [45], [46], [48].

2.4.4 Ανάλυση Benford σε Οικονομικά Δεδομένα

Τα οικονομικά και λογιστικά δεδομένα αποτελούν ένα από τα σημαντικότερα πεδία εφαρμογής του Benford. Η ύπαρξη μεγάλου αριθμού αριθμητικών παρατηρήσεων, όπως ποσά συναλλαγών, έσοδα, έξοδα, περιουσιακά στοιχεία και λογιστικές εγγραφές, δημιουργεί τις προϋποθέσεις για την εφαρμογή ψηφιακής ανάλυσης, εφόσον το εξεταζόμενο σύνολο δεδομένων διαθέτει τα κατάλληλα χαρακτηριστικά [44]–[46].

Σε εφαρμογή σε τραπεζικές οικονομικές καταστάσεις, εξετάστηκαν τόσο το πρώτο όσο και το δεύτερο ψηφίο στοιχείων ισολογισμού, αποτελεσμάτων και επιλεγμένων τραπεζικών δεικτών. Η σύγκριση των παρατηρούμενων και θεωρητικών συχνοτήτων παρουσίασε αποκλίσεις σε συγκεκριμένες κατηγορίες οικονομικών στοιχείων και χρονικές περιόδους, οι οποίες αξιολογήθηκαν ως ενδείξεις πιθανής λογιστικής χειραγώγησης και ανάγκης περαιτέρω διερεύνησης [44]. Η εφαρμογή αυτή δείχνει ότι ο Benford μπορεί να χρησιμοποιηθεί για τη διερεύνηση της αξιοπιστίας χρηματοοικονομικών στοιχείων σε τραπεζικά περιβάλλοντα.

Παράλληλα, η εφαρμογή του σε πραγματικά τραπεζικά δεδομένα αναδεικνύει την ανάγκη προσεκτικής ερμηνείας. Σε ανάλυση τραπεζικών υπολοίπων πολλών ετών, διαφορετικοί στατιστικοί έλεγχοι οδήγησαν σε αποτελέσματα που κυμαίνονταν από οριακή συμμόρφωση έως οριακή μη συμμόρφωση [45]. Το αποτέλεσμα αυτό επιβεβαιώνει ότι μια απόκλιση δεν μπορεί να ερμηνεύεται αυτόματα ως απάτη και ότι πρέπει να λαμβάνονται υπόψη το μέγεθος του δείγματος, η χρονική διάσταση και η διαδικασία δημιουργίας των δεδομένων.

Ο Benford έχει επίσης χρησιμοποιηθεί σε πραγματικές εμπορικές συναλλαγές που σχετίζονταν με νομιμοποίηση εσόδων από παράνομες δραστηριότητες. Η ψηφιακή συμπεριφορά των συναλλαγών περισσότερων από 600 προμηθευτών εξετάστηκε και μετατράπηκε σε χαρακτηριστικά τα οποία τροφοδότησαν μοντέλα Logistic Regression, Decision Trees, Neural Networks και Random Forests

[46]. Η προσέγγιση αυτή αναδεικνύει τη δυνατότητα δημιουργίας υβριδικών συστημάτων, όπου η ψηφιακή ανάλυση λειτουργεί ως συμπληρωματική πηγή πληροφορίας για μοντέλα Μηχανικής Μάθησης.

Εφαρμογές έχουν πραγματοποιηθεί επίσης στον χώρο του διεθνούς εμπορίου, όπου οι αποκλίσεις από την αναμενόμενη ψηφιακή συμπεριφορά χρησιμοποιούνται για τον εντοπισμό συναλλαγών ή οικονομικών φορέων που μπορεί να απαιτούν πρόσθετο έλεγχο [47], [48]. Νεότερες προσεγγίσεις εξετάζουν ακόμη και περιπτώσεις στις οποίες η πρώτη ψηφιακή κατανομή έχει σκόπιμα διαμορφωθεί ώστε να προσεγγίζει τον Benford, χρησιμοποιώντας πρόσθετες ιδιότητες των σημαντικών ψηφίων για την αναγνώριση πιθανής τεχνητής συμμόρφωσης [48].

Ιδιαίτερα σημαντική για περιβάλλοντα επιχειρησιακών συστημάτων είναι η εφαρμογή του Benford σε ERP δεδομένα. Σε πραγματικό ERP μεγάλης κρατικής επιχείρησης εξετάστηκαν οκτώ λειτουργικές μονάδες, περισσότερες από 1.900 βάσεις/πίνακες και δισεκατομμύρια εγγραφές, χρησιμοποιώντας την ανάλυση πρώτου σημαντικού ψηφίου, τον έλεγχο χ^2 Pearson και τη μέση απόλυτη απόκλιση [49]. Η μέθοδος αξιοποιήθηκε ως εργαλείο ελέγχου της ακεραιότητας μεγάλων βάσεων δεδομένων και εντοπισμού περιοχών που ενδέχεται να απαιτούν περαιτέρω εσωτερικό έλεγχο. Η συγκεκριμένη εφαρμογή δεν αντιμετωπίζει τη μη συμμόρφωση ως απόδειξη απάτης, αλλά ως προειδοποιητική ένδειξη για περαιτέρω διερεύνηση.

Η χρήση αυτή παρουσιάζει ιδιαίτερο ενδιαφέρον για την ανάλυση οικονομικών πληροφοριακών συστημάτων, καθώς δείχνει ότι ο Benford μπορεί να λειτουργήσει σε διαφορετικό επίπεδο από τα μοντέλα ανίχνευσης ανωμαλιών. Τα Isolation Forest, LOF, HBOS και Autoencoder αξιολογούν τη συμπεριφορά των παρατηρήσεων βάσει της δομής των χαρακτηριστικών τους, ενώ ο Benford εξετάζει τη μορφή της ψηφιακής κατανομής των αριθμητικών δεδομένων. Οι δύο κατηγορίες προσεγγίσεων μπορούν επομένως να παρέχουν διαφορετικού τύπου ενδείξεις για πιθανές αποκλίσεις.

Συνολικά, ο Νόμος του Benford μπορεί να θεωρηθεί ένα χρήσιμο εργαλείο πρώιμου ελέγχου και συμπληρωματικής ανάλυσης οικονομικών δεδομένων. Η απλότητα εφαρμογής του, η ανεξαρτησία του από την κλίμακα και η δυνατότητα εφαρμογής σε μεγάλους όγκους αριθμητικών παρατηρήσεων αποτελούν σημαντικά πλεονεκτήματα. Παράλληλα, η μη καθολική ισχύς του, η επίδραση του μεγέθους και της δομής του δείγματος και η δυνατότητα τεχνητής συμμόρφωσης περιορίζουν την ασφαλή χρήση του ως αυτόνομου μηχανισμού ανίχνευσης απάτης [45], [48]. Για τον λόγο αυτό, η αξιοποίησή του ως συμπληρωματικού στατιστικού φίλτρου μπορεί να προσφέρει πρόσθετη πληροφορία σε σχέση με τα αποτελέσματα των μοντέλων ανίχνευσης ανωμαλιών.

2.5 Ανασκόπηση Σχετικών Εργασιών

Η έρευνα γύρω από την ανίχνευση οικονομικής απάτης έχει στραφεί σταδιακά από συστήματα βασισμένα αποκλειστικά σε προκαθορισμένους κανόνες προς περισσότερο ευέλικτες προσεγγίσεις μηχανικής μάθησης και ανίχνευσης ανωμαλιών. Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι μέθοδοι μη επιβλεπόμενης μάθησης, καθώς η περιορισμένη διαθεσιμότητα αξιόπιστων ετικετών, η έντονη ανισορροπία μεταξύ νόμιμων και δόλιων συναλλαγών και η εμφάνιση νέων μορφών απάτης δυσκολεύουν την ανάπτυξη συστημάτων που βασίζονται αποκλειστικά σε ήδη επιβεβαιωμένα περιστατικά [16], [20], [27]. Παράλληλα, έχουν εξεταστεί τεχνικές συνδυασμού πολλών μοντέλων, καθώς και στατιστικές προσεγγίσεις όπως ο Νόμος του Benford, οι οποίες μπορούν να παρέχουν διαφορετικής φύσης ενδείξεις για ασυνήθιστη οικονομική συμπεριφορά.

Στο πλαίσιο της μη επιβλεπόμενης ανίχνευσης οικονομικών ανωμαλιών, ιδιαίτερο ενδιαφέρον παρουσιάζει η εφαρμογή μοντέλων Μηχανικής Μάθησης σε δεδομένα Γενικού Καθολικού. Σε πραγματικό σύνολο λογιστικών δεδομένων εφαρμόστηκαν, μεταξύ άλλων, τα Isolation Forest και Autoencoder με στόχο τον εντοπισμό ασυνήθιστων λογιστικών εγγραφών [20]. Τα αρχικά δεδομένα δεν περιείχαν ετικέτες ανωμαλιών, ενώ για την αξιολόγηση της διαδικασίας χρησιμοποιήθηκαν και τεχνητά δημιουργημένες αποκλίνουσες εγγραφές. Η εφαρμογή ανέδειξε τη δυνατότητα των μη επιβλεπόμενων μοντέλων να εξετάζουν μεγάλο αριθμό λογιστικών εγγραφών και να αποδίδουν βαθμολογίες ανωμαλίας, οι οποίες μπορούν να χρησιμοποιηθούν για την ιεράρχηση περιπτώσεων που απαιτούν περαιτέρω ελεγκτική διερεύνηση [20].

Η συγκεκριμένη προσέγγιση παρουσιάζει ιδιαίτερη χρησιμότητα σε περιβάλλοντα ERP, όπου ο μεγάλος όγκος οικονομικών συναλλαγών περιορίζει την αποτελεσματικότητα της αποκλειστικής χρήσης χειροκίνητου ελέγχου ή τυχαίας δειγματοληψίας. Σε τέτοιες περιπτώσεις, η μη επιβλεπόμενη

ανίχνευση ανωμαλιών μπορεί να λειτουργήσει ως αρχικό φίλτρο για τον εντοπισμό λογιστικών εγγραφών των οποίων η συμπεριφορά διαφοροποιείται από τη συνήθη δραστηριότητα.

Μια διαφορετική μη επιβλεπόμενη προσέγγιση εφαρμόστηκε σε τραπεζικές συναλλαγές, συνδυάζοντας Fuzzy C-Means, Isolation Forest και LOF [27]. Οι συναλλαγές αρχικά ομαδοποιήθηκαν βάσει της ομοιότητάς τους και στη συνέχεια χρησιμοποιήθηκαν μέθοδοι ανίχνευσης αποκλινουσών παρατηρήσεων για τον εντοπισμό περιπτώσεων που διαφοροποιούνταν από τα συνηθισμένα πρότυπα. Η προσέγγιση σχεδιάστηκε ώστε να μην εξαρτάται από προκαθορισμένες ετικέτες και να μπορεί να αναγνωρίζει σπάνιες ή προηγουμένως άγνωστες μορφές συμπεριφοράς [27]. Η συνδυασμένη χρήση ομαδοποίησης και ανίχνευσης ανωμαλιών υποδηλώνει ότι διαφορετικοί μη επιβλεπόμενοι μηχανισμοί μπορούν να αξιοποιηθούν συμπληρωματικά για την εξέταση οικονομικών συναλλαγών.

Η μη επιβλεπόμενη ανίχνευση απάτης έχει εξεταστεί επίσης μέσω βαθύτερων μοντέλων. Σε σχετική εφαρμογή συνδυάστηκαν Autoencoder, Self-Organizing Map και Restricted Boltzmann Machine με προσαρμόζόμενο μηχανισμό καθορισμού του κατωφλίου ανωμαλίας [16]. Η διαδικασία αξιολογήθηκε σε σύνολα δεδομένων οικονομικής απάτης, μεταξύ των οποίων και το IEEE-CIS, και ανέδειξε τη σημασία του τρόπου με τον οποίο οι συνεχείς βαθμολογίες ανωμαλίας μετατρέπονται σε τελικές αποφάσεις [16]. Η χρήση ενός σταθερού καθολικού κατωφλίου μπορεί να μην είναι κατάλληλη όταν η δομή των οικονομικών δεδομένων μεταβάλλεται ή παρουσιάζει έντονη ετερογένεια.

Οι παραπάνω εφαρμογές δείχνουν ότι η μη επιβλεπόμενη μάθηση μπορεί να αξιοποιηθεί τόσο σε τραπεζικές συναλλαγές όσο και σε λογιστικά δεδομένα χωρίς να απαιτείται πλήρης επισήμανση των παρατηρήσεων. Παράλληλα, όμως, οι διαφορετικές μέθοδοι δεν αξιολογούν τις αποκλίσεις με τον ίδιο τρόπο. Η συμπεριφορά ενός Isolation Forest, ενός LOF ή ενός Autoencoder επηρεάζεται από τη μορφή της ανωμαλίας, τη δομή των χαρακτηριστικών και τις παραμέτρους του μοντέλου [16], [20], [27]. Η διαφοροποίηση αυτή δημιουργεί ενδιαφέρον για την αξιοποίηση περισσότερων του ενός μοντέλων στο ίδιο σύστημα ανίχνευσης.

Στην κατεύθυνση αυτή, η ανίχνευση απάτης μέσω συνδυαστικών μοντέλων επιχειρεί να αξιοποιήσει τις προβλέψεις περισσότερων επιμέρους μεθόδων ώστε η τελική απόφαση να μην εξαρτάται από έναν μοναδικό μηχανισμό. Σε εφαρμογή ανίχνευσης απάτης σε πιστωτικές κάρτες χρησιμοποιήθηκε συνδυαστικό νευρωνικό μοντέλο μαζί με τεχνικές επεξεργασίας χαρακτηριστικών και αντιμετώπισης της ανισορροπίας των δεδομένων [38]. Η συγκριτική αξιολόγηση υποστήριξε τη δυνατότητα των συνδυαστικών προσεγγίσεων να αξιοποιούν διαφορετικές προβλεπτικές συμπεριφορές σε σχέση με τη χρήση ενός μόνο μοντέλου [38].

Αντίστοιχα, έχει εξεταστεί ο συνδυασμός διαφορετικών ταξινομητών, όπως SVM, KNN και Random Forest, μαζί με τεχνικές Bagging και Boosting για την ανίχνευση απάτης σε συναλλαγές πιστωτικών καρτών [39]. Η προσέγγιση συνδυάστηκε με τεχνικές αντιμετώπισης της ανισορροπίας των κλάσεων, όπως το SMOTE, και η σύγκριση με επιμέρους μοντέλα ανέδειξε τη δυνατότητα των συνδυαστικών συστημάτων να βελτιώνουν τη συνολική διαδικασία ανίχνευσης [39]. Ταυτόχρονα, η αύξηση της πολυπλοκότητας και του υπολογιστικού κόστους αποτελεί παράγοντα που πρέπει να λαμβάνεται υπόψη όταν οι τεχνικές αυτές εφαρμόζονται σε μεγάλους όγκους οικονομικών συναλλαγών.

Στην ίδια ερευνητική κατεύθυνση έχουν εξεταστεί διαφορετικοί μηχανισμοί συνδυασμού, όπως πλειοψηφική ψηφοφορία, σταθμισμένος μέσος όρος και συνδυασμός βαθμολογιών [42]. Η ύπαρξη διαφορετικών στρατηγικών υποδηλώνει ότι η αποτελεσματικότητα ενός συνδυαστικού μοντέλου δεν καθορίζεται αποκλειστικά από τα επιμέρους μοντέλα που συμμετέχουν, αλλά και από τον τρόπο με τον οποίο συγκεντρώνονται τα αποτελέσματά τους. Η πλειοψηφική ψηφοφορία βασίζεται στη συμφωνία των επιμέρους αποφάσεων, ενώ οι τεχνικές που χρησιμοποιούν συνεχείς βαθμολογίες διατηρούν περισσότερη πληροφορία σχετικά με τον βαθμό στον οποίο κάθε παρατήρηση θεωρείται ασυνήθιστη [42].

Η διαφορετικότητα μεταξύ των επιμέρους μοντέλων ανίχνευσης αποτελεί σημαντικό στοιχείο της λογικής των συνδυαστικών προσεγγίσεων. Έχει παρατηρηθεί ότι διαφορετικοί αλγόριθμοι μπορούν να εντοπίζουν μόνο μερικώς επικαλυπτόμενα σύνολα ύποπτων παρατηρήσεων [43]. Παράλληλα, η σύγκριση διαφορετικών κανόνων ψηφοφορίας, όπως OR, AND και πλειοψηφική ψηφοφορία, δείχνει ότι το επίπεδο συμφωνίας που απαιτείται μεταξύ των μοντέλων επηρεάζει άμεσα το τελικό αποτέλεσμα [43]. Στον κανόνα OR αρκεί μία θετική απόφαση ώστε η παρατήρηση να συμπεριληφθεί στο τελικό σύνολο, ενώ η πλειοψηφική ψηφοφορία απαιτεί συμφωνία των περισσότερων μοντέλων. Παρότι η συγκεκριμένη εφαρμογή δεν αφορά άμεσα την οικονομική απάτη, παρέχει χρήσιμη θεωρητική

τεκμηρίωση για τη συμπληρωματικότητα διαφορετικών μοντέλων και τον τρόπο συνδυασμού των αποφάσεών τους.

Ένας επιπλέον άξονας της σχετικής έρευνας αφορά ειδικότερα τη χρήση Autoencoders στην ανίχνευση οικονομικής απάτης. Σε δεδομένα Γενικού Καθολικού ο Autoencoder έχει χρησιμοποιηθεί ως μη επιβλεπόμενος μηχανισμός εντοπισμού λογιστικών εγγραφών που αποκλίνουν από τα συνηθισμένα πρότυπα [20]. Στην περίπτωση αυτή, το σφάλμα ανακατασκευής λειτουργεί ως βασικός δείκτης για την εκτίμηση του βαθμού ανωμαλίας μιας εγγραφής.

Ο Autoencoder έχει επίσης ενσωματωθεί σε ευρύτερο σύστημα ανίχνευσης ασυνήθιστης συμπεριφοράς για την ανάλυση οικονομικών συναλλαγών [33]. Η προσέγγιση συνδυάζει διαφορετικούς μηχανισμούς ανάλυσης και ανίχνευσης αποκλινουσών παρατηρήσεων, αναδεικνύοντας τη δυνατότητα ενός Autoencoder να χρησιμοποιείται ως τμήμα ενός ευρύτερου αναλυτικού συστήματος και όχι αποκλειστικά ως αυτόνομο μοντέλο [33].

Σε εφαρμογές πιστωτικών καρτών, οι Autoencoders έχουν επίσης ενσωματωθεί σε βαθύτερες αρχιτεκτονικές με στόχο την αποτελεσματικότερη ανίχνευση δόλιων συναλλαγών [37]. Η χρήση τους επιτρέπει την εκμάθηση εσωτερικών αναπαραστάσεων των δεδομένων και μπορεί να συμβάλει στην αποτύπωση σύνθετων σχέσεων μεταξύ των χαρακτηριστικών. Η ιδιότητα αυτή διαφοροποιεί τους Autoencoders από μεθόδους που βασίζονται αποκλειστικά σε αποστάσεις, πυκνότητες ή μονοδιάστατες κατανομές.

Η ένταξη Autoencoder σε μη επιβλεπόμενο συνδυαστικό σύστημα μαζί με άλλες τεχνικές [16] ενισχύει επίσης την άποψη ότι τα μοντέλα ανακατασκευής μπορούν να προσφέρουν συμπληρωματική πληροφορία. Παράλληλα, η αποτελεσματικότητά τους εξακολουθεί να εξαρτάται από την αρχιτεκτονική, την παρουσία ανωμαλιών στο σύνολο εκπαίδευσης και τον τρόπο καθορισμού του κατωφλίου. Συνεπώς, η συνδυασμένη χρήση τους με μοντέλα διαφορετικής αρχής λειτουργίας μπορεί να περιορίσει την εξάρτηση του τελικού συστήματος από έναν μοναδικό μηχανισμό ανίχνευσης.

Παράλληλα με τις προσεγγίσεις Μηχανικής Μάθησης, σημαντικό τμήμα της έρευνας αφορά την ανάλυση απάτης μέσω του Νόμου του Benford. Σε εφαρμογή στον τραπεζικό τομέα, ο Νόμος του Benford χρησιμοποιήθηκε για την εξέταση του πρώτου και του δεύτερου ψηφίου οικονομικών καταστάσεων και επιλεγμένων τραπεζικών μεταβλητών πριν και μετά από περίοδο χρηματοοικονομικής μηχανικής [44]. Οι αποκλίσεις από τις αναμενόμενες συχνότητες αξιολογήθηκαν με στατιστικούς ελέγχους και εντοπίστηκαν διαφοροποιήσεις σε στοιχεία που σχετίζονταν με την κεφαλαιακή επάρκεια, τη ρευστότητα και την ποιότητα του ενεργητικού [44]. Τα αποτελέσματα χρησιμοποιήθηκαν ως ενδείξεις πιθανής χειραγώγησης και όχι ως οριστική επιβεβαίωση απάτης.

Η ερμηνεία των αποτελεσμάτων του Benford απαιτεί, ωστόσο, ιδιαίτερη προσοχή. Σε πολυετή τραπεζικά δεδομένα, διαφορετικοί στατιστικοί έλεγχοι οδήγησαν σε αποτελέσματα που κυμαίνονταν μεταξύ οριακής συμμόρφωσης και οριακής μη συμμόρφωσης [45]. Παράλληλα, το μέγεθος του δείγματος μπορεί να επηρεάσει σημαντικά ορισμένους κλασικούς ελέγχους και να οδηγήσει σε υπερβολικά αυστηρά συμπεράσματα. Για τον λόγο αυτό, η απόκλιση από τη θεωρητική κατανομή δεν μπορεί να θεωρηθεί από μόνη της απόδειξη οικονομικής απάτης [45]. Η ανάλυση βάσει του Benford είναι περισσότερο κατάλληλη ως εργαλείο εντοπισμού περιπτώσεων που απαιτούν περαιτέρω διερεύνηση.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η συνδυασμένη χρήση του Benford με τεχνικές μηχανικής μάθησης σε πραγματική δικαστική υπόθεση νομιμοποίησης εσόδων από παράνομες δραστηριότητες [46]. Η ψηφιακή συμπεριφορά των λογιστικών συναλλαγών των εξεταζόμενων επιχειρήσεων μετατράπηκε σε αριθμητικά χαρακτηριστικά βασισμένα στον Νόμο του Benford και τα χαρακτηριστικά αυτά χρησιμοποιήθηκαν ως είσοδοι σε Logistic Regression, Neural Networks, Decision Trees και Random Forests. Η διαδικασία επέτρεψε την επισήμανση επιπλέον επιχειρήσεων που μπορούσαν να απαιτούν λεπτομερέστερη διερεύνηση [46]. Η εφαρμογή αυτή αποτελεί χαρακτηριστικό παράδειγμα του τρόπου με τον οποίο η ψηφιακή ανάλυση μπορεί να λειτουργήσει συμπληρωματικά προς μοντέλα Μηχανικής Μάθησης και όχι αποκλειστικά ως ανεξάρτητος στατιστικός έλεγχος.

Ο Νόμος του Benford έχει εξεταστεί επίσης σε περιβάλλοντα ERP και μεγάλων βάσεων δεδομένων. Σε πραγματικό σύστημα ERP μεγάλης κρατικής επιχείρησης εφαρμόστηκαν έλεγχοι συμμόρφωσης με τον Benford σε πολλαπλές λειτουργικές μονάδες και μεγάλο αριθμό εγγραφών [49]. Η ανάλυση χρησιμοποιήθηκε για τον εντοπισμό τμημάτων των δεδομένων που παρουσίαζαν αποκλίσεις και μπορούσαν να αποτελέσουν αντικείμενο περαιτέρω εσωτερικού ελέγχου. Η συγκεκριμένη εφαρμογή

τονίζει ότι η μη συμμόρφωση πρέπει να αντιμετωπίζεται ως ένδειξη που απαιτεί περαιτέρω διερεύνηση και όχι ως απόδειξη ύπαρξης απάτης [49].

Νεότερες προσεγγίσεις εξετάζουν ακόμη και το ενδεχόμενο ένας δράστης να γνωρίζει τον τρόπο λειτουργίας του Benford και να προσαρμόζει σκόπιμα τα δεδομένα ώστε η κατανομή του πρώτου σημαντικού ψηφίου να φαίνεται φυσιολογική [48]. Σε τέτοιες περιπτώσεις, η εξέταση αποκλειστικά του πρώτου ψηφίου μπορεί να μην επαρκεί για την αναγνώριση της χειραγώγησης. Για την αντιμετώπιση του προβλήματος έχουν αναπτυχθεί στατιστικές τεχνικές που αξιοποιούν περισσότερη πληροφορία από τη δομή των ψηφίων και συνδυάζουν διαφορετικούς ελέγχους συμμόρφωσης [48]. Η κατεύθυνση αυτή ενισχύει τη χρησιμότητα υβριδικών συστημάτων που συνδυάζουν ψηφιακή ανάλυση, στατιστικούς ελέγχους και μοντέλα Μηχανικής Μάθησης.

Η συνολική εικόνα των σχετικών εργασιών αναδεικνύει επομένως μια σταδιακή μετάβαση από τη χρήση μεμονωμένων τεχνικών προς περισσότερο συνδυαστικές και υβριδικές προσεγγίσεις. Οι μη επιβλεπόμενες μέθοδοι προσφέρουν τη δυνατότητα αναγνώρισης αποκλίσεων χωρίς πλήρη εξάρτηση από διαθέσιμες ετικέτες, αλλά διαφορετικά μοντέλα ανταποκρίνονται σε διαφορετικές μορφές ανωμαλιών [16], [20], [27]. Οι τεχνικές συνδυασμού επιχειρούν να αξιοποιήσουν αυτή τη διαφορετικότητα, ενώ ο τρόπος με τον οποίο ενώνονται οι επιμέρους αποφάσεις ή βαθμολογίες επηρεάζει άμεσα τη συμπεριφορά του τελικού συστήματος [38]–[43]. Οι Autoencoders προσθέτουν τη δυνατότητα εκμάθησης σύνθετων μη γραμμικών αναπαραστάσεων, ενώ ο Νόμος του Benford παρέχει μια διαφορετικής φύσης στατιστική ένδειξη, η οποία βασίζεται στην κατανομή των σημαντικών ψηφίων [44]–[49].

Στο σύνολο των εργασιών που εξετάστηκαν στη παρούσα ανασκόπηση, οι παραπάνω κατευθύνσεις εμφανίζονται κυρίως είτε ανεξάρτητα είτε σε επιμέρους συνδυασμούς. Έχουν εξεταστεί συνδυασμοί διαφορετικών μη επιβλεπόμενων μοντέλων [16], [27], συνδυαστικά συστήματα για την ανίχνευση οικονομικής απάτης [38], [39], [42], προσεγγίσεις βασισμένες σε Autoencoders [20], [33], [37] και συνδυασμένη αξιοποίηση του Νόμου του Benford με Μηχανική Μάθηση [46]. Στο σύνολο των εργασιών που εξετάστηκαν στο πλαίσιο της παρούσας ανασκόπησης δεν εντοπίστηκε κοινή πειραματική προσέγγιση που να αξιολογεί παράλληλα τα Isolation Forest, LOF, HBOS και Autoencoder ως ετερογενή μοντέλα ανίχνευσης ανωμαλιών, να συγκρίνει τις τεχνικές OR/Union, πλειοψηφικής ψηφοφορίας και σταθμισμένου συνδυασμού βαθμολογιών και, παράλληλα, να εξετάζει τον Νόμο του Benford ως συμπληρωματικό στατιστικό μηχανισμό.

Κεφάλαιο 3ο: Δεδομένα και Μεθοδολογία

Στο παρόν κεφάλαιο παρουσιάζονται τα σύνολα δεδομένων και η μεθοδολογική διαδικασία που ακολουθήθηκε για την ανάπτυξη και την αξιολόγηση του προτεινόμενου πλαισίου ανίχνευσης οικονομικών ανωμαλιών. Η πειραματική διαδικασία βασίζεται σε δύο σύνολα δεδομένων με διαφορετικό ρόλο. Το IEEE-CIS Fraud Detection χρησιμοποιείται ως σύνολο αναφοράς για την εκπαίδευση, τη ρύθμιση και την ποσοτική αξιολόγηση των μοντέλων, καθώς διαθέτει επισημασμένες περιπτώσεις απάτης. Αντίθετα, το δεύτερο σύνολο προέρχεται από πραγματικό σύστημα Enterprise Resource Planning (ERP) και περιλαμβάνει λογιστικές εγγραφές χωρίς διαθέσιμες ετικέτες απάτης.

Η χρήση των δύο συνόλων επιτρέπει τον διαχωρισμό της πειραματικής αξιολόγησης από την εφαρμογή σε πραγματικές επιχειρησιακές συνθήκες. Αρχικά, τα μοντέλα διερευνώνται στο IEEE-CIS, όπου οι διαθέσιμες ετικέτες καθιστούν δυνατή την αντικειμενική σύγκριση μέσω κατάλληλων μετρικών. Στη συνέχεια, η μεθοδολογία εφαρμόζεται στο μη επισημασμένο σύνολο ERP χωρίς ετικέτες, με στόχο την παραγωγή βαθμολογιών ανωμαλίας και την ιεράρχηση λογιστικών εγγραφών προς περαιτέρω διερεύνηση. Με τον τρόπο αυτό εξετάζεται όχι μόνο η πειραματική συμπεριφορά των τεχνικών ανίχνευσης, αλλά και η δυνατότητα εφαρμογής τους σε ένα διαφορετικό και περισσότερο επιχειρησιακό περιβάλλον.

3.1 Περιγραφή των Δεδομένων

3.1.1 Σύνολο Δεδομένων IEEE-CIS Fraud Detection

Το σύνολο δεδομένων IEEE-CIS Fraud Detection αποτελεί ένα εκτεταμένο σύνολο δεδομένων ηλεκτρονικών συναλλαγών, το οποίο έχει χρησιμοποιηθεί για τη μελέτη προβλημάτων ανίχνευσης οικονομικής απάτης. Το σύνολο περιλαμβάνει πληροφορίες που σχετίζονται τόσο με τα χαρακτηριστικά των συναλλαγών όσο και με στοιχεία ταυτοποίησης της συσκευής και του περιβάλλοντος από το οποίο πραγματοποιήθηκε η συναλλαγή. Η παρουσία πολλών και ετερογενών μεταβλητών δημιουργεί ένα πρόβλημα υψηλής διαστασιμότητας, στο οποίο τα ύποπτα πρότυπα δεν είναι απαραίτητο να εκφράζονται μέσω μίας μόνο ακραίας τιμής, αλλά ενδέχεται να προκύπτουν από σύνθετους συνδυασμούς χαρακτηριστικών. Η χρήση του συγκεκριμένου συνόλου σε πρόσφατες εργασίες ανίχνευσης απάτης υποστηρίζει επίσης την καταλληλότητά του ως απαιτητικού συνόλου αναφοράς για την αξιολόγηση διαφορετικών προσεγγίσεων [50].

Τα αρχικά δεδομένα οργανώνονται σε δύο βασικούς πίνακες, Transaction και Identity. Για τις ανάγκες της παρούσας εργασίας πραγματοποιήθηκε συνένωση των δύο πινάκων με βάση το κοινό αναγνωριστικό TransactionID. Μετά τη συνένωση προέκυψαν 590.540 συναλλαγές, οι οποίες αποτέλεσαν τη βάση της περαιτέρω προεπεξεργασίας και των πειραμάτων.

Η μεταβλητή isFraud παρέχει την πραγματική κατηγορία κάθε συναλλαγής και χρησιμοποιείται για την αξιολόγηση της ικανότητας των μοντέλων να εντοπίζουν τις περιπτώσεις απάτης. Από τις 590.540 συναλλαγές, οι 569.877 χαρακτηρίζονται ως νόμιμες και οι 20.663 ως περιπτώσεις απάτης. Επομένως, οι περιπτώσεις απάτης αντιστοιχούν περίπου στο 3,5% του συνολικού συνόλου δεδομένων.

Πίνακας 3.1. Βασικά χαρακτηριστικά του IEEE-CIS

Χαρακτηριστικό	Τιμή
Συνολικές συναλλαγές	590.540
Νόμιμες συναλλαγές	569.877
Συναλλαγές απάτης	20.663
Ποσοστό απάτης	≈3,5%
Μεταβλητή στόχου	isFraud
Πηγή χαρακτηριστικών	Transaction και Identity

Η κατανομή αυτή παρουσιάζει έντονη ανισορροπία μεταξύ των κατηγοριών, καθώς η μεγάλη πλειονότητα των παρατηρήσεων αφορά νόμιμες συναλλαγές. Το χαρακτηριστικό αυτό επηρεάζει άμεσα τόσο την επιλογή των μετρικών αξιολόγησης όσο και την ερμηνεία των αποτελεσμάτων. Για παράδειγμα, μία υψηλή συνολική ακρίβεια δεν αποτελεί από μόνη της επαρκή ένδειξη αποτελεσματικής ανίχνευσης, καθώς μπορεί να προκύψει ακόμη και όταν ένα μοντέλο αποτυγχάνει να εντοπίσει σημαντικό ποσοστό της μειοψηφικής κατηγορίας. Για τον λόγο αυτό, στην παρούσα εργασία δίνεται μεγαλύτερη έμφαση στις μετρικές Precision, Recall, F1-score και κυρίως στην περιοχή κάτω από την καμπύλη Precision–Recall, όπως αναλύεται στην Ενότητα 3.4.

Η επιλογή του IEEE-CIS ως κύριου συνόλου για την εκπαίδευση και την πειραματική αξιολόγηση βασίστηκε κυρίως στην κλίμακα και στην πολυπλοκότητά του. Ο μεγάλος αριθμός συναλλαγών και η ύπαρξη ετερογενών μεταβλητών δημιουργούν ένα περιβάλλον στο οποίο μπορούν να διερευνηθούν διαφορετικοί μηχανισμοί ανίχνευσης ανωμαλιών. Παράλληλα, η ύπαρξη πραγματικών ετικετών επιτρέπει την ποσοτική αποτίμηση των αποτελεσμάτων, χωρίς οι ίδιες οι ετικέτες να αποτελούν απαραίτητα είσοδο στους μηχανισμούς μη επιβλεπόμενης ανίχνευσης.

Επιπλέον, η δομή του συνόλου παρουσιάζει μεγαλύτερη πολυπλοκότητα από απλούστερα σύνολα δεδομένων οικονομικής απάτης, καθώς συνδυάζει πληροφορίες συναλλαγής με στοιχεία που σχετίζονται με την ταυτότητα και τη συσκευή. Το χαρακτηριστικό αυτό θεωρήθηκε χρήσιμο για την παρούσα εργασία, καθώς επιτρέπει την αξιολόγηση των μοντέλων σε δεδομένα όπου η ανωμαλία ενδέχεται να προκύπτει από πολλαπλές διαστάσεις και όχι αποκλειστικά από το χρηματικό ποσό μιας συναλλαγής. Πρόσφατες εφαρμογές στο ίδιο σύνολο αναδεικνύουν επίσης τη σημασία της αξιοποίησης σύνθετων σχέσεων μεταξύ συναλλαγών, συσκευών και άλλων χαρακτηριστικών για την ανάλυση προτύπων απάτης [50].

3.1.2 Σύνολο Δεδομένων ERP

Το δεύτερο σύνολο δεδομένων προέρχεται από πραγματικό πληροφοριακό σύστημα ERP και περιλαμβάνει εγγραφές του Γενικού Καθολικού. Σε αντίθεση με το IEEE-CIS, το οποίο χρησιμοποιείται ως ελεγχόμενο πειραματικό περιβάλλον με διαθέσιμες ετικέτες, το συγκεκριμένο σύνολο αντιπροσωπεύει ένα πραγματικό λογιστικό περιβάλλον στο οποίο δεν υπάρχει προκαθορισμένη πληροφορία σχετικά με το αν μία εγγραφή είναι κανονική ή ύποπτη. Η χρήση δεδομένων ERP για την ανίχνευση λογιστικών ανωμαλιών αποτελεί σχετικό πεδίο εφαρμογής τεχνικών μηχανικής μάθησης, καθώς οι εγγραφές των επιχειρησιακών συστημάτων μπορούν να περιλαμβάνουν μεγάλο όγκο και ετερογένεια λογιστικών κινήσεων [51].

Το αρχικό αρχείο περιείχε 65.506 γραμμές δεδομένων μετά την ανάγνωση του λογιστικού πίνακα. Κατά τον αρχικό έλεγχο εντοπίστηκαν 27 γραμμές οι οποίες δεν αντιστοιχούσαν σε πραγματικές λογιστικές εγγραφές, αλλά σε γραμμές μεταφοράς ή συγκεντρωτικών συνόλων. Οι εγγραφές αυτές αφαιρέθηκαν κατά την προεπεξεργασία, με αποτέλεσμα το τελικό σύνολο να αποτελείται από 65.479 λογιστικές εγγραφές.

Το διαθέσιμο αρχείο περιλαμβάνει, μεταξύ άλλων, ημερομηνία, παραστατικό, τύπο εγγραφής, κωδικό και περιγραφή λογαριασμού, καθώς και τα οικονομικά μεγέθη της χρέωσης, της πίστωσης και του υπολοίπου. Στο σύνολο εντοπίστηκαν περίπου 17.173 διαφορετικά παραστατικά, 262 διαφορετικοί λογαριασμοί και 6 τύποι λογιστικών εγγραφών. Μετά την ολοκλήρωση της προεπεξεργασίας και της δημιουργίας πρόσθετων χαρακτηριστικών, το τελικό διάνυσμα εισόδου αποτελείται από 15 χαρακτηριστικά.

Πίνακας 3.2. Βασικά χαρακτηριστικά του συνόλου ERP

Χαρακτηριστικό	Τιμή
Τελικές λογιστικές εγγραφές	65.479
Διαφορετικά παραστατικά	≈ 17.173
Διαφορετικοί λογαριασμοί	262
Τύποι εγγραφών	6
Τελικά χαρακτηριστικά μοντελοποίησης	15
Ετικέτες απάτης	Δεν υπάρχουν

Η χρονική πληροφορία του αρχείου απαιτεί ιδιαίτερη προσοχή. Παρότι η επικεφαλίδα του εξαγόμενου αρχείου αναφέρει περίοδο από 01/01/2026 έως 31/03/2026, οι πραγματικά διαθέσιμες εγγραφές εκτείνονται έως περίπου τις 27/01/2026. Συνεπώς, στην παρούσα εργασία η ανάλυση βασίζεται αποκλειστικά στις εγγραφές που ήταν πράγματι διαθέσιμες στο αρχείο και όχι στην ονομαστική χρονική περίοδο που αναφέρεται στην επικεφαλίδα του.

Η σημαντικότερη διαφορά του ERP από το IEEE-CIS είναι η απουσία πραγματικών ετικετών. Στο IEEE-CIS είναι δυνατός ο υπολογισμός True Positives, False Positives, False Negatives και True Negatives, καθώς είναι γνωστό ποιες συναλλαγές έχουν επισημανθεί ως απάτη. Αντίθετα, στο ERP δεν υπάρχει αντίστοιχη πληροφορία. Ως εκ τούτου, μία εγγραφή με υψηλή βαθμολογία ανωμαλίας δεν μπορεί να χαρακτηριστεί αυτομάτως ως περίπτωση οικονομικής απάτης. Αντιμετωπίζεται ως εγγραφή αυξημένης προτεραιότητας για περαιτέρω έλεγχο.

Υπάρχουν επίσης ουσιαστικές διαφορές ως προς τη φύση των χαρακτηριστικών. Το IEEE-CIS αφορά ηλεκτρονικές συναλλαγές και περιλαμβάνει μεταβλητές σχετικές με τη συναλλαγή, την ταυτότητα και τη συσκευή. Αντίθετα, το ERP περιέχει λογιστική πληροφορία, όπως λογαριασμούς, τύπους εγγραφών, παραστατικά, χρεώσεις, πιστώσεις και υπόλοιπα. Επομένως, τα δύο σύνολα δεν διαθέτουν κοινό χώρο χαρακτηριστικών και δεν επιχειρείται άμεση μεταφορά ενός ήδη εκπαιδευμένου μοντέλου από το IEEE-CIS στο ERP.

Ο ρόλος του ERP στην εργασία είναι συνεπώς διαφορετικός από εκείνον του IEEE-CIS. Το IEEE-CIS χρησιμοποιείται για την πειραματική διερεύνηση, ρύθμιση και ποσοτική αξιολόγηση της συμπεριφοράς των αλγορίθμων. Στο ERP εφαρμόζεται στη συνέχεια η ίδια γενική μεθοδολογική προσέγγιση, με τα επιμέρους μοντέλα να προσαρμόζονται εκ νέου στα πραγματικά λογιστικά δεδομένα. Η μεταφορά αφορά επομένως το πλαίσιο ανίχνευσης και τις επιλεγμένες διαμορφώσεις και όχι τις εκπαιδευμένες παραμέτρους των μοντέλων του IEEE-CIS.

Με αυτή τη διαδικασία διερευνάται κατά πόσο ένα πλαίσιο που αξιολογήθηκε αρχικά σε επισημασμένα δεδομένα μπορεί να χρησιμοποιηθεί για την παραγωγή χρήσιμων βαθμολογιών κινδύνου σε πραγματικές λογιστικές εγγραφές χωρίς διαθέσιμες ετικέτες. Η εφαρμογή αυτή δεν αποσκοπεί στην αυτόματη απόδοση χαρακτηρισμού «απάτη», αλλά στη δημιουργία ενός μηχανισμού ιεράρχησης που μπορεί να περιορίσει το σύνολο των εγγραφών που απαιτούν ανθρώπινο έλεγχο.

3.2 Προεπεξεργασία Δεδομένων

Η προεπεξεργασία αποτέλεσε απαραίτητο στάδιο πριν από την εφαρμογή των μοντέλων ανίχνευσης ανωμαλιών, καθώς τα δύο σύνολα δεδομένων παρουσίαζαν διαφορετική δομή, τύπους μεταβλητών και επίπεδα πληρότητας. Για τον λόγο αυτό δεν εφαρμόστηκε μία κοινή διαδικασία μετασχηματισμού, αλλά διαμορφώθηκε ξεχωριστή ακολουθία προεπεξεργασίας για το IEEE-CIS και για το σύνολο ERP.

Στο IEEE-CIS κύριες προκλήσεις αποτέλεσαν ο μεγάλος αριθμός χαρακτηριστικών, η υψηλή παρουσία ελλিপών τιμών και η συνύπαρξη αριθμητικών και κατηγορικών μεταβλητών. Αντίθετα, στο ERP το βασικό ζητούμενο ήταν η μετατροπή των αρχικών λογιστικών πεδίων σε ένα συνεπές αριθμητικό

σύνολο χαρακτηριστικών, κατάλληλο για την εφαρμογή των τεσσάρων μοντέλων ανίχνευσης ανωμαλιών.

Η συνολική διαδικασία περιλάμβανε καθαρισμό των δεδομένων, διαχείριση ελλিপών τιμών, κωδικοποίηση των κατηγορικών μεταβλητών και κλιμάκωση των τελικών χαρακτηριστικών.

3.2.1 Καθαρισμός Δεδομένων

IEEE-CIS

Στο IEEE-CIS πραγματοποιήθηκε αρχικά συνένωση των πινάκων Transaction και Identity μέσω του μοναδικού αναγνωριστικού TransactionID. Μετά τη διαδικασία αυτή προέκυψαν 590.540 παρατηρήσεις, οι οποίες αποτέλεσαν τη βάση για τα επόμενα στάδια προεπεξεργασίας.

Ακολούθησε έλεγχος της πληρότητας των μεταβλητών. Χαρακτηριστικά με ποσοστό ελλিপών τιμών μεγαλύτερο από 50% αφαιρέθηκαν, καθώς η περιορισμένη διαθέσιμη πληροφορία θα καθιστούσε τη συμπλήρωσή τους ιδιαίτερα εξαρτώμενη από τεχνητές εκτιμήσεις. Μετά το συγκεκριμένο φιλτράρισμα παρέμειναν 220 μεταβλητές πριν από τη διαμόρφωση του τελικού πίνακα εισόδου.

Το TransactionID χρησιμοποιήθηκε για την ενοποίηση των αρχικών πινάκων, αλλά δεν αξιοποιήθηκε ως χαρακτηριστικό των μοντέλων, καθώς λειτουργεί ως μοναδικό αναγνωριστικό και δεν αποτελεί χαρακτηριστικό της συμπεριφοράς μιας συναλλαγής. Αντίστοιχα, η μεταβλητή isFraud διατηρήθηκε αποκλειστικά για την αξιολόγηση των αποτελεσμάτων και δεν συμπεριλήφθηκε στο διάλυσμα εισόδου. Ως αποτέλεσμα, ο τελικός πίνακας που χρησιμοποιήθηκε στα πειράματα περιείχε 218 χαρακτηριστικά, όπως επιβεβαιώνεται και από το τελικό σχήμα του δείγματος που χρησιμοποιήθηκε στα πειράματα του Autoencoder, (100000, 218).

Στο τελευταίο στάδιο της προετοιμασίας, τα αριθμητικά δεδομένα μετατράπηκαν σε τύπο float32. Η επιλογή αυτή μείωσε τις απαιτήσεις μνήμης σε σχέση με τον προεπιλεγμένο τύπο float64, γεγονός ιδιαίτερα σημαντικό λόγω του μεγέθους του IEEE-CIS και της εκτέλεσης πολλαπλών πειραμάτων στο υπολογιστικό περιβάλλον Google Colab.

ERP

Η διαδικασία καθαρισμού του συνόλου ERP διαφοροποιήθηκε λόγω της λογιστικής φύσης των δεδομένων. Το αρχικό φύλλο περιείχε τις μεταβλητές:

- A/A,
- Ημερομηνία,
- Παραστατικό,
- Τύπος εγγραφής,
- Κωδικός λογαριασμού,
- Περιγραφή λογαριασμού,
- Χρέωση,
- Πίστωση,
- Υπόλοιπο.

Μετά την ανάγνωση του πραγματικού πίνακα δεδομένων εντοπίστηκαν 65.506 γραμμές. Από αυτές αφαιρέθηκαν 27 γραμμές οι οποίες δεν αντιστοιχούσαν σε πραγματικές λογιστικές κινήσεις. Συγκεκριμένα, περιλαμβάνονταν μία γραμμή μεταφοράς και 26 γραμμές συγκεντρωτικών συνόλων. Η απομάκρυνσή τους πραγματοποιήθηκε μέσω του ελέγχου εγκυρότητας της ημερομηνίας, με αποτέλεσμα να παραμείνουν 65.479 πραγματικές λογιστικές εγγραφές.

Η στήλη A/A αφαιρέθηκε από το σύνολο χαρακτηριστικών, καθώς αποτελεί σειριακό αναγνωριστικό και δεν μεταφέρει πληροφορία σχετικά με τη λογιστική συμπεριφορά της εγγραφής. Παρόμοια, η Περιγραφή λογαριασμού δεν χρησιμοποιήθηκε ως μεταβλητή μοντελοποίησης, επειδή περιέχει σε μεγάλο βαθμό την ίδια πληροφορία με τον Κωδικό λογαριασμού. Ωστόσο, διατηρήθηκε στο αρχικό σύνολο για την περιγραφική ερμηνεία των εγγραφών που επισημαίνονται από τα μοντέλα.

Το πεδίο Παραστατικό επίσης δεν χρησιμοποιήθηκε απευθείας ως κωδικοποιημένη μεταβλητή. Αντί αυτού, αξιοποιήθηκε για τη δημιουργία ενός χαρακτηριστικού συχνότητας, ώστε να διατηρηθεί πληροφορία σχετικά με το πόσο συχνά εμφανίζεται κάθε παραστατικό χωρίς να εισάγεται ο ίδιος ο μοναδικός ή ημιμοναδικός κωδικός στο μοντέλο.

Από την ημερομηνία δημιουργήθηκαν επιπλέον χρονικά χαρακτηριστικά:

- Ημέρα μήνα,

- Ημέρα εβδομάδας,
- Εβδομάδα,
- Ένδειξη Σαββατοκύριακου.

Παράλληλα, από τα οικονομικά και λογιστικά πεδία δημιουργήθηκαν τα ακόλουθα παράγωγα χαρακτηριστικά:

$$\text{Ποσό} = |\text{Χρέωση}| + |\text{Πίστωση}| \quad (3.1)$$

καθώς και δυαδικές μεταβλητές που δηλώνουν αν η συγκεκριμένη εγγραφή περιλαμβάνει χρέωση ή πίστωση. Επιπλέον, χρησιμοποιήθηκε η απόλυτη τιμή του υπολοίπου:

$$\text{Απόλυτο Υπόλοιπο} = |\text{Υπόλοιπο}| \quad (3.2)$$

και υπολογίστηκαν οι μεταβλητές Συχνότητα Λογαριασμού και Συχνότητα Παραστατικού, οι οποίες αποτυπώνουν πόσο συχνά εμφανίζεται κάθε αντίστοιχη τιμή στο σύνολο των λογιστικών εγγραφών. Μετά την ολοκλήρωση των παραπάνω μετασχηματισμών, το τελικό ERP σύνολο που χρησιμοποιήθηκε για τη μοντελοποίηση περιείχε 15 χαρακτηριστικά.

Πίνακας 3.3. Χαρακτηριστικά του τελικού συνόλου ERP

A/A	Χαρακτηριστικό
1	Τύπος εγγραφής
2	Κωδικός λογαριασμού
3	Χρέωση
4	Πίστωση
5	Υπόλοιπο
6	Ημέρα μήνα
7	Ημέρα εβδομάδας
8	Εβδομάδα
9	Σαββατοκύριακο
10	Ποσό
11	Είναι χρέωση
12	Είναι πίστωση
13	Απόλυτο υπόλοιπο
14	Συχνότητα λογαριασμού
15	Συχνότητα παραστ ατικού

Η συγκεκριμένη διαμόρφωση επιτρέπει την ταυτόχρονη αξιοποίηση οικονομικών, χρονικών και λογιστικών χαρακτηριστικών, χωρίς να εισάγονται άμεσα μοναδικά αναγνωριστικά που θα μπορούσαν να περιορίσουν τη γενίκευση της ανάλυσης.

3.2.2 Διαχείριση Ελλιπών Τιμών

Η παρουσία ελλιπών τιμών αντιμετωπίστηκε διαφορετικά ανάλογα με τον τύπο κάθε μεταβλητής. Στο IEEE-CIS, μετά την αφαίρεση των χαρακτηριστικών που παρουσίαζαν ποσοστό απουσίας μεγαλύτερο από 50%, εφαρμόστηκε συμπλήρωση των ελλιπών τιμών στις μεταβλητές που παρέμειναν. Για τις αριθμητικές μεταβλητές χρησιμοποιήθηκε η διάμεσος της αντίστοιχης μεταβλητής. Η επιλογή της διαμέσου κρίθηκε κατάλληλη λόγω της μικρότερης ευαισθησίας της σε ακραίες τιμές σε σύγκριση με τον αριθμητικό μέσο.

Για τις κατηγορικές μεταβλητές, οι ελλείψεις τιμές αντικαταστάθηκαν από την ειδική κατηγορία Unknown. Με τον τρόπο αυτό η απουσία μιας κατηγορικής πληροφορίας δεν οδηγεί σε διαγραφή ολόκληρης της συναλλαγής, αλλά διατηρείται ως ξεχωριστή κατάσταση στο σύνολο δεδομένων. Η συγκεκριμένη διαδικασία καταγράφεται και στο αρχείο των πειραμάτων της εργασίας.

Αντίστοιχη λογική εφαρμόστηκε και στο ERP. Για πιθανές ελλείψεις αριθμητικές τιμές χρησιμοποιήθηκε η διάμεσος της αντίστοιχης μεταβλητής, ενώ στις κατηγορικές μεταβλητές χρησιμοποιήθηκε η κατηγορία Unknown.

Μετά την ολοκλήρωση της διαδικασίας πραγματοποιήθηκε έλεγχος του τελικού πίνακα ERP, χωρίς να εντοπιστούν τιμές NaN ή άπειρες αριθμητικές τιμές (Inf). Το τελικό σύνολο είχε διαστάσεις: 65.479×15 και ήταν κατάλληλο για την εφαρμογή των επιλεγμένων αλγορίθμων.

3.2.3 Κωδικοποίηση Κατηγορικών Μεταβλητών

Τα μοντέλα που χρησιμοποιούνται στην εργασία απαιτούν αριθμητική αναπαράσταση των μεταβλητών. Για τον λόγο αυτό οι κατηγορικές μεταβλητές μετατράπηκαν σε αριθμητική μορφή μέσω κωδικοποίησης ετικετών (Label Encoding).

Στο IEEE-CIS, κάθε διακριτή κατηγορία μιας κατηγορικής μεταβλητής αντιστοιχίστηκε σε έναν ακέραιο αριθμό. Η συγκεκριμένη προσέγγιση επιλέχθηκε αντί της κωδικοποίησης One-Hot κυρίως λόγω της υψηλής διαστασιμότητας του συνόλου και του μεγάλου αριθμού διαφορετικών κατηγοριών που εμφανίζονται σε ορισμένες μεταβλητές. Η χρήση της κωδικοποίησης μοναδικής ένδειξης (One-Hot Encoding) θα αύξανε σημαντικά τον αριθμό των τελικών χαρακτηριστικών και, κατ' επέκταση, τις απαιτήσεις μνήμης και υπολογισμού.

Στο ERP, κωδικοποίηση ετικετών εφαρμόστηκε στις δύο βασικές κατηγορικές μεταβλητές που παρέμειναν στο σύνολο μοντελοποίησης:

- Τύπος εγγραφής,
- Κωδικός λογαριασμού.

Οι αντιστοιχίσεις αποθηκεύτηκαν ώστε να είναι δυνατή η μεταγενέστερη ανακατασκευή και ερμηνεία των αρχικών κατηγοριών.

Η επιλογή της συγκεκριμένης κωδικοποίησης συνοδεύεται από έναν μεθοδολογικό περιορισμό. Οι ακέραιες τιμές που παράγονται από την κωδικοποίηση ετικετών δεν αντιπροσωπεύουν πραγματική αριθμητική διάταξη ή απόσταση μεταξύ των κατηγοριών. Επομένως, μοντέλα που βασίζονται περισσότερο σε αποστάσεις ή στη γεωμετρική δομή του χώρου χαρακτηριστικών, όπως ο LOF και σε διαφορετικό βαθμό ο Autoencoder, ενδέχεται να επηρεάζονται από την τεχνητή αριθμητική αναπαράσταση των κατηγορικών μεταβλητών. Ο περιορισμός αυτός λαμβάνεται υπόψη κατά την ερμηνεία των αποτελεσμάτων.

3.2.4 Κανονικοποίηση / Κλιμάκωση Χαρακτηριστικών

Μετά την ολοκλήρωση του καθαρισμού, της συμπλήρωσης και της κωδικοποίησης, εφαρμόστηκε τυποποίηση των χαρακτηριστικών μέσω του StandardScaler.

Για κάθε αριθμητικό χαρακτηριστικό x , ο μετασχηματισμός πραγματοποιείται σύμφωνα με τη σχέση:

$$\frac{x-\mu}{\sigma} \quad (3.3)$$

όπου x είναι η αρχική τιμή, μ είναι η μέση τιμή του χαρακτηριστικού, σ είναι η τυπική του απόκλιση, z είναι η τυποποιημένη τιμή.

Με τον συγκεκριμένο μετασχηματισμό, τα χαρακτηριστικά αποκτούν κατά προσέγγιση μέση τιμή μηδέν και τυπική απόκλιση ίση με ένα.

Η κλιμάκωση ήταν ιδιαίτερα σημαντική καθώς τα αρχικά χαρακτηριστικά παρουσίαζαν σημαντικές διαφορές ως προς την αριθμητική τους κλίμακα. Για παράδειγμα, ένα χρηματικό ποσό μπορεί να λαμβάνει τιμές πολλών χιλιάδων μονάδων, ενώ μια κωδικοποιημένη κατηγορική μεταβλητή μπορεί να περιορίζεται σε σχετικά μικρούς ακέραιους αριθμούς. Χωρίς κλιμάκωση, τα χαρακτηριστικά μεγαλύτερου αριθμητικού εύρους θα μπορούσαν να επηρεάζουν δυσανάλογα μοντέλα που χρησιμοποιούν αποστάσεις ή σφάλματα ανακατασκευής.

Στο IEEE-CIS, ο StandardScaler εφαρμόστηκε στον τελικό πίνακα των 218 χαρακτηριστικών πριν από την εκπαίδευση των μοντέλων. Η προεπεξεργασία κατέληξε επομένως σε έναν πλήρως αριθμητικό και τυποποιημένο πίνακα κατάλληλο για την πειραματική διαδικασία. Η τελική μετατροπή του πίνακα σε float32 χρησιμοποιήθηκε επιπλέον για τη μείωση της κατανάλωσης μνήμης.

Στο ERP, εφαρμόστηκε νέος και ανεξάρτητος StandardScaler, ο οποίος προσαρμόστηκε αποκλειστικά στα πραγματικά λογιστικά δεδομένα. Ο scaler που είχε προσαρμοστεί στο IEEE-CIS δεν μεταφέρθηκε στο ERP, καθώς τα δύο σύνολα παρουσιάζουν διαφορετικά χαρακτηριστικά και διαφορετικές αριθμητικές κατανομές.

Η επιλογή αυτή είναι ουσιώδης για τη μεθοδολογία της εργασίας. Η εφαρμογή στα πραγματικά δεδομένα ERP δεν αποτελεί άμεση πρόβλεψη ενός μοντέλου εκπαιδευμένου στο IEEE-CIS. Αντίθετα, μεταφέρεται η μεθοδολογική διαμόρφωση του πλαισίου, ενώ η προεπεξεργασία και η προσαρμογή των επιμέρους μοντέλων πραγματοποιούνται εκ νέου στο ERP.

Μετά την ολοκλήρωση της διαδικασίας, ο τελικός πίνακας ERP είχε διαστάσεις $X_{ERP} \in \mathbb{R}^{65479 \times 15}$ με όλες τις τιμές αποθηκευμένες σε μορφή float32 και χωρίς ελλείψεις ή άπειρες τιμές.

3.3 Πειραματική Μεθοδολογία

Η πειραματική μεθοδολογία σχεδιάστηκε με στόχο τη συστηματική διερεύνηση της συμπεριφοράς των τεσσάρων επιλεγμένων μοντέλων ανίχνευσης ανωμαλιών, δηλαδή των Isolation Forest, Local Outlier Factor (LOF), Histogram-Based Outlier Score (HBOS) και Autoencoder. Ως βασικό πειραματικό σύνολο χρησιμοποιήθηκε το IEEE-CIS Fraud Detection, καθώς η ύπαρξη πραγματικών ετικετών επέτρεψε την ποσοτική αξιολόγηση των αποτελεσμάτων και τη σύγκριση διαφορετικών πειραματικών διαμορφώσεων.

Η εκπαίδευση των Isolation Forest, LOF και HBOS πραγματοποιήθηκε χωρίς τη χρήση της μεταβλητής isFraud ως χαρακτηριστικού εισόδου ή στόχου επιβλεπόμενης μάθησης. Οι πραγματικές ετικέτες αξιοποιήθηκαν για τη στρωματοποιημένη δημιουργία του κοινού δείγματος αξιολόγησης, για τον υπολογισμό των μετρικών απόδοσης και για την επιλογή των τελικών πειραματικών διαμορφώσεων. Στην τελική διαμόρφωση του Autoencoder χρησιμοποιήθηκαν επιπλέον για τον σχηματισμό συνόλου εκπαίδευσης που περιλαμβάνει μόνο κανονικές παρατηρήσεις. Η συγκεκριμένη περίπτωση αντιμετωπίζεται επομένως ως στρατηγική μίας κλάσης και διακρίνεται από την πλήρως μη επιβλεπόμενη εκπαίδευση.

Η πειραματική διερεύνηση δεν περιορίστηκε σε μία προκαθορισμένη διαμόρφωση ανά μοντέλο. Εξετάστηκαν διαφορετικοί συνδυασμοί υπερπαραμέτρων και διαφορετικά σημεία λειτουργίας, με στόχο τη διερεύνηση της επίδρασής τους στις μετρικές αξιολόγησης και ιδιαίτερα στον συμβιβασμό μεταξύ Precision και Recall. Τα αποτελέσματα των επιμέρους μοντέλων παρουσιάζονται αναλυτικά στο Κεφάλαιο 4.

3.3.1 Πειραματικό Πρωτόκολλο

Μετά την ολοκλήρωση της διαδικασίας προεπεξεργασίας που περιγράφηκε στην Ενότητα 3.2, προέκυψε ο τελικός αριθμητικός πίνακας χαρακτηριστικών του IEEE-CIS. Η γενική πειραματική διαδικασία οργανώθηκε σε διαδοχικά στάδια.

Αρχικά, πραγματοποιήθηκε ανεξάρτητη εφαρμογή των τεσσάρων μοντέλων ανίχνευσης ανωμαλιών. Για κάθε παρατήρηση, το αντίστοιχο μοντέλο παρήγαγε είτε άμεσα είτε μέσω κατάλληλου

μετασχηματισμού μία αριθμητική βαθμολογία ανωμαλίας. Υψηλότερες τιμές της βαθμολογίας αντιστοιχούν σε παρατηρήσεις που το εκάστοτε μοντέλο θεωρεί περισσότερο ασυνήθιστες σε σχέση με τη δομή των δεδομένων που έχει εκτιμήσει.

Η έκταση του συνόλου δεδομένων που χρησιμοποιήθηκε διαφοροποιήθηκε ανάλογα με τις υπολογιστικές απαιτήσεις κάθε αλγορίθμου. Το Isolation Forest και το HBOS ήταν δυνατό να εφαρμοστούν στο πλήρες σύνολο των 590.540 συναλλαγών κατά τη διερεύνηση των υπερπαραμέτρων τους. Αντίθετα, για το LOF χρησιμοποιήθηκε στρωματοποιημένο υποσύνολο 100.000 συναλλαγών, λόγω των σημαντικά αυξημένων απαιτήσεων σε χρόνο και μνήμη που συνεπάγεται ο υπολογισμός των τοπικών γειτονιών σε δεδομένα υψηλής διαστασιμότητας.

Το δείγμα των 100.000 συναλλαγών δημιουργήθηκε με στρωματοποιημένη δειγματοληψία και σταθερή τιμή `random_state=42`, ώστε να διατηρηθεί η αρχική αναλογία των δύο κατηγοριών. Το δείγμα περιλαμβάνει 96.501 νόμιμες συναλλαγές και 3.499 περιπτώσεις απάτης, διατηρώντας ποσοστό απάτης περίπου 3,5%.

Η ίδια λογική χρησιμοποιήθηκε και στα πειράματα του Autoencoder, ο οποίος αξιολογήθηκε στο ίδιο στρωματοποιημένο σύνολο των 100.000 συναλλαγών. Επιπλέον, για τις συγκρίσεις μεταξύ των διαφορετικών μοντέλων χρησιμοποιήθηκε κοινό δείγμα, ώστε οι διαφορές στις μετρικές να μην οφείλονται σε διαφορετική σύνθεση των δεδομένων αξιολόγησης.

Για τον Autoencoder διερευνήθηκαν δύο στρατηγικές εκπαίδευσης. Στην πρώτη, η εκπαίδευση πραγματοποιήθηκε χρησιμοποιώντας ολόκληρο το στρωματοποιημένο δείγμα, χωρίς διάκριση μεταξύ νόμιμων και δόλιων συναλλαγών. Στη δεύτερη στρατηγική, η εκπαίδευση περιορίστηκε στις 96.501 συναλλαγές που χαρακτηρίζονταν ως νόμιμες, ενώ η αξιολόγηση πραγματοποιήθηκε στο σύνολο των 100.000 παρατηρήσεων. Η δεύτερη προσέγγιση επιτρέπει στον Autoencoder να μοντελοποιήσει κυρίως τη φυσιολογική δομή των δεδομένων, ώστε παρατηρήσεις με υψηλότερο σφάλμα ανακατασκευής να αντιμετωπίζονται ως περισσότερο ασυνήθιστες.

Πρέπει να επισημανθεί ότι στη δεύτερη στρατηγική οι πραγματικές ετικέτες χρησιμοποιήθηκαν για τον σχηματισμό του συνόλου εκπαίδευσης που περιλαμβάνει μόνο κανονικές συναλλαγές. Συνεπώς, η συγκεκριμένη διαμόρφωση προσεγγίζει στρατηγική ανίχνευσης μίας κλάσης (one-class), παρότι ο Autoencoder δεν εκπαιδεύεται να προβλέπει άμεσα την κατηγορία της συναλλαγής. Η διάκριση αυτή λαμβάνεται υπόψη τόσο στην παρουσίαση της μεθοδολογίας όσο και στην ερμηνεία των αποτελεσμάτων.

Μετά την παραγωγή των τελικών βαθμολογιών ανωμαλίας, οι προβλέψεις των μοντέλων συγκρίθηκαν με την πραγματική μεταβλητή `isFraud`. Η αξιολόγηση πραγματοποιήθηκε με Precision, Recall, F1-score, ROC-AUC, AUPRC και Πίνακα Σύγχυσης, οι οποίες παρουσιάζονται αναλυτικά στην Ενότητα 3.4. Η Accuracy δεν χρησιμοποιήθηκε ως κριτήριο επιλογής, καθώς η έντονη ανισορροπία των κατηγοριών μπορεί να οδηγήσει σε υψηλή τιμή ακόμη και όταν η μειοψηφική κατηγορία ανιχνεύεται ανεπαρκώς.

Με στόχο την αναπαραγωγικότητα των πειραμάτων, όπου ήταν δυνατό χρησιμοποιήθηκαν σταθερές τιμές τυχαιότητας και κοινές διαδικασίες προεπεξεργασίας. Παράλληλα, τα δεδομένα μετατράπηκαν σε μορφή `float32`, ώστε να περιοριστούν οι απαιτήσεις μνήμης κατά την εκτέλεση των πειραμάτων στο Google Colab.

3.3.2 Εκπαίδευση και Αξιολόγηση των Μοντέλων

Η διαδικασία εκπαίδευσης και αξιολόγησης πραγματοποιήθηκε ξεχωριστά για κάθε μοντέλο, καθώς οι τέσσερις αλγόριθμοι βασίζονται σε διαφορετικούς μηχανισμούς υπολογισμού της ανωμαλίας.

Για τα Isolation Forest, LOF και HBOS, οι πραγματικές ετικέτες χρησιμοποιήθηκαν μετά την παραγωγή των βαθμολογιών και των προβλέψεων, ώστε να υπολογιστούν οι μετρικές αξιολόγησης. Στον τελικό Autoencoder ισχύει επιπλέον η ειδική χρήση των ετικετών για τον σχηματισμό του συνόλου εκπαίδευσης μόνο με κανονικές παρατηρήσεις, όπως περιγράφηκε στην προηγούμενη ενότητα.

Στην περίπτωση του Autoencoder, η βαθμολογία ανωμαλίας ορίστηκε από το σφάλμα ανακατασκευής της κάθε συναλλαγής. Μετά την εκπαίδευση του νευρωνικού δικτύου, κάθε παρατήρηση εισαγόταν στο μοντέλο και υπολογιζόταν η μέση τετραγωνική απόκλιση μεταξύ της αρχικής και της ανακατασκευασμένης αναπαράστασης. Μεγαλύτερο σφάλμα ανακατασκευής αντιστοιχούσε σε υψηλότερη βαθμολογία ανωμαλίας.

Η διαδικασία αξιολόγησης εξυπηρετούσε δύο διαφορετικούς σκοπούς. Ο πρώτος ήταν η αποτίμηση της ικανότητας κάθε διαμόρφωσης να κατατάσσει τις πραγματικές περιπτώσεις απάτης υψηλότερα από τις νόμιμες συναλλαγές. Ο δεύτερος ήταν η διερεύνηση της συμπεριφοράς των μοντέλων σε συγκεκριμένα λειτουργικά σημεία, όπου μέρος των συναλλαγών χαρακτηριζόταν ως ύποπτο.

Για τον λόγο αυτό, η επιλογή των τελικών διαμορφώσεων δεν βασίστηκε σε μία μεμονωμένη μετρική, αλλά στη συνολική συνεκτίμηση των Precision, Recall, F1-score και AUPRC/AP. Ιδιαίτερη σημασία αποδόθηκε στο Recall και στην AUPRC/AP λόγω της έντονης ανισορροπίας των κλάσεων, χωρίς όμως να παραβλέπεται ο αντίστοιχος συμβιβασμός με την Precision και τη συνολική συμπεριφορά των μοντέλων.

Οι αναλυτικές επιδόσεις που προέκυψαν από την παραπάνω διαδικασία δεν παρουσιάζονται στην παρούσα ενότητα, αλλά στο Κεφάλαιο 4, όπου πραγματοποιείται η συγκριτική αξιολόγηση των τεσσάρων μοντέλων.

3.3.3 Διαμόρφωση των Μοντέλων

Η επιλογή των υπερπαραμέτρων πραγματοποιήθηκε μέσω συστηματικής πειραματικής διερεύνησης διαφορετικών διαμορφώσεων. Η έκταση της διερεύνησης προσαρμόστηκε στη φύση κάθε αλγορίθμου και στον αριθμό των ουσιαστικών υπερπαραμέτρων του.

Για το Isolation Forest εξετάστηκαν τέσσερις βασικές παράμετροι: ο αριθμός των δέντρων `n_estimators`, το μέγεθος του υποδείγματος που χρησιμοποιείται σε κάθε δέντρο `max_samples`, το ποσοστό των χαρακτηριστικών `max_features` και η παράμετρος `contamination`. Οι τιμές που εξετάστηκαν ήταν:

- `n_estimators`: 100, 200, 300
- `max_samples`: 128, 256, 512
- `max_features`: 0,4, 0,6, 0,8
- `contamination`: 0,035, 0,06, 0,08, 0,10

Ο πλήρης συνδυασμός των παραπάνω τιμών οδήγησε σε 108 διαφορετικές πειραματικές διαμορφώσεις. Στόχος δεν ήταν μόνο η αναζήτηση μίας μέγιστης τιμής απόδοσης, αλλά και η εξέταση της ευαισθησίας του μοντέλου ως προς κάθε υπερπάρμετρο.

Για το LOF εξετάστηκαν οι παράμετροι `n_neighbors` και `contamination`. Οι τιμές ήταν:

- `n_neighbors`: 5, 10, 20, 50
- `contamination`: 0,035, 0,06, 0,08, 0,10

Προέκυψαν επομένως 16 διαφορετικές διαμορφώσεις. Η πειραματική διερεύνηση πραγματοποιήθηκε στο στρωματοποιημένο δείγμα των 100.000 συναλλαγών.

Αντίστοιχα, για το HBOS εξετάστηκε η επίδραση του αριθμού διαστημάτων των ιστογραμμάτων `n_bins` και του `contamination`:

- `n_bins`: 10, 20, 30, 50
- `contamination`: 0,035, 0,06, 0,08, 0,10

Ο συνδυασμός τους δημιούργησε επίσης 16 πειραματικές διαμορφώσεις.

Για τον Autoencoder εξετάστηκαν τρεις βασικοί παράγοντες: η διάσταση της λανθάνουσας αναπαράστασης, ο μέγιστος αριθμός εποχών εκπαίδευσης και το λειτουργικό κατώφλι:

- `latent_dim`: 8, 16, 32
- `epochs`: 20, 50, 100
- `contamination`: 0,035, 0,06, 0,08, 0,10

Για κάθε στρατηγική εκπαίδευσης εξετάστηκαν επομένως 36 διαμορφώσεις λειτουργίας. Η αρχιτεκτονική του δικτύου διατηρήθηκε σταθερή κατά τη διερεύνηση και περιλάμβανε δύο κρυφά επίπεδα κωδικοποίησης 128 και 64 νευρώνων, τη λανθάνουσα αναπαράσταση, και συμμετρικά επίπεδα αποκωδικοποίησης 64 και 128 νευρώνων. Στα κρυφά επίπεδα χρησιμοποιήθηκε συνάρτηση ενεργοποίησης ReLU, ενώ στο επίπεδο εξόδου χρησιμοποιήθηκε γραμμική ενεργοποίηση. Η εκπαίδευση πραγματοποιήθηκε με βελτιστοποιητή Adam και συνάρτηση απώλειας το μέσο τετραγωνικό σφάλμα.

Για τον περιορισμό της υπερπροσαρμογής χρησιμοποιήθηκε μηχανισμός Early Stopping με παρακολούθηση του σφάλματος επικύρωσης, `'patience=5'` και επαναφορά των βαρών της καλύτερης εποχής. Το 10% των δεδομένων εκπαίδευσης χρησιμοποιήθηκε ως σύνολο επικύρωσης.

Ιδιαίτερη προσοχή απαιτεί η ερμηνεία της παραμέτρου `contamination`. Στην παρούσα εργασία η παράμετρος δεν αντιμετωπίζεται ως εκτίμηση του πραγματικού ποσοστού απάτης. Χρησιμοποιείται ως λειτουργική παράμετρος ευαισθησίας, η οποία καθορίζει το ποσοστό των παρατηρήσεων που θα χαρακτηριστεί ως ανώμαλο σε ένα συγκεκριμένο λειτουργικό σημείο.

Για μία βαθμολογία ανωμαλίας (s_i) και μία επιλεγμένη τιμή contamination (c), το αντίστοιχο κατώφλι μπορεί να εκφραστεί γενικά ως:

$$\tau_c = Q_1 - c(S) \quad (3.4)$$

όπου $Q_1 - c$ είναι το αντίστοιχο ποσοστημόριο της κατανομής των βαθμολογιών ανωμαλίας. Η τελική απόφαση δίνεται από:

$$\hat{y}_i = \begin{cases} 1, & s_i \geq \tau_c \\ 0, & s_i < \tau_c \end{cases} \quad (3.5)$$

όπου η τιμή 1 αντιστοιχεί σε εγγραφή που επισημαίνεται ως ανωμαλία.

Για παράδειγμα, τιμή ευαισθησίας contamination=0,10 σημαίνει ότι το λειτουργικό κατώφλι τοποθετείται έτσι ώστε περίπου το 10% των παρατηρήσεων με τις υψηλότερες βαθμολογίες ανωμαλίας να επισημαίνονται για περαιτέρω εξέταση. Δεν σημαίνει ότι θεωρείται εκ των προτέρων πως το 10% των συναλλαγών αποτελεί απάτη.

Η μεταβολή του συγκεκριμένου λειτουργικού σημείου επηρεάζει κυρίως τις μετρικές που βασίζονται σε δυαδικές αποφάσεις, όπως Precision, Recall και F1-score. Αντίθετα, μετρικές που αξιολογούν την κατάταξη των συνεχών βαθμολογιών, όπως η AUPRC και η ROC-AUC, δεν μεταβάλλονται υποχρεωτικά όταν αλλάζει αποκλειστικά το κατώφλι χωρίς να αλλάζει η σειρά κατάταξης των παρατηρήσεων.

Με βάση τη συνολική πειραματική διερεύνηση, οι διαμορφώσεις που επιλέχθηκαν για την τελική αξιολόγηση των μεμονωμένων μοντέλων ήταν:

Πίνακας 3.4. Τελικές διαμορφώσεις των μοντέλων

Μοντέλο	Τελική Διαμόρφωση
IF	n_estimators=200, contamination=0.100, max_features=0.4, max_samples=128
LOF	n_neighbors=50, contamination=0.100
HBOS	n_bins=50, contamination=0.100
AE	latent_dim=32, epochs=100, contamination=0.100

Η επιλογή των παραπάνω διαμορφώσεων βασίστηκε στη συνολική συνεκτίμηση των Precision, Recall, F1-score και AUPRC/AP, με αυξημένη έμφαση στο Recall και στην AUPRC/AP λόγω της έντονης ανισορροπίας των κλάσεων του IEEE-CIS. Η ποσοτική συμπεριφορά των τελικών διαμορφώσεων και η μεταξύ τους σύγκριση παρουσιάζονται στο επόμενο κεφάλαιο.

3.3.4 Ανάλυση Συμπληρωματικότητας και Συνεισφοράς των Μοντέλων

Πέρα από τη σύγκριση των συνολικών μετρικών απόδοσης, πραγματοποιήθηκε συμπληρωματική ανάλυση με στόχο να διερευνηθεί κατά πόσο τα Isolation Forest, LOF, HBOS και Autoencoder παρέχουν διαφορετική ή πλεονάζουσα πληροφορία. Η ανάλυση πραγματοποιήθηκε στο κοινό δείγμα των 100.000 συναλλαγών του IEEE-CIS, το οποίο περιλαμβάνει 96.501 κανονικές και 3.499 πραγματικές περιπτώσεις απάτης. Στο πλαίσιο της συγκεκριμένης ανάλυσης συμπληρωματικότητας, οι πραγματικές ετικέτες χρησιμοποιήθηκαν εκ των υστέρων για την αξιολόγηση των ανιχνεύσεων και τη σύγκριση της συμπεριφοράς των μοντέλων. Οι ετικέτες δεν χρησιμοποιήθηκαν ως στόχος επιβλεπόμενης μάθησης για την εκπαίδευση των βασικών μοντέλων, ενώ η ειδική χρήση τους για τον σχηματισμό του συνόλου εκπαίδευσης μόνο με κανονικές παρατηρήσεις στον τελικό Autoencoder περιγράφεται στην προηγούμενη ενότητα.

Η συμπληρωματικότητα διερευνήθηκε αρχικά στο επίπεδο των συνεχών βαθμολογιών ανωμαλίας. Για κάθε ζεύγος μοντέλων υπολογίστηκε ο συντελεστής συσχέτισης Pearson [54]. Πριν από τη σύγκριση εξασφαλίστηκε κοινή κατεύθυνση των βαθμολογιών, έτσι ώστε υψηλότερη τιμή να αντιστοιχεί σε μεγαλύτερο βαθμό ανωμαλίας για όλα τα μοντέλα. Ο Pearson χρησιμοποιήθηκε για να εξεταστεί σε ποιο βαθμό δύο ανιχνευτές κατατάσσουν τις ίδιες συναλλαγές με γραμμικά παρόμοιο τρόπο. Υψηλή θετική συσχέτιση ερμηνεύεται ως ένδειξη παρόμοιας συμπεριφοράς των συνεχών βαθμολογιών, ενώ χαμηλότερη συσχέτιση υποδηλώνει μεγαλύτερη διαφοροποίηση στη βαθμολόγηση των συναλλαγών. Σε δεύτερο επίπεδο εξετάστηκε η επικάλυψη των πραγματικών θετικών ανιχνεύσεων. Για κάθε μοντέλο δημιουργήθηκε το σύνολο των συναλλαγών με $isFraud=1$ που χαρακτηρίστηκαν σωστά ως ανωμαλίες. Για δύο μοντέλα A και B, η επικάλυψη υπολογίστηκε μέσω του δείκτη ομοιότητας Jaccard [55]:

$$J(A, B) = \frac{|TP_A \cap TP_B|}{|TP_A \cup TP_B|} \quad (3.6)$$

όπου TP_A και TP_B είναι τα σύνολα των Αληθώς Θετικών Περιπτώσεων (True Positives-TP) που εντοπίζονται από τα δύο μοντέλα. Παράλληλα, καταγράφηκαν οι κοινές πραγματικές θετικές ανιχνεύσεις και οι περιπτώσεις που εντοπίζονταν αποκλειστικά από καθένα από τα δύο μοντέλα. Με αυτόν τον τρόπο η συμπληρωματικότητα δεν εξετάστηκε μόνο ως συνολική ομοιότητα, αλλά και ως δυνατότητα ενός ανιχνευτή να εντοπίζει περιπτώσεις απάτης που δεν εντοπίζονται από έναν δεύτερο.

Ως τρίτο επίπεδο σύγκρισης εφαρμόστηκε ο έλεγχος McNemar[56]. Ο έλεγχος πραγματοποιήθηκε αποκλειστικά στις 3.499 πραγματικές περιπτώσεις απάτης και επομένως αφορά τη διαφοροποίηση των μοντέλων ως προς την ανίχνευση της θετικής κλάσης και όχι τη συνολική τους ταξινομητική συμπεριφορά στις 100.000 συναλλαγές. Για κάθε ζεύγος μοντέλων, ως b ορίστηκε ο αριθμός των πραγματικών περιπτώσεων απάτης που εντοπίστηκαν αποκλειστικά από το πρώτο μοντέλο και ως c ο αριθμός των περιπτώσεων που εντοπίστηκαν αποκλειστικά από το δεύτερο. Χρησιμοποιήθηκε η εκδοχή του McNemar με διόρθωση συνέχειας:

$$X^2 = \frac{(|b-c|-1)^2}{b+c} \quad (3.7)$$

με έναν βαθμό ελευθερίας.

Καθώς τα τέσσερα μοντέλα δημιουργούν συνολικά έξι διαφορετικές ζευγαρωτές συγκρίσεις, εφαρμόστηκε η διόρθωση Holm στα αντίστοιχα p-values [57]. Η διόρθωση χρησιμοποιήθηκε ώστε να περιοριστεί η αύξηση του σφάλματος Τύπου I που προκύπτει από την ταυτόχρονη εκτέλεση πολλαπλών στατιστικών ελέγχων. Ως επίπεδο στατιστικής σημαντικότητας χρησιμοποιήθηκε $\alpha=0,05$. Τα διορθωμένα p-values παρουσιάζονται μαζί με τις αποκλειστικές ανιχνεύσεις στο Κεφάλαιο 4.

Μετά την αξιολόγηση των τεχνικών συνδυασμού πραγματοποιήθηκε επιπλέον post-hoc ανάλυση συνεισφοράς μέσω Shapley Values. Η ανάλυση αυτή δεν χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων ούτε για τον καθορισμό των βαρών του Weighted Score-Based Ensemble. Στόχος της ήταν να εξεταστεί εκ των υστέρων η μέση οριακή συνεισφορά κάθε ανιχνευτή όταν αυτός συμμετέχει σε διαφορετικούς συνασπισμούς.

Ως σύνολο παικτών ορίστηκαν τα τέσσερα μοντέλα:

$$N = \{IF, LOF, HBOS, AE\} \quad (3.8)$$

Με τέσσερις παίκτες προκύπτουν $2^4 = 16$ δυνατοί συνασπισμοί, συμπεριλαμβανομένου του κενού συνόλου. Επειδή ο αριθμός των παικτών είναι μικρός, ήταν δυνατή η ακριβής αξιολόγηση όλων των συνασπισμών χωρίς να απαιτηθεί δειγματοληπτική προσέγγιση τον υπολογισμό των τιμών Shapley.

Πριν από τη δημιουργία των συνασπισμών, οι συνεχείς βαθμολογίες ανωμαλιών κάθε μοντέλου κανονικοποιήθηκαν ανεξάρτητα στο διάστημα $[0,1]$ μέσω Min-Max normalization. Για κάθε μη κενό συνασπισμό S , η κοινή βαθμολογία μιας συναλλαγής υπολογίστηκε ως ο ισότιμος αριθμητικός μέσος των κανονικοποιημένων βαθμολογιών ανωμαλίας των μοντέλων που συμμετέχουν στον συγκεκριμένο συνασπισμό:

$$Score_S(x) = \frac{1}{|S|} \sum_{i \in S} \tilde{S}_i(x) \quad (3.9)$$

Με αυτόν τον τρόπο κανένα μοντέλο δεν λαμβάνει εκ των προτέρων μεγαλύτερο βάρος μέσα στην ανάλυση Shapley.

Η απόδοση κάθε συνασπισμού αξιολογήθηκε μέσω της Average Precision (AP), η οποία υπολογίστηκε με βάση τις πραγματικές ετικέτες του IEEE-CIS. Η αναλογία της θετικής κλάσης στο κοινό δείγμα ήταν:

$$\pi = \frac{3499}{100000} = 0,03499 \quad (3.10)$$

και χρησιμοποιήθηκε ως τιμή αναφοράς μιας τυχαίας κατάταξης. Η χαρακτηριστική συνάρτηση του συνεργατικού παιγνίου ορίστηκε ως $u(S)=AP(S)-\pi$ για κάθε μη κενό συνασπισμό, ενώ $u(\emptyset)=0$. Με την επιλογή αυτή, η αξία ενός συνασπισμού εκφράζει το κέρδος της AUPRC/AP σε σχέση με τη βασική αναλογία της θετικής κλάσης. Στη συνέχεια, η συνεισφορά κάθε μοντέλου υπολογίστηκε μέσω της τιμής Shapley [58]:

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} [u(S \cup \{i\}) - u(S)] \quad (3.11)$$

Η τιμή φ_i εκφράζει επομένως τη μέση οριακή μεταβολή της αξίας των συνασπισμών όταν προστίθεται το μοντέλο i , λαμβάνοντας υπόψη όλες τις δυνατές συνθέσεις των υπόλοιπων ανιχνευτών. Θετική τιμή Shapley υποδηλώνει θετική μέση οριακή συνεισφορά σύμφωνα με τη συγκεκριμένη συνάρτηση αξίας, ενώ είναι δυνατό να προκύψει και αρνητική τιμή όταν η προσθήκη ενός μοντέλου μειώνει κατά μέσο όρο την AP των συνασπισμών.

Η συγκεκριμένη ανάλυση χρησιμοποιείται αποκλειστικά ως μηχανισμός ερμηνείας της συνεισφοράς των τεσσάρων ανιχνευτών στο ελεγχόμενο σύνολο δεδομένων αναφοράς IEEE-CIS. Επειδή η συνάρτηση αξίας βασίζεται στην AP, η διαδικασία απαιτεί τη χρήση των πραγματικών ετικετών και επομένως δεν εφαρμόζεται με τον ίδιο τρόπο στα μη επισημασμένα ERP δεδομένα. Επιπλέον, οι τιμές Shapley δεν μετατράπηκαν σε βάρη και δεν χρησιμοποιήθηκαν για εκ των υστέρων βελτιστοποίηση της τελικής τεχνικής συνδυασμού. Τα αποτελέσματα της συγκεκριμένης ανάλυσης παρουσιάζονται στην Ενότητα 5.5.

3.4 Μετρικές Αξιολόγησης

Η αξιολόγηση της απόδοσης των μοντέλων αποτελεί κρίσιμο στάδιο της πειραματικής διαδικασίας, ιδιαίτερα σε προβλήματα οικονομικής απάτης όπου η κατανομή των κατηγοριών είναι έντονα μη ισορροπημένη. Στο IEEE-CIS, οι συναλλαγές που έχουν επισημανθεί ως περιπτώσεις απάτης αντιστοιχούν περίπου στο 3,5% του συνολικού συνόλου δεδομένων. Συνεπώς, η χρήση μιας μετρικής όπως η συνολική ακρίβεια ταξινόμησης δεν επαρκεί για την αποτίμηση της πραγματικής ικανότητας ενός μοντέλου να εντοπίζει τη μειοψηφική κατηγορία [52].

Για τον λόγο αυτό, η αξιολόγηση στην παρούσα εργασία βασίστηκε στις μετρικές Precision, Recall, F1-score, ROC-AUC και AUPRC, καθώς και στην εξέταση του Πίνακα Σύγχυσης. Οι συγκεκριμένες μετρικές παρέχουν συμπληρωματική πληροφορία σχετικά με την ικανότητα των μοντέλων να εντοπίζουν πραγματικές περιπτώσεις απάτης, το πλήθος των ψευδών συναγεργμών και τη συνολική ποιότητα κατάταξης των συναλλαγών.

3.4.1 Precision

Η μετρική Precision εκφράζει το ποσοστό των συναλλαγών που χαρακτηρίστηκαν από το μοντέλο ως ύποπτες και οι οποίες αντιστοιχούν πράγματι σε περιπτώσεις απάτης [52]. Ορίζεται ως:

$$Precision = \frac{TP}{TP+FP} \quad (3.12)$$

όπου:

TP είναι οι πραγματικά θετικές περιπτώσεις (True Positives), δηλαδή συναλλαγές απάτης που εντοπίστηκαν σωστά,

FP είναι οι ψευδώς θετικές περιπτώσεις (False Positives), δηλαδή νόμιμες συναλλαγές που επισημάνθηκαν λανθασμένα ως ύποπτες.

Υψηλή τιμή Precision σημαίνει ότι μεγάλο μέρος των ειδοποιήσεων που παράγει το μοντέλο αντιστοιχεί σε πραγματικές περιπτώσεις της θετικής κατηγορίας. Στο πλαίσιο της οικονομικής απάτης, η συγκεκριμένη ιδιότητα είναι σημαντική επειδή ένας μεγάλος αριθμός ψευδών θετικών αποτελεσμάτων αυξάνει τον φόρτο του ανθρώπινου ελέγχου και μπορεί να περιορίσει την πρακτική χρησιμότητα ενός συστήματος ανίχνευσης.

Ωστόσο, η Precision δεν πρέπει να εξετάζεται ανεξάρτητα. Ένα πολύ αυστηρό λειτουργικό κατώφλι μπορεί να αυξήσει την Precision, επιτρέποντας στο μοντέλο να επισημαίνει μόνο τις πλέον ακραίες περιπτώσεις, αλλά ταυτόχρονα να οδηγήσει σε απώλεια σημαντικού αριθμού πραγματικών περιπτώσεων απάτης. Επομένως, στην παρούσα εργασία η Precision εξετάζεται σε συνδυασμό κυρίως με το Recall.

3.4.2 Recall

Η μετρική Recall, η οποία αναφέρεται επίσης ως ευαισθησία ή True Positive Rate, εκφράζει το ποσοστό των πραγματικών περιπτώσεων απάτης που εντοπίστηκαν επιτυχώς από το μοντέλο [52]. Υπολογίζεται ως:

$$Recall = \frac{TP}{TP+FN} \quad (3.13)$$

όπου (FN) είναι οι ψευδώς αρνητικές περιπτώσεις (False Negatives), δηλαδή πραγματικές περιπτώσεις απάτης οι οποίες δεν επισημάνθηκαν από το μοντέλο.

Η συγκεκριμένη μετρική έχει ιδιαίτερη σημασία στην παρούσα εργασία. Σε ένα σύστημα ανίχνευσης οικονομικών ανωμαλιών, μία ψευδώς αρνητική περίπτωση αντιπροσωπεύει μία πραγματική περίπτωση απάτης που δεν προωθήθηκε για περαιτέρω διερεύνηση. Κατά συνέπεια, χαμηλό Recall μπορεί να περιορίσει σημαντικά τη χρησιμότητα του συστήματος ως μηχανισμού αρχικού ελέγχου.

Από την άλλη πλευρά, η αύξηση του Recall δεν αποτελεί από μόνη της επαρκή στόχο. Η μείωση του λειτουργικού κατωφλίου οδηγεί συνήθως στην επισήμανση περισσότερων συναλλαγών, αυξάνοντας την πιθανότητα εντοπισμού πραγματικών περιπτώσεων απάτης αλλά ταυτόχρονα και το πλήθος των ψευδών θετικών αποτελεσμάτων. Πρόκειται επομένως για έναν συμβιβασμό μεταξύ Recall και Precision, ο οποίος εξετάστηκε συστηματικά κατά την επιλογή των λειτουργικών σημείων των μοντέλων.

Στην παρούσα εργασία η αξιολόγηση δεν βασίστηκε σε μία μεμονωμένη μετρική, αλλά στη συνολική συνεκτίμηση των Precision, Recall, F1-score και AUPRC/AP. Ιδιαίτερη σημασία αποδόθηκε στο Recall, λόγω της ανάγκης περιορισμού των μη εντοπισμένων περιπτώσεων απάτης, χωρίς όμως να παραβλέπεται ο αντίστοιχος συμβιβασμός με την Precision και η συνολική ποιότητα κατάταξης που αποτυπώνεται από την AUPRC/AP. Υπό αυτή την έννοια, η αποτυχία ανάδειξης μιας πραγματικά ύποπτης συναλλαγής μπορεί να είναι ιδιαίτερα σημαντική.

3.4.3 F1-score

Η μετρική F1-score συνδυάζει τις μετρικές Precision και Recall σε μία ενιαία τιμή και υπολογίζεται ως ο αρμονικός μέσος των δύο μετρικών [52]:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.14)$$

Σε αντίθεση με έναν απλό αριθμητικό μέσο, ο αρμονικός μέσος επηρεάζεται περισσότερο όταν μία από τις δύο επιμέρους μετρικές έχει χαμηλή τιμή. Επομένως, ένα υψηλό F1-score προϋποθέτει σχετικά ισορροπημένη συμπεριφορά μεταξύ Precision και Recall.

Η μετρική είναι ιδιαίτερα χρήσιμη όταν δεν είναι επιθυμητή η αποκλειστική βελτιστοποίηση ούτε της Precision ούτε του Recall. Για παράδειγμα, ένα μοντέλο μπορεί να επιτυγχάνει πολύ υψηλό Recall επειδή χαρακτηρίζει μεγάλο αριθμό συναλλαγών ως ύποπτες, αλλά να παρουσιάζει ταυτόχρονα πολύ

χαμηλή Precision λόγω των πολλών ψευδών θετικών περιπτώσεων. Στην περίπτωση αυτή, το F1-score αποτυπώνει την έλλειψη ισορροπίας μεταξύ των δύο μετρικών.

Στην παρούσα εργασία το F1-score χρησιμοποιείται ως συμπληρωματικό κριτήριο κατά τη σύγκριση των διαφορετικών λειτουργικών σημείων των μοντέλων. Δεν αποτελεί όμως τη μοναδική βάση επιλογής, καθώς η πρακτική σημασία του Recall και η έντονη ανισορροπία των κατηγοριών απαιτούν την παράλληλη εξέταση και άλλων μετρικών.

3.4.4 ROC-AUC και AUPRC

Οι μετρικές ROC-AUC και AUPRC επιτρέπουν την αξιολόγηση της ποιότητας των συνεχών βαθμολογιών που παράγουν τα μοντέλα σε ένα εύρος πιθανών κατωφλίων και, συνεπώς, παρέχουν διαφορετική πληροφορία από τις Precision, Recall και F1-score που υπολογίζονται σε ένα συγκεκριμένο λειτουργικό σημείο.

Η καμπύλη Receiver Operating Characteristic (ROC) απεικονίζει τη σχέση μεταξύ του True Positive Rate και του False Positive Rate καθώς μεταβάλλεται το κατώφλι απόφασης. Το True Positive Rate ταυτίζεται με το Recall:

$$TPR = \frac{TP}{TP+FN} \quad (3.15)$$

ενώ το False Positive Rate ορίζεται ως:

$$FPR = \frac{FP}{FP+TN} \quad (3.16)$$

Η περιοχή κάτω από την καμπύλη ROC, δηλαδή η ROC-AUC, συνοψίζει την ικανότητα του μοντέλου να κατατάσσει θετικές περιπτώσεις υψηλότερα από αρνητικές σε διαφορετικά κατώφλια. Τιμές που προσεγγίζουν το 1 υποδηλώνουν ισχυρότερη διαχωριστική ικανότητα, ενώ τιμή περίπου 0,5 αντιστοιχεί σε συμπεριφορά που δεν διαφοροποιείται ουσιαστικά από μία τυχαία κατάταξη.

Παρότι η ROC-AUC αποτελεί ευρέως χρησιμοποιούμενη μετρική, απαιτεί προσεκτική ερμηνεία όταν η θετική κατηγορία αποτελεί πολύ μικρό μέρος του συνόλου. Σε έντονα μη ισορροπημένα δεδομένα, ο πολύ μεγάλος αριθμός True Negatives αυξάνει σημαντικά τον παρονομαστή του False Positive Rate. Ως αποτέλεσμα, ακόμη και ένας σημαντικός αριθμός False Positives μπορεί να αντιστοιχεί σε σχετικά μικρό FPR και η ROC-AUC να εμφανίζει περισσότερο αισιόδοξη εικόνα της πρακτικής απόδοσης του μοντέλου [14], [53].

Η καμπύλη Precision–Recall, αντίθετα, παρουσιάζει άμεσα τη σχέση μεταξύ Precision και Recall καθώς μεταβάλλεται το κατώφλι απόφασης. Η περιοχή κάτω από αυτή την καμπύλη AUPRC, εστιάζει επομένως περισσότερο στη συμπεριφορά του μοντέλου ως προς τη θετική και σπάνια κατηγορία.

Η συγκεκριμένη ιδιότητα είναι ιδιαίτερα σημαντική στο πρόβλημα της οικονομικής απάτης. Επειδή ο αριθμός των πραγματικών περιπτώσεων απάτης είναι μικρός σε σχέση με τις νόμιμες συναλλαγές, η Precision επηρεάζεται άμεσα όταν αυξάνονται τα False Positives. Η AUPRC μπορεί επομένως να αποκαλύψει υποβάθμιση της ποιότητας των ύποπτων ειδοποιήσεων η οποία ενδέχεται να είναι λιγότερο εμφανής μέσω της ROC-AUC [14].

Η χρησιμότητα της Precision–Recall σε μη ισορροπημένα προβλήματα έχει επίσης υποστηριχθεί σε συγκριτικές μελέτες, όπου η ROC μπορεί να παρέχει υπερβολικά αισιόδοξη εικόνα όταν η πλειοψηφική κατηγορία κυριαρχεί αριθμητικά, ενώ η Precision–Recall παρέχει περισσότερο άμεση εικόνα της συμπεριφοράς στη μειοψηφική κατηγορία [53]. Αντίστοιχα, η χρήση της περιοχής κάτω από την Precision–Recall έχει προταθεί ως κατάλληλο κριτήριο για προβλήματα μη ισορροπημένης ταξινόμησης, στα οποία περιλαμβάνεται και η ανίχνευση οικονομικής απάτης [52].

Για το IEEE-CIS το συγκεκριμένο ζήτημα είναι ιδιαίτερα σημαντικό, καθώς μόνο περίπου το 3,5% των συναλλαγών ανήκει στην κατηγορία απάτης. Για τον λόγο αυτό, η AUPRC χρησιμοποιείται στην παρούσα εργασία ως βασικότερο κριτήριο συνολικής κατάταξης των μοντέλων σε σχέση με τη ROC-AUC, χωρίς όμως η δεύτερη να απορρίπτεται. Οι δύο μετρικές αναφέρονται παράλληλα, αλλά η AUPRC λαμβάνει μεγαλύτερη βαρύτητα κατά την ερμηνεία της απόδοσης.

Για λόγους συνέπειας μεταξύ των διαφορετικών πειραμάτων, η αριθμητική αποτίμηση της Precision–Recall στην τελική αξιολόγηση πραγματοποιείται μέσω της Average Precision (AP), η οποία υπολογίζεται απευθείας από τις βαθμολογίες ανωμαλίας και τις πραγματικές ετικέτες. Στους πίνακες και στην ανάλυση των αποτελεσμάτων η συγκεκριμένη ποσότητα αναφέρεται ως AUPRC, ώστε να διατηρείται ενιαία ορολογία σε όλη την εργασία.

Συνεπώς, μία υψηλότερη ROC-AUC δεν θεωρείται αυτομάτως ένδειξη καλύτερου μοντέλου για το συγκεκριμένο πρόβλημα. Η τελική αξιολόγηση πραγματοποιείται εξετάζοντας κυρίως την AUPRC σε συνδυασμό με Recall, Precision και F1-score στο επιλεγμένο λειτουργικό σημείο.

3.4.5 Πίνακας Σύγχυσης

Ο Πίνακας Σύγχυσης (Confusion Matrix) παρέχει αναλυτική εικόνα των δυαδικών αποφάσεων ενός μοντέλου, καταγράφοντας τις προβλέψεις σε σχέση με τις πραγματικές κατηγορίες [52]. Για το πρόβλημα της παρούσας εργασίας οι τέσσερις δυνατές περιπτώσεις είναι:

	Προβλεπόμενη απάτη	Προβλεπόμενη νόμιμη συναλλαγή
Πραγματική Απάτη	TP	FN
Πραγματικά Νόμιμη	FP	TN

Οι τέσσερις κατηγορίες ερμηνεύονται ως εξής:

- True Positive (TP): μία πραγματική περίπτωση απάτης επισημαίνεται σωστά από το μοντέλο ως ύποπτη.
- False Positive (FP): μία νόμιμη συναλλαγή επισημαίνεται λανθασμένα ως ύποπτη. Η περίπτωση αυτή δημιουργεί έναν ψευδή συναγερμό και αυξάνει τον φόρτο περαιτέρω ελέγχου.
- False Negative (FN): μία πραγματική περίπτωση απάτης δεν εντοπίζεται από το μοντέλο. Η συγκεκριμένη κατηγορία είναι ιδιαίτερα σημαντική στην παρούσα εφαρμογή, καθώς αντιπροσωπεύει ύποπτες συναλλαγές που διαφεύγουν του μηχανισμού ανίχνευσης.
- True Negative (TN): μία νόμιμη συναλλαγή αναγνωρίζεται σωστά ως μη ύποπτη.

Ο Πίνακας Σύγχυσης δεν αποτελεί απλώς μία πρόσθετη παρουσίαση των αποτελεσμάτων, αλλά τη βάση για τον υπολογισμό των Precision, Recall και F1-score. Επιτρέπει επίσης την άμεση εξέταση του αριθμού των False Positives και False Negatives, στοιχείο ιδιαίτερα χρήσιμο για την επιχειρησιακή ερμηνεία ενός συστήματος ανίχνευσης απάτης.

Στην παρούσα εργασία ο Πίνακας Σύγχυσης χρησιμοποιείται για την αξιολόγηση των μοντέλων στο IEEE-CIS, όπου είναι διαθέσιμες οι πραγματικές ετικέτες. Αντίθετα, στο πραγματικό σύνολο ERP δεν μπορεί να υπολογιστεί, καθώς δεν υπάρχουν πραγματικές ετικέτες που να επιτρέπουν τον διαχωρισμό των εγγραφών σε TP, FP, FN και TN. Στην περίπτωση αυτή η αξιολόγηση μετατοπίζεται στην ανάλυση των βαθμολογιών κινδύνου, στην ιεράρχηση των εγγραφών και στη συμπληρωματική στατιστική ανάλυση που παρουσιάζεται στο Κεφάλαιο 6.

Κεφάλαιο 4ο: Αξιολόγηση και Συγκριτική Ανάλυση των Μοντέλων Ανίχνευσης Ανωμαλιών

4.1 Πειραματικό Περιβάλλον

Η πειραματική αξιολόγηση των μοντέλων πραγματοποιήθηκε σε περιβάλλον Google Colab, με χρήση της γλώσσας προγραμματισμού Python. Η συγκεκριμένη πλατφόρμα χρησιμοποιήθηκε για την εκτέλεση των διαδικασιών προεπεξεργασίας, εκπαίδευσης, παραμετροποίησης και αξιολόγησης των μοντέλων, καθώς και για την παραγωγή των πινάκων και των γραφικών απεικονίσεων που παρουσιάζονται στο παρόν κεφάλαιο. Η αποθήκευση των ενδιάμεσων και τελικών αποτελεσμάτων πραγματοποιήθηκε στο Google Drive, ώστε τα παραγόμενα αρχεία να μπορούν να επαναχρησιμοποιηθούν χωρίς να απαιτείται εκ νέου εκπαίδευση των μοντέλων σε κάθε συνεδρία εκτέλεσης.

Για την επεξεργασία και ανάλυση των δεδομένων χρησιμοποιήθηκαν καθιερωμένες βιβλιοθήκες του οικοσυστήματος της Python. Οι βιβλιοθήκες NumPy και pandas αξιοποιήθηκαν για τη διαχείριση των δεδομένων και των αριθμητικών πινάκων, ενώ εργαλεία της scikit-learn χρησιμοποιήθηκαν για διαδικασίες προεπεξεργασίας, υπολογισμό μετρικών αξιολόγησης και υλοποίηση επιμέρους αλγορίθμων. Για το Histogram-Based Outlier Score αξιοποιήθηκε αντίστοιχη υλοποίηση βιβλιοθήκης ανίχνευσης ανωμαλιών, ενώ η ανάπτυξη του Autoencoder πραγματοποιήθηκε μέσω του οικοσυστήματος TensorFlow/Keras. Η οπτικοποίηση των αποτελεσμάτων πραγματοποιήθηκε με βιβλιοθήκες γραφικής απεικόνισης της Python, μέσω των οποίων δημιουργήθηκαν διαγράμματα απόδοσης, πίνακες συσχέτισης και απεικονίσεις επικάλυψης των ανιχνεύσεων.

Για τη συγκριτική αξιολόγηση των τεσσάρων μοντέλων χρησιμοποιήθηκε κοινό στρωματοποιημένο δείγμα 100.000 συναλλαγών από το σύνολο δεδομένων IEEE-CIS Fraud Detection. Το δείγμα περιλαμβάνει 96.501 νόμιμες συναλλαγές και 3.499 περιπτώσεις απάτης, διατηρώντας την έντονη ανισορροπία κλάσεων του αρχικού προβλήματος. Η χρήση του ίδιου ακριβώς δείγματος για τα Isolation Forest, Local Outlier Factor, HBOS και Autoencoder επιτρέπει την άμεση σύγκριση των αποτελεσμάτων τους πάνω στις ίδιες παρατηρήσεις και περιορίζει διαφοροποιήσεις που θα μπορούσαν να προκύψουν από διαφορετική δειγματοληψία.

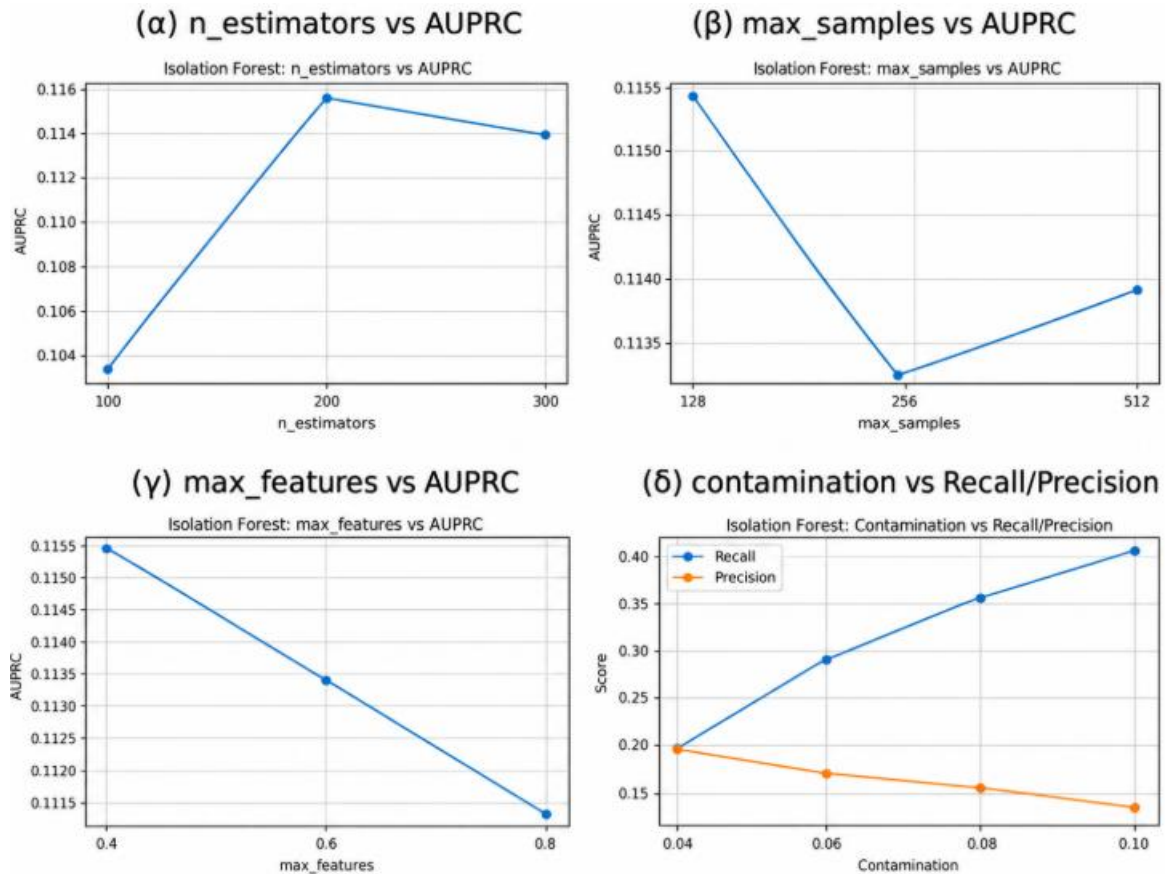
Τα δεδομένα που χρησιμοποιήθηκαν από τα μοντέλα είχαν προηγουμένως υποβληθεί στις διαδικασίες προεπεξεργασίας που περιγράφηκαν στο Κεφάλαιο 3. Για λόγους συνέπειας, το ίδιο επεξεργασμένο σύνολο χαρακτηριστικών χρησιμοποιήθηκε στις συγκριτικές αναλύσεις, ενώ οι ετικέτες της μεταβλητής isFraud δεν χρησιμοποιήθηκαν ως χαρακτηριστικό εισόδου κατά την εφαρμογή των μοντέλων ανίχνευσης ανωμαλιών. Αντίθετα, αξιοποιήθηκαν εκ των υστέρων για τον υπολογισμό των μετρικών αξιολόγησης και τη σύγκριση των παραγόμενων ανιχνεύσεων με τις πραγματικές περιπτώσεις απάτης. Για κάθε μοντέλο πραγματοποιήθηκε αρχικά διερεύνηση διαφορετικών συνδυασμών υπερπαραμέτρων, σύμφωνα με το πειραματικό πρωτόκολλο που παρουσιάστηκε στην Ενότητα 3.3. Στη συνέχεια επιλέχθηκε μία τελική διαμόρφωση ανά μοντέλο και αποθηκεύτηκαν τόσο οι δυαδικές προβλέψεις ανωμαλίας όσο και οι αντίστοιχες συνεχείς βαθμολογίες ανωμαλίας. Η διαδικασία αυτή επέτρεψε όχι μόνο τη σύγκριση των μοντέλων βάσει των Precision, Recall, F1-score, ROC-AUC και AUPRC, αλλά και τη μεταγενέστερη διερεύνηση του βαθμού επικάλυψης και συμπληρωματικότητας μεταξύ των διαφορετικών μηχανισμών ανίχνευσης.

Με βάση το παραπάνω πειραματικό περιβάλλον, στις επόμενες ενότητες παρουσιάζονται διαδοχικά τα αποτελέσματα παραμετροποίησης και αξιολόγησης των Isolation Forest, LOF, HBOS και Autoencoder, πριν ακολουθήσει η συνολική συγκριτική τους ανάλυση.

4.2 Isolation Forest

Για την παραμετροποίηση του Isolation Forest εξετάστηκαν διαφορετικές τιμές για τις υπερπαραμέτρους `n_estimators`, `max_samples`, `max_features` και `contamination`. Συγκεκριμένα, το `n_estimators` έλαβε τις τιμές 100, 200 και 300, το `max_samples` τις τιμές 128, 256 και 512, το `max_features` τις τιμές 0,4, 0,6 και 0,8, ενώ για το `contamination` εξετάστηκαν οι τιμές 0,035, 0,06, 0,08 και 0,10. Από τον συνδυασμό τους προέκυψαν συνολικά 108 διαφορετικές διαμορφώσεις, οι οποίες παρουσιάζονται αναλυτικά στο Παράρτημα Α.

Η επίδραση των βασικών υπερπαραμέτρων στην πειραματική συμπεριφορά του μοντέλου παρουσιάζεται στο Σχήμα 4.1. Ως προς το `n_estimators`, η μετάβαση από 100 σε 200 δέντρα συνοδεύτηκε από αύξηση της AUPRC, ενώ στα 300 δέντρα παρατηρήθηκε μικρή υποχώρηση. Η εικόνα αυτή υποδηλώνει ότι, για το συγκεκριμένο σύνολο δεδομένων και τις εξεταζόμενες διαμορφώσεις, η αύξηση του αριθμού των δέντρων πέρα από ένα σημείο δεν συνοδεύτηκε από περαιτέρω βελτίωση της συγκεκριμένης μετρικής.



Σχήμα 4.1. Επίδραση των υπερπαραμέτρων του Isolation Forest στην AUPRC και στη σχέση Recall–Precision.

Αντίστοιχα, το `max_samples`=128 συνδέθηκε με την υψηλότερη AUPRC μεταξύ των εξεταζόμενων τιμών, ενώ οι τιμές 256 και 512 παρουσίασαν ελαφρώς χαμηλότερα αποτελέσματα. Ως προς το `max_features`, η υψηλότερη τιμή AUPRC εμφανίστηκε όταν χρησιμοποιήθηκε το 40% των χαρακτηριστικών, ενώ παρατηρήθηκε σταδιακή μείωση για τις τιμές 0,6 και 0,8. Το συγκεκριμένο αποτέλεσμα φαίνεται να υποστηρίζει ότι, στο συγκεκριμένο πειραματικό πλαίσιο, η χρήση μικρότερου υποσυνόλου χαρακτηριστικών ενδέχεται να ευνοεί τη διαφοροποίηση μεταξύ των επιμέρους δέντρων. Η πιο σαφής μεταβολή παρατηρήθηκε στην παράμετρο `contamination`. Η αύξησή της από 0,035 σε 0,10 συνοδεύτηκε από σταδιακή αύξηση του Recall και ταυτόχρονη μείωση του Precision. Η συμπεριφορά αυτή είναι συμβατή με τον τρόπο με τον οποίο μεταβάλλεται το σημείο λειτουργίας του μοντέλου: όταν μεγαλύτερο ποσοστό των παρατηρήσεων χαρακτηρίζεται ως ανώμαλο, αυξάνεται η πιθανότητα εντοπισμού πραγματικών περιπτώσεων απάτης, αλλά παράλληλα αυξάνεται και ο αριθμός των ψευδών θετικών επισημάνσεων (FP). Για τον λόγο αυτό, το `contamination` αντιμετωπίζεται στην παρούσα εργασία ως παράμετρος ευαισθησίας του μηχανισμού ανίχνευσης και όχι ως εκτίμηση του πραγματικού ποσοστού απάτης.

Η τελική διαμόρφωση που χρησιμοποιήθηκε στην κοινή αξιολόγηση των μοντέλων ήταν:

- `n_estimators` = 200
- `max_samples` = 128
- `max_features` = 0,4
- `contamination` = 0,10

Στο κοινό δείγμα των 100.000 συναλλαγών, το Isolation Forest επισήμανε 10.000 συναλλαγές ως ανωμαλίες και εντόπισε 1.420 από τις 3.499 πραγματικές περιπτώσεις απάτης. Το Precision διαμορφώθηκε σε 0,1420, το Recall σε 0,4058 και το F1-score σε περίπου 0,2104.

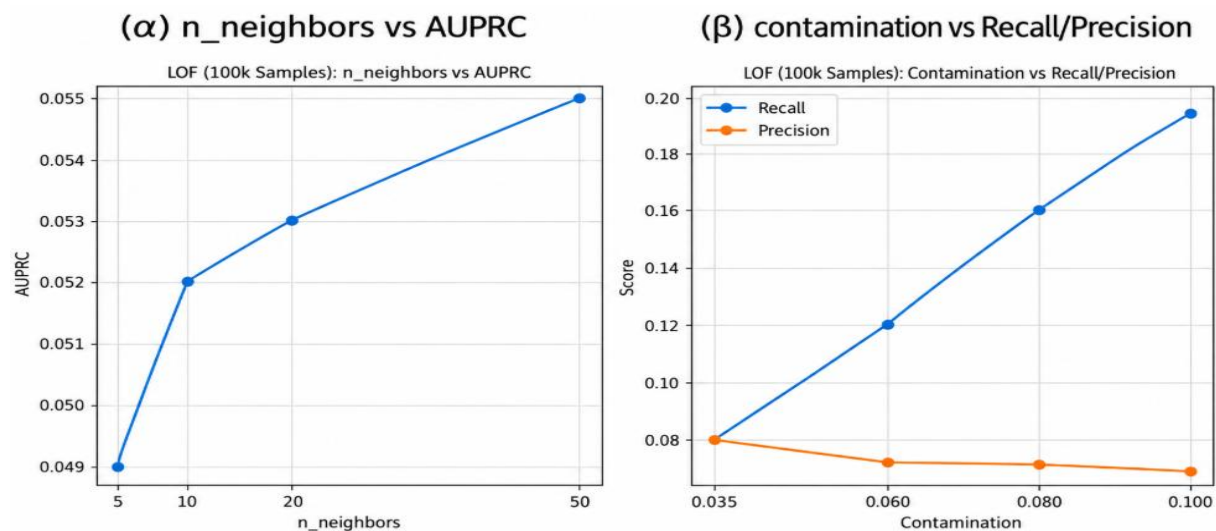
Τα αποτελέσματα υποδηλώνουν ότι το Isolation Forest παρουσίασε σχετικά αυξημένη ικανότητα ανάκτησης της θετικής κλάσης, καθώς εντόπισε περίπου το 40,6% των πραγματικών περιπτώσεων απάτης. Η επίδοση αυτή συνοδεύτηκε, ωστόσο, από περιορισμένο Precision, γεγονός που σημαίνει ότι σημαντικό μέρος των συναλλαγών που επισημάνθηκαν ως ανώμαλες δεν αντιστοιχούσε σε καταγεγραμμένη περίπτωση απάτης.

Η συγκεκριμένη σχέση μεταξύ Recall και Precision δεν επιτρέπει την αξιολόγηση του μοντέλου αποκλειστικά βάσει μίας μετρικής. Το αυξημένο Recall μπορεί να θεωρηθεί χρήσιμο σε ένα πλαίσιο όπου δίνεται προτεραιότητα στον εντοπισμό μεγαλύτερου μέρους των ύποπτων συναλλαγών, συνεπάγεται όμως μεγαλύτερο αριθμό περιπτώσεων που θα απαιτούσαν μεταγενέστερο έλεγχο. Επομένως, η συνολική αποτίμηση του Isolation Forest απαιτεί τη συνεκτίμηση τόσο της ικανότητας ανίχνευσης όσο και του αντίστοιχου αριθμού ψευδώς θετικών επισημάνσεων (false positives-FP).

4.3 Local Outlier Factor

Για το Local Outlier Factor εξετάστηκε η επίδραση των υπερπαραμέτρων $n_neighbors$ και $contamination$. Το $n_neighbors$ έλαβε τις τιμές 5, 10, 20 και 50, ενώ το $contamination$ εξετάστηκε στις τιμές 0,035, 0,06, 0,08 και 0,10. Συνολικά προέκυψαν 16 διαφορετικές διαμορφώσεις, οι οποίες παρατίθενται αναλυτικά στο Παράρτημα Α.

Η AUPRC παρουσίασε ανοδική τάση καθώς αυξανόταν ο αριθμός των γειτονικών παρατηρήσεων. Η χαμηλότερη τιμή καταγράφηκε για $n_neighbors=5$, ενώ η υψηλότερη μεταξύ των εξεταζόμενων τιμών εμφανίστηκε για $n_neighbors=50$. Η παρατήρηση αυτή υποδηλώνει ότι, στο συγκεκριμένο σύνολο δεδομένων, η εκτίμηση της τοπικής πυκνότητας σε μεγαλύτερη γειτονιά φαίνεται να συνδέεται με αποτελεσματικότερη διάκριση των παρατηρήσεων ως προς τη βαθμολογία ανωμαλίας.



Σχήμα 4.2. Επίδραση του αριθμού των γειτόνων και του $contamination$ στην απόδοση του LOF.

Η μεταβολή του $contamination$ επηρέασε κυρίως τη σχέση Recall–Precision. Με την αύξηση της παραμέτρου παρατηρήθηκε σταδιακή αύξηση του Recall, ενώ το Precision παρουσίασε μικρή αλλά συνεχή μείωση. Παρά την αύξηση του Recall, η τελική τιμή του παρέμεινε χαμηλότερη σε σχέση με εκείνες που παρατηρήθηκαν σε ορισμένα από τα υπόλοιπα μοντέλα.

Η τελική διαμόρφωση που χρησιμοποιήθηκε ήταν:

- $n_neighbors = 50$
- $contamination = 0,10$

Στο κοινό δείγμα των 100.000 συναλλαγών, το LOF χαρακτήρισε 10.000 συναλλαγές ως ανωμαλίες και εντόπισε 714 πραγματικές περιπτώσεις απάτης. Το Precision διαμορφώθηκε σε 0,0714, το Recall σε 0,2041 και το F1-score σε περίπου 0,1058.

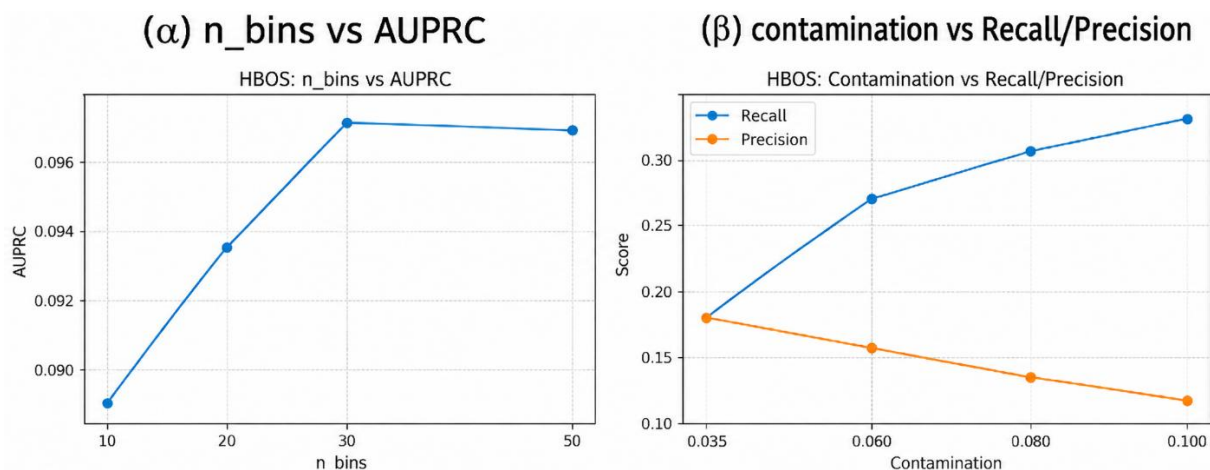
Τα αποτελέσματα υποδηλώνουν περιορισμένη ατομική ικανότητα του LOF ως προς την ανάκτηση της θετικής κλάσης στο συγκεκριμένο πειραματικό πλαίσιο. Μία πιθανή ερμηνεία σχετίζεται με τον τοπικό χαρακτήρα του αλγορίθμου, καθώς η βαθμολογία μιας παρατήρησης εξαρτάται από την πυκνότητα της περιοχής στην οποία βρίσκεται σε σχέση με τις γειτονικές της παρατηρήσεις. Σε ένα πολυδιάστατο και ετερογενές σύνολο συναλλαγών, ορισμένες περιπτώσεις απάτης ενδέχεται να βρίσκονται σε περιοχές που δεν διαφοροποιούνται επαρκώς από την τοπική δομή των νόμιμων συναλλαγών.

Η χαμηλότερη ατομική επίδοση του LOF δεν συνεπάγεται απαραίτητα ότι οι πληροφορίες που παράγει είναι πλήρως πλεονάζουσες ή μη αξιοποιήσιμες. Η διαφορετική λογική εντοπισμού που εφαρμόζει ενδέχεται να οδηγεί στην επισήμανση διαφορετικού υποσυνόλου συναλλαγών, ζήτημα που εξετάζεται αναλυτικότερα στην Ενότητα 4.7.

4.4 Histogram-Based Outlier Score

Η πειραματική παραμετροποίηση του HBOS επικεντρώθηκε στις υπερπαραμέτρους n_bins και $contamination$. Για το n_bins εξετάστηκαν οι τιμές 10, 20, 30 και 50, ενώ για το $contamination$ χρησιμοποιήθηκαν οι τιμές 0,035, 0,06, 0,08 και 0,10. Συνολικά εξετάστηκαν 16 διαφορετικοί συνδυασμοί.

Η αύξηση του αριθμού των bins από 10 σε 20 και στη συνέχεια σε 30 συνοδεύτηκε από βελτίωση της AUPRC. Στα 50 bins παρατηρήθηκε μικρή μείωση σε σχέση με τα 30, χωρίς όμως μεγάλη διαφοροποίηση της συνολικής επίδοσης. Το αποτέλεσμα αυτό υποδηλώνει ότι η αύξηση της διακριτικής ικανότητας των ιστογραμμάτων μέχρι ένα συγκεκριμένο σημείο ενδέχεται να συμβάλλει στην καλύτερη αποτύπωση των κατανομών των χαρακτηριστικών, χωρίς η περαιτέρω αύξηση του αριθμού των bins να συνεπάγεται απαραίτητα επιπλέον όφελος.



Σχήμα 4.3. Επίδραση του αριθμού των bins και του $contamination$ στην απόδοση του HBOS.

Ως προς το $contamination$, παρατηρήθηκε αντίστοιχη τάση με εκείνη των προηγούμενων μοντέλων. Η αύξησή του συνοδεύτηκε από υψηλότερο Recall και χαμηλότερο Precision. Η συγκεκριμένη μεταβολή αντανάκλα τη διαφορετική θέση του σημείου λειτουργίας και όχι απαραίτητα μεταβολή της εγγενούς ικανότητας κατάταξης του μοντέλου.

Η τελική διαμόρφωση που χρησιμοποιήθηκε στην κοινή αξιολόγηση ήταν:

- $n_bins = 50$
- $contamination = 0,10$

Παρότι στο γράφημα της AUPRC η τιμή των 30 bins εμφανίζει οριακά υψηλότερο αποτέλεσμα, η διαφορά σε σχέση με τα 50 bins είναι μικρή. Επομένως, η συγκεκριμένη εικόνα δεν υποστηρίζει την ύπαρξη ουσιαστικής υπεροχής της μίας τιμής έναντι της άλλης αποκλειστικά βάσει της AUPRC.

Στο κοινό δείγμα των 100.000 συναλλαγών, το HBOS επισήμανε 10.000 παρατηρήσεις ως ανωμαλίες και εντόπισε 1.196 πραγματικές περιπτώσεις απάτης. Το Precision διαμορφώθηκε σε 0,1196, το Recall σε 0,3418 και το F1-score σε περίπου 0,1772.

Η συγκεκριμένη επίδοση υποδηλώνει ότι το HBOS κατάφερε να εντοπίσει σημαντικό μέρος της θετικής κλάσης, χωρίς όμως το Precision να διαμορφωθεί σε ιδιαίτερα υψηλό επίπεδο. Μία πιθανή ερμηνεία

σχετίζεται με τον τρόπο λειτουργίας του μοντέλου, το οποίο αξιολογεί την πιθανότητα ανωμαλίας βάσει των κατανομών των επιμέρους χαρακτηριστικών. Με τον τρόπο αυτό, συναλλαγές που παρουσιάζουν ακραίες τιμές σε συγκεκριμένες διαστάσεις μπορούν να λάβουν υψηλή βαθμολογία ανωμαλίας.

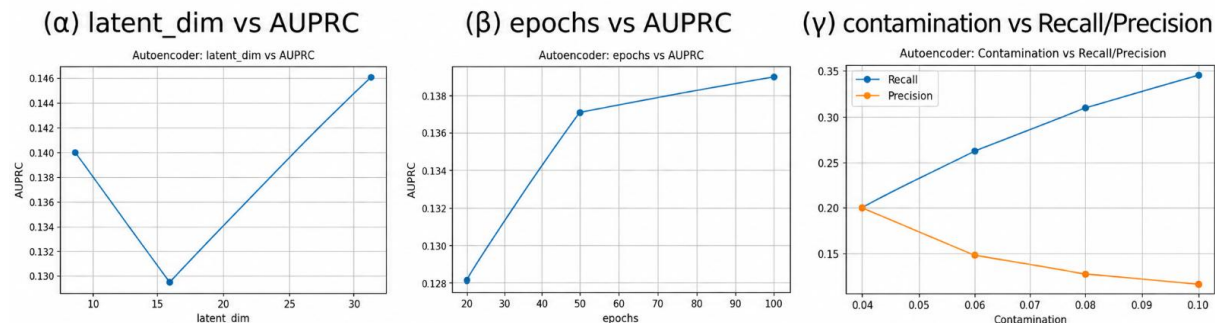
Παράλληλα, η βασική υπόθεση ανεξαρτησίας μεταξύ των χαρακτηριστικών περιορίζει τη δυνατότητα άμεσης αποτύπωσης σύνθετων πολυμεταβλητών αλληλεπιδράσεων. Συνεπώς, η παρατηρούμενη επίδοση πιθανώς αντανακλά τόσο την ικανότητα του μοντέλου να εντοπίζει ασυνήθιστες περιθωριακές κατανομές όσο και τους περιορισμούς που προκύπτουν όταν οι ανωμαλίες εξαρτώνται από συνδυασμούς χαρακτηριστικών.

4.5 Autoencoder

Η παραμετροποίηση του Autoencoder περιλάμβανε τη διερεύνηση της διάστασης του λανθάνοντος χώρου (latent_dim), του αριθμού των εποχών εκπαίδευσης και του contamination. Για τον λανθάνοντα χώρο εξετάστηκαν οι τιμές 8, 16 και 32, για τον αριθμό των εποχών οι τιμές 20, 50 και 100, ενώ το contamination έλαβε τις τιμές 0,035, 0,06, 0,08 και 0,10.

Παράλληλα, εξετάστηκαν δύο διαφορετικές στρατηγικές εκπαίδευσης: μία στην οποία χρησιμοποιήθηκε το σύνολο των διαθέσιμων παρατηρήσεων και μία στην οποία η εκπαίδευση πραγματοποιήθηκε αποκλειστικά σε συναλλαγές της κανονικής κλάσης. Η δεύτερη περίπτωση προϋποθέτει τη χρήση των πραγματικών ετικετών για τον σχηματισμό του συνόλου εκπαίδευσης και, επομένως, προσεγγίζει περισσότερο μία στρατηγική μιας κλάσης ή ανίχνευση νέων ή μη αναμενόμενων παρατηρήσεων διαδικασία παρά μία αυστηρά πλήρως μη επιβλεπόμενη στρατηγική. Η συγκεκριμένη διαφοροποίηση λαμβάνεται υπόψη κατά την ερμηνεία των αποτελεσμάτων.

Η μεταβολή του λανθάνοντος χώρου δεν παρουσίασε μονοτονική σχέση με την AUPRC. Η μετάβαση από 8 σε 16 διαστάσεις συνοδεύτηκε από μικρή μείωση, ενώ η τιμή 32 παρουσίασε την υψηλότερη AUPRC μεταξύ των εξεταζόμενων επιλογών. Το αποτέλεσμα αυτό ενδέχεται να υποδηλώνει ότι μία μεγαλύτερη λανθάνουσα αναπαράσταση διατηρεί περισσότερη πληροφορία που είναι χρήσιμη για την ανακατασκευή των συναλλακτικών προτύπων, χωρίς ωστόσο από τις συγκεκριμένες τρεις τιμές να μπορεί να συναχθεί γενικό συμπέρασμα για τη βέλτιστη διάσταση του λανθάνοντος χώρου.



Σχήμα 4.4. Επίδραση της διάστασης του λανθάνοντος χώρου, των εποχών εκπαίδευσης και του contamination στην απόδοση του Autoencoder.

Ως προς τον αριθμό των εποχών, η AUPRC αυξήθηκε μεταξύ των 20 και 50 εποχών και παρουσίασε μικρότερη περαιτέρω αύξηση στις 100. Η συγκεκριμένη τάση φαίνεται να υποστηρίζει ότι η μεγαλύτερη διάρκεια εκπαίδευσης συνέβαλε στη βελτίωση της αναπαράστασης των δεδομένων μέχρι το εξεταζόμενο εύρος, ενώ το μικρότερο πρόσθετο όφελος μεταξύ 50 και 100 εποχών ενδέχεται να υποδηλώνει σταδιακή σύγκλιση.

Η αύξηση του contamination συνοδεύτηκε, όπως και στα προηγούμενα μοντέλα, από αύξηση του Recall και μείωση του Precision. Επομένως, η συγκεκριμένη παράμετρος φαίνεται να επηρεάζει κυρίως το σημείο λειτουργίας του μοντέλου και τη σχέση μεταξύ ευαισθησίας και ακρίβειας των θετικών προβλέψεων.

Η τελική διαμόρφωση που χρησιμοποιήθηκε ήταν:

- latent_dim = 32
- epochs = 100
- contamination = 0,10

Η αρχιτεκτονική του δικτύου διαμορφώθηκε ως 218–128–64–32–64–128–218, με ReLU στα ενδιάμεσα επίπεδα και γραμμική ενεργοποίηση στο επίπεδο εξόδου. Για τη βελτιστοποίηση χρησιμοποιήθηκε ο αλγόριθμος Adam με συνάρτηση απώλειας το MSE, ενώ εφαρμόστηκε ο μηχανισμός πρόωρης διακοπής (Early Stopping) ώστε να περιορίζεται η συνέχιση της εκπαίδευσης όταν δεν παρατηρούνταν περαιτέρω βελτίωση στην επικύρωση.

Στο κοινό δείγμα των 100.000 συναλλαγών, ο Autoencoder χαρακτήρισε 10.000 συναλλαγές ως ανωμαλίες και εντόπισε 1.245 από τις 3.499 πραγματικές περιπτώσεις απάτης. Το Precision διαμορφώθηκε σε 0,1245, το Recall σε 0,3558 και το F1-score σε περίπου 0,1845. Παράλληλα, η AUPRC/AP ανήλθε σε 0,1395 και η ROC-AUC σε 0,6905.

Τα αποτελέσματα υποδηλώνουν ότι ο Autoencoder παρουσίασε αξιοσημείωτη ικανότητα ανάκτησης της θετικής κλάσης σε σχέση με ορισμένα από τα υπόλοιπα μοντέλα. Η συμπεριφορά αυτή ενδέχεται να σχετίζεται με τη δυνατότητα του νευρωνικού δικτύου να αναπαριστά μη γραμμικές σχέσεις μεταξύ των χαρακτηριστικών και να αποδίδει υψηλότερο σφάλμα ανακατασκευής σε παρατηρήσεις που αποκλίνουν από τα πρότυπα που έμαθε κατά την εκπαίδευση.

Ωστόσο, η παρατηρούμενη επίδοση δεν μπορεί να αποδοθεί αποκλειστικά στην αρχιτεκτονική του μοντέλου, καθώς επηρεάζεται επίσης από τη στρατηγική εκπαίδευσης, τη διάσταση του λανθάνοντος χώρου, τη διάρκεια εκπαίδευσης και τον τρόπο καθορισμού του κατωφλίου ανωμαλίας. Για τον λόγο αυτό, η τελική αποτίμηση του Autoencoder πραγματοποιείται σε συνδυασμό με τις υπόλοιπες μετρικές και με τη συγκριτική ανάλυση που ακολουθεί στην Ενότητα 4.6.

4.6 Συγκριτική Αξιολόγηση Μοντέλων

Μετά την ολοκλήρωση της διαδικασίας παραμετροποίησης και την επιλογή της τελικής διαμόρφωσης για κάθε μοντέλο, πραγματοποιήθηκε συγκριτική αξιολόγηση των Isolation Forest, LOF, HBOS και Autoencoder στο κοινό δείγμα των 100.000 συναλλαγών. Η χρήση των ίδιων παρατηρήσεων επιτρέπει τη σύγκριση των μοντέλων κάτω από κοινές συνθήκες αξιολόγησης. Οι τελικές τιμές των βασικών μετρικών παρουσιάζονται στον Πίνακα 4.1.

Πίνακας 4.1. Συγκριτική αξιολόγηση των τελικών διαμορφώσεων των μοντέλων ανίχνευσης ανωμαλιών.

Model	Recall	Precision	F1-score	AUPRC/AP	ROC-AUC
Isolation Forest	0,4058	0,1420	0,2104	0,1115	0,7235
LOF	0,2041	0,0714	0,1058	0,0549	0,6057
HBOS	0,3418	0,1196	0,1772	0,1009	0,6557
Autoencoder	0,3558	0,1245	0,1845	0,1395	0,6905

Η σύγκριση των αποτελεσμάτων δεν αναδεικνύει ένα μοντέλο που να υπερέχει σε όλες τις εξεταζόμενες μετρικές. Το Isolation Forest παρουσίασε το υψηλότερο Recall, με τιμή 0,4058, εντοπίζοντας το μεγαλύτερο ποσοστό των πραγματικών περιπτώσεων απάτης μεταξύ των τεσσάρων μοντέλων. Ακολούθησαν ο Autoencoder με Recall 0,3558 και το HBOS με 0,3418, ενώ το LOF παρουσίασε αισθητά χαμηλότερη τιμή, 0,2041. Η συγκεκριμένη κατάταξη υποδηλώνει ότι, στο επιλεγμένο σημείο λειτουργίας, το Isolation Forest εμφάνισε μεγαλύτερη ευαισθησία ως προς την ανίχνευση της θετικής κλάσης.

Παρόμοια εικόνα παρατηρείται ως προς το Precision. Το Isolation Forest κατέγραψε την υψηλότερη τιμή, 0,1420, ακολουθούμενο από τον Autoencoder με 0,1245 και το HBOS με 0,1196. Το LOF εμφάνισε Precision 0,0714. Παρά το γεγονός ότι οι απόλυτες τιμές του Precision παραμένουν σχετικά χαμηλές για όλα τα μοντέλα, η ερμηνεία τους πρέπει να πραγματοποιείται υπό το πρίσμα της έντονης ανισορροπίας του συνόλου δεδομένων και του γεγονότος ότι τα μοντέλα λειτουργούν ως μηχανισμοί επισήμανσης ύποπτων συναλλαγών και όχι ως μηχανισμοί οριστικής επιβεβαίωσης απάτης.

Η συνεξέταση Precision και Recall μέσω του F1-score οδηγεί σε αντίστοιχη κατάταξη. Το Isolation Forest παρουσίασε την υψηλότερη τιμή F1-score, ίση με 0,2104, ενώ ακολούθησαν ο Autoencoder με 0,1845 και το HBOS με 0,1772. Το LOF παρουσίασε τη χαμηλότερη τιμή, 0,1058. Επομένως, με βάση τις μετρικές που εξαρτώνται από το επιλεγμένο κατώφλι ταξινόμησης, το Isolation Forest φαίνεται να επιτυγχάνει την πιο ισορροπημένη σχέση μεταξύ ανάκτησης πραγματικών περιπτώσεων απάτης και ακρίβειας των θετικών προβλέψεων.

Διαφορετική εικόνα προκύπτει, ωστόσο, από την AUPRC/AP. Ο Autoencoder κατέγραψε την υψηλότερη τιμή, 0,1395, ενώ ακολούθησαν το Isolation Forest με 0,1115, το HBOS με 0,1009 και το LOF με 0,0549. Το εύρημα αυτό είναι ιδιαίτερα σημαντικό, καθώς η AUPRC/AP εξετάζει τη σχέση Precision–Recall σε διαφορετικά σημεία λειτουργίας και δεν περιορίζεται στο συγκεκριμένο κατώφλι που χρησιμοποιήθηκε για την παραγωγή των τελικών δυαδικών προβλέψεων. Κατά συνέπεια, η υψηλότερη τιμή του Autoencoder υποδηλώνει ότι οι συνεχείς βαθμολογίες ανωμαλίας του φαίνεται να επιτρέπουν αποτελεσματικότερη κατάταξη των πραγματικών περιπτώσεων απάτης σε ένα ευρύτερο εύρος κατωφλίων, παρά το γεγονός ότι στο τελικό operating point το Isolation Forest παρουσιάζει υψηλότερο Recall, Precision και F1-score.

Η διαφοροποίηση μεταξύ των δύο μοντέλων γίνεται ακόμη περισσότερο εμφανής όταν εξεταστεί παράλληλα η ROC-AUC. Το Isolation Forest παρουσίασε την υψηλότερη ROC-AUC, με τιμή 0,7235, ακολουθούμενο από τον Autoencoder με 0,6905, το HBOS με 0,6557 και το LOF με 0,6057. Η υψηλότερη ROC-AUC του Isolation Forest υποδηλώνει καλύτερη συνολική ικανότητα διάκρισης μεταξύ των δύο κλάσεων ως προς τη σχέση True Positive Rate και False Positive Rate. Ωστόσο, σε ένα ιδιαίτερα μη ισορροπημένο πρόβλημα, η ROC-AUC μπορεί να επηρεάζεται από τον μεγάλο αριθμό αρνητικών παρατηρήσεων. Για τον λόγο αυτό, η AUPRC/AP αποκτά ιδιαίτερη σημασία στην παρούσα αξιολόγηση, καθώς επικεντρώνεται περισσότερο στη συμπεριφορά του μοντέλου ως προς τη θετική και σπανιότερη κλάση.

Ο Autoencoder παρουσιάζει επομένως μία διαφορετική εικόνα από εκείνη που προκύπτει αποκλειστικά από τις threshold-dependent μετρικές. Παρότι το Recall και το F1-score του είναι χαμηλότερα από εκείνα του Isolation Forest στο επιλεγμένο κατώφλι, η υψηλότερη AUPRC/AP υποδηλώνει ότι η βαθμολογία ανακατασκευαστικού σφάλματος ενδέχεται να παρέχει αποτελεσματικότερη σχετική κατάταξη των ύποπτων παρατηρήσεων. Η συμπεριφορά αυτή μπορεί να σχετίζεται με τη δυνατότητα του Autoencoder να αναπαριστά μη γραμμικές σχέσεις μεταξύ των χαρακτηριστικών και να αποδίδει υψηλότερες βαθμολογίες ανωμαλίας σε συναλλαγές που αποκλίνουν από τα πρότυπα που έχουν αποτυπωθεί κατά την εκπαίδευση.

Το HBOS καταλαμβάνει ενδιάμεση θέση στις περισσότερες μετρικές. Με Recall 0,3418, Precision 0,1196, F1-score 0,1772 και AUPRC/AP 0,1009, η απόδοσή του είναι σαφώς υψηλότερη από εκείνη του LOF, αλλά χαμηλότερη από τις καλύτερες επιδόσεις των Isolation Forest και Autoencoder. Η συμπεριφορά αυτή ενδέχεται να συνδέεται με τον σχετικά απλό μηχανισμό του HBOS, ο οποίος μπορεί να εντοπίζει αποτελεσματικά παρατηρήσεις που εμφανίζουν ασυνήθιστες τιμές στις περιθωριακές κατανομές των χαρακτηριστικών, χωρίς όμως να μοντελοποιεί άμεσα πολύπλοκες αλληλεπιδράσεις μεταξύ τους.

Αντίθετα, το LOF παρουσίασε τις χαμηλότερες τιμές σε όλες τις βασικές μετρικές, με Recall 0,2041, Precision 0,0714, F1-score 0,1058, AUPRC/AP 0,0549 και ROC-AUC 0,6057. Τα αποτελέσματα αυτά υποδηλώνουν ότι η τοπική προσέγγιση πυκνότητας του LOF δεν ήταν εξίσου αποτελεσματική για την ατομική ανίχνευση της θετικής κλάσης στο συγκεκριμένο σύνολο δεδομένων. Η συγκεκριμένη παρατήρηση, ωστόσο, αφορά την αυτόνομη επίδοση του μοντέλου και δεν συνεπάγεται ότι οι συναλλαγές που εντοπίζει είναι κατ' ανάγκη ίδιες με εκείνες που εντοπίζονται από τα ισχυρότερα μοντέλα.

Συνολικά, τα αποτελέσματα αναδεικνύουν δύο διαφορετικές μορφές ισχυρής επίδοσης. Το Isolation Forest εμφανίζει την καλύτερη συμπεριφορά στο επιλεγμένο operating point, καθώς παρουσιάζει το υψηλότερο Recall, Precision και F1-score, καθώς και την υψηλότερη ROC-AUC. Ο Autoencoder, από την άλλη πλευρά, παρουσιάζει την υψηλότερη AUPRC/AP, γεγονός που υποδηλώνει ισχυρότερη συμπεριφορά ως προς την κατάταξη των πραγματικών περιπτώσεων απάτης κατά μήκος της καμπύλης Precision–Recall. Το HBOS παρουσιάζει ενδιάμεση επίδοση, ενώ το LOF εμφανίζει τη χαμηλότερη ατομική απόδοση.

Η παραπάνω σύγκριση δεν επαρκεί, ωστόσο, για να προσδιοριστεί κατά πόσο τα τέσσερα μοντέλα εντοπίζουν τις ίδιες ή διαφορετικές περιπτώσεις απάτης. Δύο μοντέλα μπορεί να εμφανίζουν

διαφορετική συνολική απόδοση αλλά ταυτόχρονα να παρέχουν μη επικαλυπτόμενη πληροφορία σε επίπεδο συναλλαγής. Για τον λόγο αυτό, στην επόμενη ενότητα εξετάζεται η συμπληρωματικότητα των μοντέλων, μέσω της ανάλυσης της επικάλυψης των πραγματικών θετικών ανιχνεύσεων και της σχέσης μεταξύ των παραγόμενων βαθμολογιών ανωμαλίας.

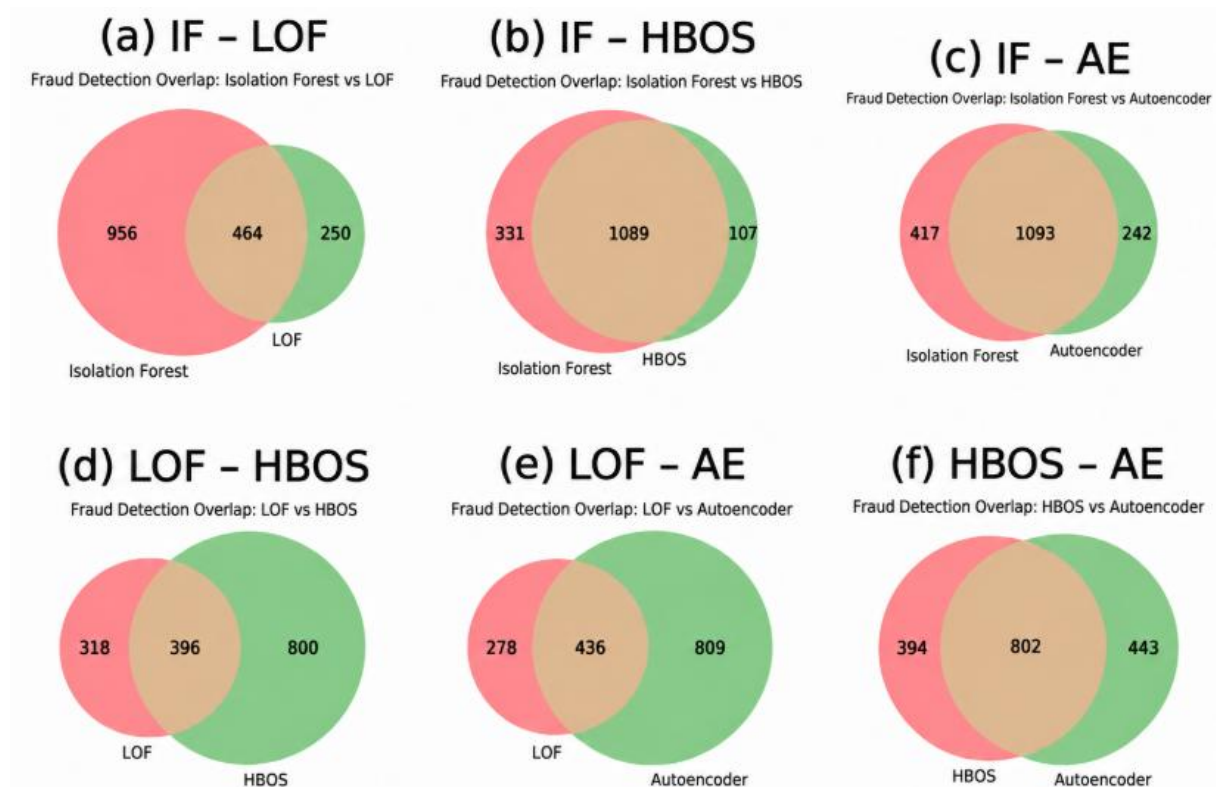
4.7 Ανάλυση Συμπληρωματικότητας των Μοντέλων

Η συγκριτική αξιολόγηση των επιμέρους μοντέλων επιτρέπει την αποτίμηση της απόδοσής τους μέσω συνολικών μετρικών, δεν αποτυπώνει όμως πλήρως τον βαθμό στον οποίο οι ανιχνευτές βασίζονται σε κοινή ή διαφορετική πληροφορία. Δύο μοντέλα μπορούν να εμφανίζουν παρόμοια επίδοση, αλλά να εντοπίζουν διαφορετικά υποσύνολα πραγματικών περιπτώσεων απάτης ή να αποδίδουν διαφορετικό βαθμό ανωμαλίας στις ίδιες συναλλαγές. Η διερεύνηση αυτών των διαφορών είναι ιδιαίτερα σημαντική, καθώς η ύπαρξη μη πλήρως επικαλυπτόμενης συμπεριφοράς ενδέχεται να δημιουργεί περιθώριο αξιοποίησης της συμπληρωματικότητας μεταξύ διαφορετικών μηχανισμών ανίχνευσης.

Για τον σκοπό αυτό, τα Isolation Forest, Local Outlier Factor, HBOS και Autoencoder εξετάστηκαν στο ίδιο στρωματοποιημένο δείγμα 100.000 συναλλαγών του IEEE-CIS, το οποίο περιλαμβάνει 3.499 πραγματικές περιπτώσεις απάτης. Η χρήση κοινού δείγματος εξασφαλίζει ότι οι προβλέψεις και οι βαθμολογίες ανωμαλίας των τεσσάρων μοντέλων αντιστοιχούν στις ίδιες ακριβώς παρατηρήσεις και, επομένως, μπορούν να συγκριθούν σε επίπεδο συναλλαγής. Η ανάλυση πραγματοποιήθηκε μέσω της επικάλυψης των πραγματικών θετικών ανιχνεύσεων, της συσχέτισης των συνεχών βαθμολογιών ανωμαλίας και της ζευγαρωτής στατιστικής σύγκρισης των ανιχνεύσεων.

4.7.1 Ανάλυση Επικάλυψης

Αρχικά εξετάστηκε η επικάλυψη μεταξύ των πραγματικών περιπτώσεων απάτης που εντοπίστηκαν από κάθε δυνατό ζεύγος μοντέλων. Για κάθε σύγκριση καταγράφηκαν οι περιπτώσεις που εντοπίστηκαν από κοινού, καθώς και εκείνες που εντοπίστηκαν αποκλειστικά από το ένα από τα δύο μοντέλα. Η ομοιότητα των αντίστοιχων συνόλων ποσοτικοποιήθηκε επιπλέον μέσω του δείκτη Jaccard [55], σύμφωνα με την διαδικασία που περιγράφηκε στην ενότητα 3.3.



Σχήμα 4.5. Επικάλυψη των πραγματικών περιπτώσεων απάτης που εντοπίστηκαν από τα επιμέρους ζεύγη μοντέλων.

Στη σύγκριση μεταξύ Isolation Forest και LOF, τα δύο μοντέλα εντόπισαν από κοινού 464 πραγματικές περιπτώσεις απάτης. Παράλληλα, 956 περιπτώσεις εντοπίστηκαν αποκλειστικά από το Isolation Forest και 250 αποκλειστικά από το LOF. Ο δείκτης Jaccard διαμορφώθηκε σε 0,2778, καταγράφοντας σχετικά περιορισμένη επικάλυψη μεταξύ των δύο συνόλων ανίχνευσης.

Για το ζεύγος Isolation Forest–HBOS παρατηρήθηκε σαφώς μεγαλύτερη επικάλυψη. Συγκεκριμένα, 1.089 πραγματικές περιπτώσεις απάτης εντοπίστηκαν και από τα δύο μοντέλα, ενώ 331 περιπτώσεις καταγράφηκαν αποκλειστικά από το Isolation Forest και 107 αποκλειστικά από το HBOS. Η αντίστοιχη τιμή του δείκτη Jaccard ανήλθε σε 0,7132 και αποτέλεσε τη μεγαλύτερη μεταξύ των έξι ζευγαρωτών συγκρίσεων.

Στην περίπτωση των Isolation Forest και Autoencoder, οι κοινές πραγματικές θετικές ανιχνεύσεις ανήλθαν σε 1.003. Επιπλέον, το Isolation Forest εντόπισε 417 περιπτώσεις που δεν επισημάνθηκαν από τον Autoencoder, ενώ ο Autoencoder εντόπισε 242 περιπτώσεις που δεν περιλαμβάνονταν στις ανιχνεύσεις του Isolation Forest. Ο δείκτης Jaccard ήταν 0,6035, παρουσιάζοντας τη δεύτερη υψηλότερη τιμή μεταξύ των εξεταζόμενων ζευγών.

Η σύγκριση LOF–HBOS παρουσίασε 396 κοινές πραγματικές θετικές περιπτώσεις, ενώ 318 περιπτώσεις εντοπίστηκαν αποκλειστικά από το LOF και 800 αποκλειστικά από το HBOS. Η τιμή Jaccard ήταν 0,2616, η χαμηλότερη μεταξύ των έξι συγκρίσεων.

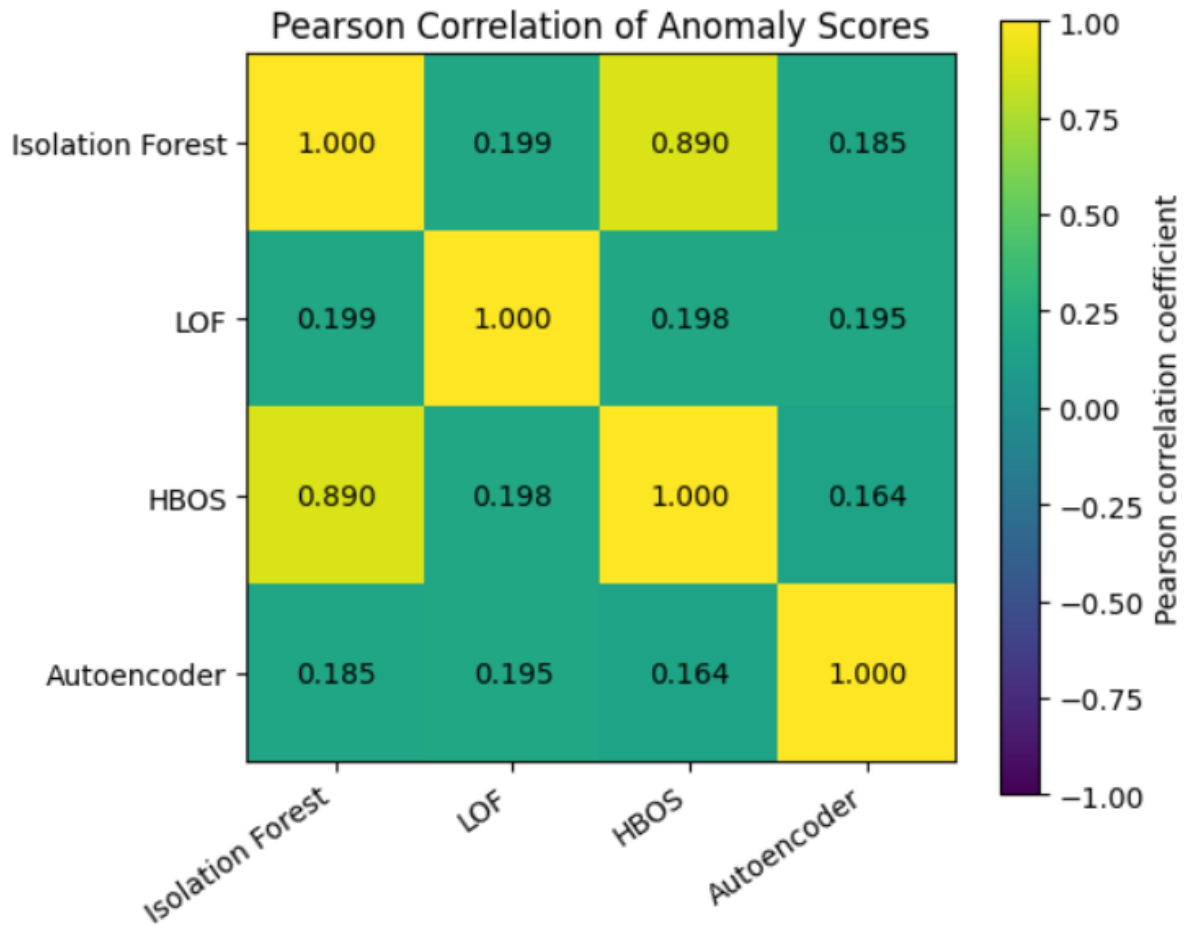
Παρόμοια εικόνα χαμηλής επικάλυψης προέκυψε και για το ζεύγος LOF–Autoencoder. Τα δύο μοντέλα εντόπισαν από κοινού 436 πραγματικές περιπτώσεις απάτης, ενώ 278 περιπτώσεις εντοπίστηκαν αποκλειστικά από το LOF και 809 αποκλειστικά από τον Autoencoder. Ο δείκτης Jaccard διαμορφώθηκε σε 0,2863.

Τέλος, μεταξύ HBOS και Autoencoder καταγράφηκαν 802 κοινές πραγματικές θετικές ανιχνεύσεις. Το HBOS εντόπισε επιπλέον 394 περιπτώσεις και ο Autoencoder 443 περιπτώσεις που δεν εντοπίστηκαν από το άλλο μοντέλο. Η τιμή Jaccard ανήλθε σε 0,4893, τοποθετώντας το συγκεκριμένο ζεύγος σε ενδιάμεσο επίπεδο επικάλυψης. Τα αποτελέσματα της ανάλυσης επικάλυψης συνοψίζονται στον Πίνακα 4.2.

Πίνακας 4.2. Επικάλυψη πραγματικών θετικών ανιχνεύσεων και δείκτης Jaccard ανά ζεύγος μοντέλων.

Ζεύγος μοντέλων	Μόνο A	Κοινές ανιχνεύσεις	Μόνο B	Jaccard
IF – LOF	956	464	250	0,2778
IF – HBOS	331	1.089	107	0,7132
IF – AE	417	1.003	242	0,6035
LOF – HBOS	318	396	800	0,2616
LOF – AE	278	436	809	0,2863
HBOS – AE	394	802	443	0,4893

Ως δεύτερο επίπεδο διερεύνησης εξετάστηκε η γραμμική συσχέτιση μεταξύ των συνεχών βαθμολογιών ανωμαλίας μέσω του συντελεστή συσχέτισης Pearson [54]. Η συγκεκριμένη ανάλυση εξετάζει κατά πόσο δύο μοντέλα αποδίδουν παρόμοια σχετική βαθμολογία στις ίδιες συναλλαγές.



Σχήμα 4.6. Συσχέτιση Pearson μεταξύ των βαθμολογιών ανωμαλίας των τεσσάρων μοντέλων.

Πίνακας 4.3. Συντελεστής συσχέτισης Pearson μεταξύ των βαθμολογιών ανωμαλίας.

Ζεύγος μοντέλων	Pearson r
IF – LOF	0,1988
IF – HBOS	0,8898
IF – AE	0,1849
LOF – HBOS	0,1975
LOF – AE	0,1947
HBOS – AE	0,1635

Η ισχυρότερη γραμμική συσχέτιση καταγράφηκε μεταξύ Isolation Forest και HBOS, με $r=0,8898$. Για όλα τα υπόλοιπα ζεύγη οι τιμές ήταν αισθητά χαμηλότερες και κυμάνθηκαν από 0,1635 έως 0,1988. Ειδικότερα, οι συσχετίσεις IF–LOF, IF–AE, LOF–HBOS, LOF–AE και HBOS–AE διαμορφώθηκαν αντίστοιχα σε 0,1988, 0,1849, 0,1975, 0,1947 και 0,1635. Η διαφορά μεταξύ της υψηλής τιμής IF–HBOS και των υπολοίπων ζευγών αποτελεί το κυριότερο εύρημα της συγκεκριμένης ανάλυσης και θα συνεκτιμηθεί στη συζήτηση της συμπληρωματικότητας στην επόμενη ενότητα.

Ως τρίτο επίπεδο σύγκρισης εφαρμόστηκε ο έλεγχος McNemar, ο οποίος είναι κατάλληλος για ζευγαρωτές συγκρίσεις δυαδικών αποτελεσμάτων πάνω στις ίδιες παρατηρήσεις [56]. Στην παρούσα

ανάλυση ο έλεγχος εφαρμόστηκε αποκλειστικά στις 3.499 πραγματικές περιπτώσεις απάτης. Συνεπώς, η σύγκριση αφορά τη διαφοροποίηση των μοντέλων ως προς την ανίχνευση της θετικής κλάσης και όχι τη συνολική ταξινομητική τους επίδοση στο σύνολο των 100.000 συναλλαγών. [57].

Επειδή πραγματοποιήθηκαν έξι ζευγαρωτές συγκρίσεις, εφαρμόστηκε διόρθωση Holm στα αντίστοιχα p-values για τον έλεγχο του σφάλματος που προκύπτει από τις πολλαπλές συγκρίσεις.

Πίνακας 4.4. Αποτελέσματα του ελέγχου McNemar μετά από διόρθωση Holm στις πραγματικές περιπτώσεις απάτης.

Ζεύγος μοντέλων	Μόνο Α εντόπισε	Μόνο Β εντόπισε	Holm-adjusted p-value	Στατιστικά σημαντική διαφορά
IF – LOF	956	250	<0,001	Ναι
IF – HBOS	331	107	<0,001	Ναι
IF – AE	417	242	<0,001	Ναι
LOF – HBOS	318	800	<0,001	Ναι
LOF – AE	278	809	<0,001	Ναι
HBOS – AE	394	443	0,0971	Όχι

Για πέντε από τα έξι ζεύγη το διορθωμένο p-value ήταν μικρότερο από 0,001, γεγονός που υποδηλώνει στατιστικά σημαντική ασυμμετρία μεταξύ των περιπτώσεων που εντοπίζει αποκλειστικά κάθε μοντέλο. Η μοναδική εξαίρεση ήταν το ζεύγος HBOS–Autoencoder, για το οποίο προέκυψε διορθωμένο $p=0,0971$. Στην περίπτωση αυτή, 394 πραγματικές περιπτώσεις απάτης εντοπίστηκαν αποκλειστικά από το HBOS και 443 αποκλειστικά από τον Autoencoder, χωρίς να προκύπτει στατιστικά σημαντική διαφορά μετά τη διόρθωση για πολλαπλές συγκρίσεις.

4.7.2 Συζήτηση της Συμπληρωματικότητας

Η συνδυαστική εξέταση των αποτελεσμάτων της επικάλυψης, του δείκτη Jaccard, της συσχέτισης Pearson και του ελέγχου McNemar υποδηλώνει ότι τα τέσσερα μοντέλα δεν λειτουργούν ως πλήρως ισοδύναμοι μηχανισμοί ανίχνευσης. Παρότι εφαρμόστηκαν στο ίδιο σύνολο συναλλαγών και στο ίδιο τελικό ποσοστό επίσημησης ανωμαλιών, οι περιπτώσεις απάτης που εντοπίζονται και ο τρόπος με τον οποίο βαθμολογούνται οι συναλλαγές διαφοροποιούνται μεταξύ των μοντέλων. Η διαφοροποίηση αυτή αποτελεί βασική ένδειξη πιθανής συμπληρωματικότητας, χωρίς από μόνη της να συνεπάγεται ότι οποιοσδήποτε συνδυασμός μοντέλων θα οδηγήσει απαραίτητα σε καλύτερη συνολική απόδοση.

Η μεγαλύτερη ομοιότητα παρατηρήθηκε μεταξύ του Isolation Forest και του HBOS. Το συγκεκριμένο ζεύγος παρουσίασε τον υψηλότερο δείκτη Jaccard, 0,7132, καθώς και ιδιαίτερα υψηλή συσχέτιση μεταξύ των βαθμολογιών ανωμαλίας, με Pearson $r=0,8898$. Η ταυτόχρονη εμφάνιση υψηλής επικάλυψης στις πραγματικές θετικές ανιχνεύσεις και υψηλής συσχέτισης των συνεχών βαθμολογιών υποδηλώνει ότι τα δύο μοντέλα τείνουν να αξιολογούν σημαντικό μέρος των συναλλαγών με παρόμοιο τρόπο. Παρότι οι μηχανισμοί τους είναι διαφορετικοί, και τα δύο φαίνεται να ανταποκρίνονται έντονα σε παρατηρήσεις που αποκλίνουν αισθητά από τις κυρίαρχες κατανομές του συνόλου δεδομένων. Το Isolation Forest επιδιώκει την απομόνωση παρατηρήσεων μέσω τυχαίων διαχωρισμών του χώρου χαρακτηριστικών, ενώ το HBOS βασίζεται στη σπανιότητα των τιμών στις επιμέρους κατανομές. Είναι επομένως πιθανό ορισμένες έντονα αποκλίνουσες συναλλαγές να λαμβάνουν υψηλή βαθμολογία ανωμαλίας και από τους δύο μηχανισμούς.

Η υψηλή ομοιότητα IF–HBOS υποδηλώνει παράλληλα μεγαλύτερο βαθμό πλεονασματικότητας σε σχέση με τα υπόλοιπα ζεύγη. Το HBOS εξακολουθεί να εντοπίζει περιπτώσεις που δεν εντοπίζονται από το Isolation Forest, ωστόσο ο αριθμός τους είναι σχετικά περιορισμένος. Συνεπώς, η προσθήκη

ενός μοντέλου με ιδιαίτερα παρόμοια συμπεριφορά σε έναν συνδυαστικό μηχανισμό ενδέχεται να παρέχει μικρότερο πρόσθετο πληροφοριακό όφελος από την προσθήκη ενός μοντέλου που διαφοροποιείται περισσότερο.

Διαφορετική εικόνα εμφανίζεται για το LOF. Παρότι στην ατομική αξιολόγηση παρουσίασε τις χαμηλότερες τιμές Recall, Precision, F1-score και AUPRC/AP, οι δείκτες επικάλυψης με τα υπόλοιπα μοντέλα ήταν σχετικά χαμηλοί. Ειδικότερα, οι τιμές Jaccard για τα ζεύγη LOF-IF, LOF-HBOS και LOF-AE παρέμειναν κοντά στο 0,26-0,29, ενώ και οι αντίστοιχες συσχετίσεις των βαθμολογιών ανωμαλίας ήταν χαμηλές. Η συμπεριφορά αυτή υποδηλώνει ότι το LOF εξετάζει διαφορετική πτυχή της δομής των δεδομένων.

Η διαφοροποίηση του LOF μπορεί να σχετίζεται με τον τοπικό χαρακτήρα του αλγορίθμου. Σε αντίθεση με τους μηχανισμούς που βασίζονται περισσότερο σε συνολικές ή περιθωριακές αποκλίσεις, το LOF αξιολογεί κάθε παρατήρηση σε σχέση με την πυκνότητα της άμεσης γειτονιάς της. Ως αποτέλεσμα, μία συναλλαγή μπορεί να μην εμφανίζεται ως ιδιαίτερα ακραία σε συνολικό επίπεδο, αλλά να χαρακτηρίζεται ως ανώμαλη επειδή βρίσκεται σε περιοχή μικρότερης πυκνότητας σε σχέση με τις γειτονικές της παρατηρήσεις. Αυτό ενδέχεται να εξηγεί γιατί το LOF εντοπίζει ορισμένες περιπτώσεις απάτης που δεν επισημαίνονται από τα υπόλοιπα μοντέλα, παρότι η συνολική ατομική του επίδοση είναι χαμηλότερη.

Η συγκεκριμένη παρατήρηση είναι σημαντική για την έννοια της συμπληρωματικότητας. Ένα μοντέλο δεν είναι απαραίτητο να αποτελεί το ισχυρότερο ατομικό μοντέλο για να παρέχει χρήσιμη πληροφορία σε έναν συνδυαστικό μηχανισμό. Εφόσον οι ορθές ανιχνεύσεις του δεν ταυτίζονται πλήρως με εκείνες των ισχυρότερων ανιχνευτών, μπορεί να συνεισφέρει πληροφορία που διαφορετικά θα απουσίαζε. Ωστόσο, η χρησιμότητα αυτής της διαφοροποίησης πρέπει να συνεκτιμάται με τον αυξημένο αριθμό λανθασμένων θετικών προβλέψεων και δεν μπορεί να αξιολογηθεί αποκλειστικά από τη χαμηλή τιμή του Jaccard.

Ιδιαίτερο ενδιαφέρον παρουσιάζει επίσης η σχέση μεταξύ Isolation Forest και Autoencoder. Τα δύο μοντέλα εμφάνισαν σχετικά υψηλή επικάλυψη στις πραγματικές θετικές ανιχνεύσεις, με Jaccard 0,6035, αλλά πολύ χαμηλότερη συσχέτιση στις συνεχείς βαθμολογίες ανωμαλίας, με Pearson $r=0,1849$. Η φαινομενική αυτή διαφοροποίηση δεν αποτελεί αντίφαση, καθώς οι δύο δείκτες εξετάζουν διαφορετικές πτυχές της συμπεριφοράς των μοντέλων. Ο Jaccard αφορά αποκλειστικά τα σύνολα των περιπτώσεων απάτης που τελικά χαρακτηρίστηκαν ως ανωμαλίες, ενώ ο Pearson εξετάζει τον τρόπο με τον οποίο βαθμολογούνται και οι 100.000 συναλλαγές πριν από τη μετατροπή των scores σε δυαδικές αποφάσεις. Επομένως, το IF και ο Autoencoder μπορούν να συμφωνούν σε σημαντικό αριθμό τελικών πραγματικών θετικών ανιχνεύσεων, αλλά να διαφέρουν ουσιαστικά ως προς τη σχετική κατάταξη και την ένταση με την οποία αξιολογούν τις υπόλοιπες συναλλαγές. Η συμπεριφορά αυτή φαίνεται να υποστηρίζει την ύπαρξη διαφορετικής πληροφοριακής δομής στις βαθμολογίες ανωμαλίας των δύο μοντέλων. Ο Autoencoder βασίζεται στο ανακατασκευαστικό σφάλμα και μπορεί να αποτυπώνει σύνθετες μη γραμμικές αποκλίσεις από τα πρότυπα που έχει μάθει, ενώ το Isolation Forest βασίζεται στην ευκολία απομόνωσης μιας παρατήρησης μέσω τυχαίων διαχωρισμών. Η διαφορετική αυτή λογική ενδέχεται να καθιστά τα δύο μοντέλα ιδιαίτερα ενδιαφέροντα όταν χρησιμοποιούνται σε συνδυαστικές στρατηγικές που αξιοποιούν συνεχείς βαθμολογίες και όχι μόνο δυαδικές αποφάσεις.

Το ζεύγος HBOS-Autoencoder παρουσίασε ενδιαμέσο επίπεδο επικάλυψης και τη χαμηλότερη συσχέτιση των βαθμολογιών ανωμαλίας μεταξύ όλων των ζευγών. Παράλληλα, ήταν το μοναδικό ζεύγος για το οποίο ο έλεγχος McNemar δεν ανέδειξε στατιστικά σημαντική διαφορά μετά τη διόρθωση Holm. Το αποτέλεσμα αυτό δεν σημαίνει ότι τα δύο μοντέλα παράγουν ίδιες προβλέψεις. Αντίθετα, οι προηγούμενες αναλύσεις έδειξαν ότι εντοπίζουν διαφορετικές επιμέρους περιπτώσεις. Ο έλεγχος McNemar υποδηλώνει συγκεκριμένα ότι, μεταξύ των πραγματικών περιπτώσεων απάτης, δεν προέκυψε επαρκής στατιστική ένδειξη ασυμμετρίας ως προς τον αριθμό των περιπτώσεων που εντοπίζονται αποκλειστικά από το ένα ή το άλλο μοντέλο.

Στα υπόλοιπα πέντε ζεύγη, οι στατιστικά σημαντικές διαφορές του McNemar υποδηλώνουν ότι η κατανομή των αποκλειστικών ορθών ανιχνεύσεων δεν είναι συμμετρική. Η παρατήρηση αυτή ενισχύει την εικόνα ότι τα μοντέλα δεν διαφοροποιούνται μόνο ως προς τη συνολική τιμή του Recall, αλλά και ως προς τις πιο συγκεκριμένες περιπτώσεις απάτης καταφέρνουν να εντοπίσουν. Ωστόσο, η στατιστική σημαντικότητα δεν πρέπει να ερμηνεύεται ως άμεση ένδειξη πρακτικής υπεροχής ή συμπληρωματικότητας. Για τον λόγο αυτό συνεκτιμάται με τα διαγράμματα Venn, τον Jaccard και τη συσχέτιση των βαθμολογιών ανωμαλίας.

Συνολικά, η ανάλυση υποδηλώνει την ύπαρξη διαφορετικών βαθμών συμπληρωματικότητας μεταξύ των τεσσάρων μοντέλων. Το IF και το HBOS εμφανίζουν τη μεγαλύτερη ομοιότητα και, κατά συνέπεια, μεγαλύτερο βαθμό πιθανής πλεονασματικότητας. Το LOF διαφοροποιείται περισσότερο από τα υπόλοιπα μοντέλα, χωρίς όμως η διαφοροποίηση αυτή να συνοδεύεται από ισχυρή ατομική απόδοση. Ο Autoencoder συνδυάζει σχετικά υψηλή ατομική επίδοση με διαφοροποιημένη συμπεριφορά στις συνεχείς βαθμολογίες ανωμαλίας, γεγονός που ενδέχεται να τον καθιστά περισσότερο κατάλληλο για συνδυαστική αξιοποίηση με μοντέλα διαφορετικής λογικής.

Η ύπαρξη αυτών των διαφορών παρέχει την ερμηνευτική βάση για τη χρήση τεχνικών συνδυασμού. Ένας συνδυαστικός μηχανισμός μπορεί να αξιοποιήσει τη συμφωνία μεταξύ ισχυρών μοντέλων στις περιπτώσεις όπου οι ενδείξεις ανωμαλίας είναι συνεπείς, ενώ παράλληλα μπορεί να διατηρήσει πληροφορία από μοντέλα που εντοπίζουν διαφορετικά υποσύνολα συναλλαγών. Το πιθανό όφελος επομένως δεν προκύπτει απλώς από την προσθήκη περισσότερων ανιχνευτών, αλλά από τον συνδυασμό μοντέλων που παρουσιάζουν διαφορετικούς αλλά δυνητικά συμπληρωματικούς μηχανισμούς ανίχνευσης. Η πραγματική αποτελεσματικότητα αυτής της υπόθεσης πρέπει να αξιολογηθεί μέσα από τις επιδόσεις των αντίστοιχων στρατηγικών συνδυασμού.

4.8 Συμπεράσματα Κεφαλαίου

Στο παρόν κεφάλαιο πραγματοποιήθηκε η πειραματική αξιολόγηση των τεσσάρων μοντέλων ανίχνευσης ανωμαλιών, εξετάζοντας τόσο την ατομική τους επίδοση όσο και τον βαθμό στον οποίο οι ανιχνεύσεις και οι βαθμολογίες ανωμαλίας τους παρουσιάζουν κοινά ή διαφοροποιημένα χαρακτηριστικά.

Στο επιλεγμένο σημείο λειτουργίας, το Isolation Forest παρουσίασε την ισχυρότερη συνολική συμπεριφορά ως προς τις δυαδικές προβλέψεις, καταγράφοντας τις υψηλότερες τιμές Precision, Recall και F1-score, καθώς και την υψηλότερη ROC-AUC μεταξύ των τεσσάρων μοντέλων. Ειδικότερα, το Recall διαμορφώθηκε σε 0,4058, γεγονός που υποδηλώνει ότι στο συγκεκριμένο σημείο λειτουργίας το μοντέλο εντόπισε μεγαλύτερο ποσοστό των πραγματικών περιπτώσεων απάτης σε σχέση με τους υπόλοιπους μεμονωμένους ανιχνευτές.

Ο Autoencoder παρουσίασε διαφορετικό προφίλ απόδοσης. Παρότι οι τιμές Recall και F1-score ήταν χαμηλότερες από εκείνες του Isolation Forest, κατέγραψε την υψηλότερη AUPRC/AP, με τιμή 0,1395. Το αποτέλεσμα αυτό υποδηλώνει καλύτερη συνολική ποιότητα κατάταξης των πραγματικών περιπτώσεων απάτης από τον Autoencoder σε όλο το εύρος των δυνατών σημείων λειτουργίας. Το HBOS παρουσίασε ενδιάμεση συμπεριφορά, ενώ το LOF εμφάνισε τις χαμηλότερες ατομικές επιδόσεις στις περισσότερες από τις εξεταζόμενες μετρικές.

Η ανάλυση συμπληρωματικότητας έδειξε ότι τα μοντέλα δεν εντοπίζουν πλήρως τα ίδια υποσύνολα πραγματικών περιπτώσεων απάτης και δεν αποδίδουν τις ίδιες σχετικές βαθμολογίες ανωμαλίας στις συναλλαγές. Η μεγαλύτερη ομοιότητα παρατηρήθηκε μεταξύ Isolation Forest και HBOS, όπως υποδηλώνεται τόσο από τον υψηλό δείκτη Jaccard όσο και από την ισχυρή συσχέτιση Pearson των βαθμολογιών ανωμαλίας. Αντίθετα, το LOF παρουσίασε χαμηλότερη επικάλυψη με τα υπόλοιπα μοντέλα, ενώ ο Autoencoder συνδύασε σχετικά ισχυρή ατομική επίδοση με χαμηλότερη συσχέτιση των συνεχών βαθμολογιών του με τα περισσότερα από τα υπόλοιπα μοντέλα.

Τα ευρήματα αυτά υποδηλώνουν ότι μεταξύ των μοντέλων υπάρχει διαφοροποιημένη συμπεριφορά, η οποία δημιουργεί τη βάση για τη διερεύνηση πιθανής συμπληρωματικότητας. Η ύπαρξη διαφορετικών ανιχνεύσεων ή χαμηλής συσχέτισης δεν συνεπάγεται, ωστόσο, ότι ο συνδυασμός των μοντέλων θα οδηγήσει αυτομάτως σε βελτίωση της συνολικής απόδοσης. Η υπόθεση αυτή πρέπει να αξιολογηθεί άμεσα μέσω των αποτελεσμάτων των αντίστοιχων στρατηγικών συνδυασμού.

Για τον λόγο αυτό, στο επόμενο κεφάλαιο εξετάζονται συστηματικά οι τεχνικές OR/Union, πλειοψηφικής ψηφοφορίας και σταθμισμένου συνδυασμού βαθμολογιών, με στόχο να διερευνηθεί κατά πόσο και υπό ποιες συνθήκες η συμπληρωματικότητα των επιμέρους ανιχνευτών μπορεί να αξιοποιηθεί αποτελεσματικά.

Κεφάλαιο 5ο: Αξιολόγηση των Τεχνικών Συνδυασμού Μοντέλων

Μετά την αξιολόγηση των επιμέρους μοντέλων ανίχνευσης ανωμαλιών και τη διερεύνηση της μεταξύ τους συμπληρωματικότητας στο προηγούμενο κεφάλαιο, στο παρόν κεφάλαιο εξετάζεται η συνδυαστική αξιοποίηση των Isolation Forest, LOF, HBOS και Autoencoder. Στόχος είναι να διερευνηθεί κατά πόσο διαφορετικές τεχνικές συνδυασμού μπορούν να αξιοποιήσουν τις διαφορετικές συμπεριφορές των επιμέρους μοντέλων και να μεταβάλουν τη σχέση μεταξύ της ανίχνευσης πραγματικών περιπτώσεων απάτης και του περιορισμού των ψευδώς θετικών επισημάνσεων.

Η αξιολόγηση πραγματοποιήθηκε στο ίδιο διαστρωματωμένο δείγμα των 100.000 συναλλαγών που χρησιμοποιήθηκε για τη συγκριτική αξιολόγηση των επιμέρους μοντέλων. Το δείγμα περιλαμβάνει 3.499 περιπτώσεις απάτης και 96.501 νόμιμες συναλλαγές. Με τον τρόπο αυτό διασφαλίζεται ότι όλες οι τεχνικές συνδυασμού αξιολογούνται στις ίδιες ακριβώς παρατηρήσεις και βάσει των τελικών αποτελεσμάτων των τεσσάρων μοντέλων.

Οι τεχνικές που εξετάζονται είναι το OR/Union Ensemble, το Majority Voting Ensemble με δύο διαφορετικά κατώφλια συμφωνίας και το Weighted Score-Based Ensemble. Στην τελευταία περίπτωση εξετάζονται τέσσερις διαφορετικές κατανομές βαρών, προκειμένου να διερευνηθεί η επίδραση της σχετικής συμμετοχής των επιμέρους μοντέλων στην τελική βαθμολογία ανωμαλίας.

5.1 Συνδυασμός OR/Union

Ο συνδυασμός OR/Union αποτελεί την περισσότερο συμπεριληπτική από τις εξεταζόμενες τεχνικές συνδυασμού. Σύμφωνα με τον συγκεκριμένο κανόνα, μία συναλλαγή χαρακτηρίζεται ως ύποπτη όταν τουλάχιστον ένα από τα τέσσερα επιμέρους μοντέλα την έχει χαρακτηρίσει ως ανωμαλία. Επομένως, για την τελική θετική απόφαση δεν απαιτείται συμφωνία μεταξύ περισσότερων μοντέλων.

Η συγκεκριμένη στρατηγική δίνει προτεραιότητα στην κάλυψη όσο το δυνατόν μεγαλύτερου αριθμού πραγματικών περιπτώσεων απάτης. Η λογική αυτή συνδέεται άμεσα με τα αποτελέσματα της ανάλυσης συμπληρωματικότητας του προηγούμενου κεφαλαίου, όπου διαπιστώθηκε ότι τα τέσσερα μοντέλα δεν εντοπίζουν ακριβώς τις ίδιες περιπτώσεις. Κατά συνέπεια, η ένωση των επισημάνσεών τους επιτρέπει τη διατήρηση περιπτώσεων απάτης που ενδέχεται να έχουν εντοπιστεί μόνο από έναν από τους ανιχνευτές. Τα αποτελέσματα του συνδυασμού OR/Union παρουσιάζονται στον Πίνακα 5.1.

Πίνακας 5.1. Αποτελέσματα του OR/Union Ensemble στο κοινό δείγμα των 100.000 συναλλαγών.

Μετρική	OR/Union Ensemble
Precision	0,0840
Recall	0,5519
F1-score	0,1458
AUPRC/AP	0,1007
ROC-AUC	0,6889
TP	1.931
FP	21.058
FN	1.568
TN	75.443

Ο συνδυασμός OR/Union εντόπισε 1.931 από τις 3.499 πραγματικές περιπτώσεις απάτης, επιτυγχάνοντας Ανάκληση ίση με 0,5519. Η τιμή αυτή είναι η υψηλότερη μεταξύ όλων των τεχνικών συνδυασμού που εξετάστηκαν. Το αποτέλεσμα υποδηλώνει ότι, στο συγκεκριμένο πειραματικό πλαίσιο, η ένωση των επισημάνσεων των τεσσάρων μοντέλων οδηγεί σε μεγαλύτερη κάλυψη της θετικής κλάσης.

Η αυξημένη Ανάκληση συνοδεύεται, ωστόσο, από σημαντική αύξηση των ψευδώς θετικών επισημάνσεων. Συγκεκριμένα, 21.058 νόμιμες συναλλαγές χαρακτηρίστηκαν ως ύποπτες, με αποτέλεσμα η Ακρίβεια να περιοριστεί σε 0,0840. Το F1-score διαμορφώθηκε σε 0,1458, γεγονός που αντανάκλα τη μεγάλη απόσταση μεταξύ της υψηλής Ανάκλησης και της χαμηλής Ακρίβειας.

Επομένως, η τεχνική συνδυασμού OR/Union εμφανίζει μία ιδιαίτερα ευαίσθητη συμπεριφορά ως προς την ανίχνευση, με έμφαση στην αύξηση του Recall εις βάρος της Precision. Η τεχνική μπορεί να είναι χρήσιμη σε περιπτώσεις όπου το κόστος μη εντοπισμού μιας πραγματικής απάτης θεωρείται σημαντικά υψηλότερο από το κόστος διερεύνησης πρόσθετων συναλλαγών. Αντίθετα, σε ένα περιβάλλον όπου ο αριθμός των συναλλαγών που μπορούν να εξεταστούν από ανθρώπινο ελεγκτή είναι περιορισμένος, ο μεγάλος αριθμός ψευδώς θετικών περιπτώσεων αποτελεί σημαντικό μειονέκτημα.

Για τον υπολογισμό των AUPRC/AP και ROC-AUC χρησιμοποιήθηκε ως βαθμολογία κατάταξης ο αριθμός των μοντέλων που χαρακτήρισαν κάθε συναλλαγή ως ανωμαλία. Η AUPRC/AP διαμορφώθηκε σε 0,1007, ενώ η ROC-AUC σε 0,6889.

5.2 Συνδυασμός Πλειοψηφικής Ψηφοφορίας

Ο συνδυασμός πλειοψηφικής ψηφοφορίας (Majority Voting Ensemble) εισάγει ένα αυστηρότερο κριτήριο απόφασης σε σχέση με τον OR/Union συνδυασμό, καθώς απαιτεί συμφωνία περισσότερων του ενός μοντέλων πριν μία συναλλαγή χαρακτηριστεί ως ύποπτη. Με τον τρόπο αυτό επιχειρείται ο περιορισμός των μεμονωμένων επισημάνσεων που μπορεί να προέρχονται από έναν μόνο ανιχνευτή. Στην παρούσα εργασία εξετάστηκαν δύο διαφορετικά κατώφλια συμφωνίας: η απαίτηση τουλάχιστον δύο θετικών ψήφων και η απαίτηση τουλάχιστον τριών θετικών ψήφων.

Κανόνας τουλάχιστον δύο ψήφων.

Στην πρώτη διαμόρφωση, μία συναλλαγή χαρακτηρίζεται ως ύποπτη όταν τουλάχιστον δύο από τα τέσσερα μοντέλα συμφωνούν ότι αποτελεί ανωμαλία. Το συγκεκριμένο σχήμα αποτελεί μία ενδιάμεση προσέγγιση μεταξύ του ιδιαίτερα συμπεριληπτικού συνδυασμού OR/Union και της αυστηρότερης απαίτησης τριών ψήφων.

Η εφαρμογή του κανόνα των δύο ψήφων οδήγησε στον εντοπισμό 1.402 πραγματικών περιπτώσεων απάτης, με Ανάκληση 0,4007. Παράλληλα, οι ψευδώς θετικές περιπτώσεις περιορίστηκαν στις 9.439, έναντι 21.058 στο OR/Union συνδυασμό. Ως αποτέλεσμα, η Ακρίβεια αυξήθηκε σε 0,1293, ενώ το F1-score διαμορφώθηκε σε 0,1955.

Η απαίτηση δεύτερης θετικής ψήφου φαίνεται επομένως να περιορίζει σημαντικό μέρος των επισημάνσεων που προέρχονται αποκλειστικά από ένα μοντέλο. Η μείωση των ψευδώς θετικών περιπτώσεων συνοδεύεται, όμως, από απώλεια πραγματικών περιπτώσεων απάτης που δεν συγκεντρώνουν δύο θετικές ψήφους.

Κανόνας τουλάχιστον τριών ψήφων.

Στη δεύτερη διαμόρφωση, μία συναλλαγή χαρακτηρίζεται ως ύποπτη μόνο όταν τουλάχιστον τρία από τα τέσσερα μοντέλα συμφωνούν. Το αυστηρότερο αυτό κριτήριο αυξάνει σημαντικά τον απαιτούμενο βαθμό συμφωνίας μεταξύ των ανιχνευτών.

Οι ψευδώς θετικές περιπτώσεις περιορίστηκαν στις 3.927, με αποτέλεσμα η Ακρίβεια να αυξηθεί σε 0,1928. Ταυτόχρονα, όμως, η Ανάκληση μειώθηκε σε 0,2681, καθώς εντοπίστηκαν 938 από τις 3.499 πραγματικές περιπτώσεις απάτης. Το F1-score διαμορφώθηκε σε 0,2243, αποτελώντας την υψηλότερη τιμή F1-score μεταξύ όλων των εξεταζόμενων τεχνικών συνδυασμού. Η σύγκριση των δύο διαμορφώσεων παρουσιάζεται στον Πίνακα 5.2.

Πίνακας 5.2. Συγκριτικά αποτελέσματα των δύο διαμορφώσεων του Majority Voting Ensemble.

Μετρική	Majority Voting ≥ 2	Majority Voting ≥ 3
Ακρίβεια	0,1293	0,1928
Ανάκληση	0,4007	0,2681
F1-score	0,1955	0,2243
AUPRC/AP	0,1007	0,1007
ROC-AUC	0,6889	0,6889
Αληθώς θετικές περιπτώσεις	1.402	938
Ψευδώς θετικές περιπτώσεις	9.439	3.927
Ψευδώς αρνητικές περιπτώσεις	2.097	2.561

Τα αποτελέσματα αναδεικνύουν τον συμβιβασμό μεταξύ Recall και Precision. Η απαίτηση δύο ψήφων διατηρεί μεγαλύτερη ικανότητα εντοπισμού πραγματικών περιπτώσεων απάτης, ενώ η απαίτηση τριών ψήφων περιορίζει σημαντικά τις ψευδώς θετικές επισημάνσεις.

Η μετάβαση από δύο σε τρεις απαιτούμενες ψήφους αυξάνει την Ακρίβεια από 0,1293 σε 0,1928, αλλά μειώνει την Ανάκληση από 0,4007 σε 0,2681. Η επιλογή του απαιτούμενου αριθμού ψήφων εξαρτάται επομένως από το κατά πόσο δίνεται μεγαλύτερη έμφαση στην κάλυψη πιθανών περιπτώσεων απάτης ή στον περιορισμό του αριθμού των συναλλαγών που θα χρειαστούν περαιτέρω έλεγχο.

Οι δύο διαμορφώσεις παρουσιάζουν τις ίδιες τιμές AUPRC/AP και ROC-AUC, επειδή χρησιμοποιούν την ίδια βαθμολογία κατάταξης, δηλαδή τον συνολικό αριθμό θετικών ψήφων των τεσσάρων μοντέλων. Η αλλαγή αφορά αποκλειστικά το κατώφλι με βάση το οποίο η βαθμολογία μετατρέπεται σε τελική δυαδική απόφαση.

5.3 Σταθμισμένος Συνδυασμός Βαθμολογιών

Η τεχνική του σταθμισμένου συνδυασμού βαθμολογιών (Weighted Score-Based Ensemble) διαφοροποιείται από τις προηγούμενες τεχνικές, καθώς δεν βασίζεται αποκλειστικά στις τελικές δυαδικές αποφάσεις των επιμέρους μοντέλων. Αντίθετα, αξιοποιεί τις συνεχείς βαθμολογίες ανωμαλίας που παράγονται από κάθε μοντέλο, επιτρέποντας τη διατήρηση περισσότερης πληροφορίας σχετικά με τον βαθμό στον οποίο κάθε συναλλαγή θεωρείται ανώμαλη.

Οι βαθμολογίες που παράγονται από τα Isolation Forest, HBOS, LOF και Autoencoder βρίσκονται αρχικά σε διαφορετικές αριθμητικές κλίμακες. Για τον λόγο αυτό εφαρμόστηκε κανονικοποίηση ελάχιστης-μέγιστης τιμής, ώστε οι βαθμολογίες κάθε μοντέλου να μετασχηματιστούν στο διάστημα $[0,1]$. Με αυτόν τον τρόπο περιορίζεται η πιθανότητα ένα μοντέλο να επηρεάζει δυσανάλογα την τελική βαθμολογία αποκλειστικά λόγω της αρχικής κλίμακας των τιμών του.

Για κάθε συναλλαγή x , η συνολική βαθμολογία ανωμαλίας υπολογίζεται ως:

$$S(x) = w_{IF}S_{IF}(x) + w_{HBOS}S_{HBOS}(x) + w_{LOF}S_{LOF}(x) + w_{AE}S_{AE}(x) \quad (5.1)$$

όπου:

$$w_{IF} + w_{HBOS} + w_{LOF} + w_{AE} = 1 \quad (5.2)$$

Η αρχική διαμόρφωση βασίστηκε στην ισότιμη συμμετοχή των τεσσάρων μοντέλων, με βάρος 0,25 για κάθε μοντέλο. Επιπλέον, εξετάστηκαν τρεις διαφορετικές κατανομές βαρών, με στόχο να διερευνηθεί εάν η μεταβολή της σχετικής συμμετοχής των μοντέλων μπορεί να επηρεάσει την απόδοση της συνδυαστικής βαθμολογίας. Οι τέσσερις διαμορφώσεις παρουσιάζονται στον Πίνακα 5.3.

Πίνακας 5.3. Εξεταζόμενες διαμορφώσεις του σταθμισμένου συνδυασμού βαθμολογιών.

Διαμόρφωση	IF	HBOS	LOF	AE
Equal Weights	0,25	0,25	0,25	0,25
IF Dominant	0,50	0,15	0,15	0,20
Diversity Focused	0,20	0,10	0,35	0,35
Stability Focused	0,35	0,35	0,15	0,15

Οι συγκεκριμένες διαμορφώσεις δεν αποτελούν εξαντλητική αναζήτηση όλων των δυνατών συνδυασμών βαρών, αλλά διαφορετικές πειραματικές στρατηγικές κατανομής της σχετικής επιρροής των τεσσάρων μοντέλων.

Για όλες τις διαμορφώσεις εφαρμόστηκε ποσοστό αναμενόμενων ανωμαλιών ίσο με 0,10 κατά τον καθορισμό του τελικού κατωφλίου. Συνεπώς, ως ύποπτο χαρακτηρίστηκε περίπου το 10% των συναλλαγών με τις υψηλότερες συνολικές βαθμολογίες ανωμαλίας. Σε κάθε περίπτωση προέκυψαν 10.000 τελικές επισημάνσεις. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 5.4.

Πίνακας 5.4. Συγκριτικά αποτελέσματα των τεσσάρων διαμορφώσεων του Weighted Score-Based Ensemble.

Διαμόρφωση	Precision	Recall	F1-score	AUPRC/AP	ROC-AUC
Equal Weights	0,1351	0,3861	0,2002	0,1101	0,6970
IF Dominant	0,1388	0,3967	0,2056	0,1118	0,7121
Diversity Focused	0,1376	0,3933	0,2039	0,1123	0,7081
Stability Focused	0,1356	0,3875	0,2009	0,1097	0,6963

Η ισοβαρής διαμόρφωση παρουσίασε Ανάκληση 0,3861, Ακρίβεια 0,1351 και F1-score 0,2002. Η AUPRC/AP διαμορφώθηκε σε 0,1101 και η ROC-AUC σε 0,6970. Η συγκεκριμένη περίπτωση αποτελεί τη βασική διαμόρφωση αναφοράς, καθώς κανένα από τα τέσσερα μοντέλα δεν αποκτά μεγαλύτερη επιρροή στην τελική βαθμολογία.

Η διαμόρφωση IF Dominant, στην οποία το Isolation Forest λαμβάνει βάρος 0,50, παρουσίασε την υψηλότερη Ανάκληση μεταξύ των τεσσάρων διαμορφώσεων, ίση με 0,3967, καθώς και το υψηλότερο F1-score, ίσο με 0,2056. Παράλληλα, κατέγραψε την υψηλότερη ROC-AUC, ίση με 0,7121. Η συμπεριφορά αυτή είναι συμβατή με την ισχυρή ατομική απόδοση του Isolation Forest που παρατηρήθηκε στο προηγούμενο κεφάλαιο.

Η διαμόρφωση Diversity Focused παρουσίασε Ανάκληση 0,3933 και F1-score 0,2039, τιμές πολύ κοντά στις αντίστοιχες της IF Dominant. Παράλληλα, κατέγραψε την υψηλότερη AUPRC/AP μεταξύ των τεσσάρων διαμορφώσεων, με τιμή 0,1123. Το αποτέλεσμα υποδηλώνει ότι η συγκεκριμένη κατανομή βαρών οδήγησε σε ελαφρώς καλύτερη συνολική κατάταξη των συναλλαγών ως προς τη σχέση μεταξύ Ακρίβειας και Ανάκλησης.

Τέλος, η διαμόρφωση Stability Focused παρουσίασε AUPRC/AP ίση με 0,1097 και ROC-AUC ίση με 0,6963, τις χαμηλότερες αντίστοιχες τιμές μεταξύ των τεσσάρων διαμορφώσεων. Η μεγαλύτερη

συνολική βαρύτητα που αποδόθηκε στα Isolation Forest και HBOS δεν συνοδεύτηκε από βελτίωση της συνολικής ικανότητας κατάταξης στο συγκεκριμένο δείγμα.

Συνολικά, καμία από τις τέσσερις διαμορφώσεις δεν υπερέχει σε όλες τις μετρικές. Η IF Dominant παρουσιάζει την υψηλότερη Ανάκληση, το υψηλότερο F1-score και τη μεγαλύτερη ROC-AUC μεταξύ των σταθμισμένων διαμορφώσεων, ενώ η Diversity Focused παρουσιάζει την υψηλότερη AUPRC/AP. Το εύρημα αυτό καθιστά αναγκαία τη συγκριτική αξιολόγηση των τεχνικών με βάση περισσότερες από μία μετρικές.

5.4 Συγκριτική Αξιολόγηση των Τεχνικών Συνδυασμού

Η συνολική αξιολόγηση των τεχνικών συνδυασμού πραγματοποιήθηκε με βάση τις Precision, Recall, F1-score, AUPRC/AP και ROC-AUC. Στον Πίνακα 5.5 παρουσιάζονται συγκεντρωτικά τα αποτελέσματα όλων των εξεταζόμενων διαμορφώσεων.

Πίνακας 5.5. Συγκριτική αξιολόγηση των εξεταζόμενων τεχνικών συνδυασμού.

Τεχνική Συνδυασμού	Recall	Precision	F1-score	AUPRC/AP	ROC-AUC
OR/Union	0,5519	0,0840	0,1458	0,1007	0,6889
Majority Voting ≥ 2	0,4007	0,1293	0,1955	0,1007	0,6889
Majority Voting ≥ 3	0,2681	0,1928	0,2243	0,1007	0,6889
Weighted – Equal Weights	0,3861	0,1351	0,2002	0,1101	0,6970
Weighted – IF Dominant	0,3967	0,1388	0,2056	0,1118	0,7121
Weighted – Diversity Focused	0,3933	0,1376	0,2039	0,1123	0,7081
Weighted – Stability Focused	0,3875	0,1356	0,2009	0,1097	0,6963

Η συγκριτική αξιολόγηση αναδεικνύει διαφορετικές συμπεριφορές μεταξύ των τεχνικών, χωρίς να προκύπτει μία προσέγγιση που να υπερέχει σε όλες τις εξεταζόμενες μετρικές.

Η τεχνική OR/Union παρουσιάζει το υψηλότερο Recall, ίσο με 0,5519, γεγονός που αντιστοιχεί στον εντοπισμό του μεγαλύτερου ποσοστού πραγματικών περιπτώσεων απάτης μεταξύ των εξεταζόμενων τεχνικών συνδυασμού. Η αυξημένη αυτή κάλυψη συνοδεύεται, ωστόσο, από χαμηλό Precision, ίσο με 0,0840, και από 21.058 ψευδώς θετικές επισημάνσεις. Ο αριθμός αυτός είναι σημαντικά μεγαλύτερος σε σχέση με τις υπόλοιπες εξεταζόμενες διαμορφώσεις.

Η απαίτηση συμφωνίας περισσότερων μοντέλων στην πλειοψηφική ψηφοφορία μειώνει προοδευτικά τον αριθμό των ψευδώς θετικών επισημάνσεων. Με απαίτηση τουλάχιστον δύο θετικών ψήφων, οι ψευδώς θετικές επισημάνσεις περιορίζονται στις 9.439, ενώ με απαίτηση τουλάχιστον τριών θετικών ψήφων μειώνονται περαιτέρω στις 3.927. Η μείωση αυτή συνοδεύεται από αύξηση του Precision, αλλά και από αντίστοιχη υποχώρηση του Recall.

Η διαμόρφωση Majority Voting ≥ 3 παρουσιάζει το υψηλότερο Precision, ίσο με 0,1928, καθώς και το υψηλότερο F1-score, ίσο με 0,2243. Η επίδοση αυτή αντιστοιχεί στον ευνοϊκότερο συμβιβασμό μεταξύ Precision και Recall μεταξύ των εξεταζόμενων τεχνικών στο συγκεκριμένο σημείο λειτουργίας. Ωστόσο, το Recall περιορίζεται σε 0,2681, γεγονός που σημαίνει ότι σημαντικό ποσοστό των πραγματικών περιπτώσεων απάτης δεν εντοπίζεται.

Οι διαμορφώσεις του σταθμισμένου συνδυασμού βαθμολογιών παρουσιάζουν περισσότερο ενδιάμεση συμπεριφορά. Το Recall κυμαίνεται από 0,3861 έως 0,3967 και το Precision από 0,1351 έως 0,1388. Με τον τρόπο αυτό αποφεύγεται τόσο η ιδιαίτερα υψηλή ευαισθησία της τεχνικής OR/Union όσο και η σημαντική μείωση του Recall που παρατηρείται στη διαμόρφωση Majority Voting ≥ 3 .

Μεταξύ των τεσσάρων σταθμισμένων διαμορφώσεων, η IF-Dominant παρουσιάζει το υψηλότερο Recall, το υψηλότερο F1-score και την υψηλότερη ROC-AUC, με τιμές 0,3967, 0,2056 και 0,7121 αντίστοιχα. Αντίθετα, η Diversity Focused παρουσιάζει την υψηλότερη AUPRC/AP, ίση με 0,1123, διατηρώντας παράλληλα Recall 0,3933 και F1-score 0,2039.

Η AUPRC/AP αποκτά ιδιαίτερη σημασία στο συγκεκριμένο πρόβλημα λόγω της έντονης ανισορροπίας μεταξύ νόμιμων και δόλιων συναλλαγών. Οι τρεις τεχνικές που βασίζονται στον αριθμό θετικών ψήφων εμφανίζουν AUPRC/AP ίση με 0,1007, ενώ οι σταθμισμένες διαμορφώσεις παρουσιάζουν τιμές μεταξύ 0,1097 και 0,1123. Η υψηλότερη τιμή της Diversity Focused υποδηλώνει ελαφρώς καλύτερη συνολική ποιότητα κατάταξης των πραγματικών περιπτώσεων απάτης σε σχέση με τις υπόλοιπες εξεταζόμενες τεχνικές συνδυασμού.

Οι ίδιες τιμές AUPRC/AP και ROC-AUC που εμφανίζονται στις διαμορφώσεις OR/Union, Majority Voting ≥ 2 και Majority Voting ≥ 3 οφείλονται στο γεγονός ότι και στις τρεις περιπτώσεις η κατάταξη των συναλλαγών βασίζεται στον ίδιο αριθμό θετικών ψήφων. Η διαφοροποίηση των τεχνικών προκύπτει από την αλλαγή του απαιτούμενου αριθμού ψήφων για την τελική δυαδική απόφαση και όχι από διαφορετική κατάταξη των συναλλαγών.

Συνολικά, τα αποτελέσματα υποδηλώνουν σαφή συμβιβασμό μεταξύ της κάλυψης πραγματικών περιπτώσεων απάτης και του περιορισμού των ψευδώς θετικών επισημάνσεων. Η τεχνική OR/Union ευνοεί την αύξηση του Recall, η διαμόρφωση Majority Voting ≥ 3 δίνει μεγαλύτερη έμφαση στο Precision, ενώ οι σταθμισμένες διαμορφώσεις κινούνται μεταξύ των δύο αυτών συμπεριφορών και παρέχουν επιπλέον συνεχή βαθμολογία ανωμαλίας για κάθε συναλλαγή.

Η επιλογή της τελικής τεχνικής δεν βασίζεται επομένως αποκλειστικά στη μεγιστοποίηση μίας μεμονωμένης μετρικής. Απαιτείται η συνεκτίμηση των Precision, Recall, F1-score, AUPRC/AP και ROC-AUC, της ποιότητας κατάταξης των ύποπτων συναλλαγών και της δυνατότητας εφαρμογής της τεχνικής σε πραγματικά δεδομένα ERP. Πριν από την τελική επιλογή, στην επόμενη ενότητα εξετάζεται συμπληρωματικά η οριακή συνεισφορά των τεσσάρων επιμέρους μοντέλων μέσω Θεωρίας Παιγνίων

5.5 Ανάλυση Συνεισφοράς των Μοντέλων με Θεωρία Παιγνίων

Μετά τη συγκριτική αξιολόγηση των τεχνικών συνδυασμού πραγματοποιήθηκε συμπληρωματική ανάλυση της συνεισφοράς των Isolation Forest, LOF, HBOS και Autoencoder μέσω τιμών Shapley. Η θεωρητική βάση της μεθόδου παρουσιάστηκε στην Ενότητα 2.3.6, ενώ η διαδικασία υπολογισμού που εφαρμόστηκε στην παρούσα εργασία περιγράφηκε στην Ενότητα 3.3.4. Στην παρούσα ενότητα παρουσιάζονται και ερμηνεύονται τα αποτελέσματα της συγκεκριμένης ανάλυσης. Τα αποτελέσματα για τους επιμέρους συνασπισμούς παρουσιάζονται στον Πίνακα 5.6.

Πίνακας 5.6. Απόδοση των δυνατών συνασπισμών των τεσσάρων μοντέλων με βάση την AUPRC/AP.

Συνασπισμός	AUPRC/AP	Αξία πάνω από την τιμή αναφοράς
AE	0,1395	0,1045
IF	0,1115	0,0765
HBOS	0,1009	0,0659

LOF	0,0549	0,0199
IF + AE	0,1130	0,0780
IF + LOF	0,1114	0,0764
IF + HBOS	0,1092	0,0742
HBOS + AE	0,1031	0,0681
LOF + HBOS	0,1007	0,0657
LOF + AE	0,0694	0,0345
IF + LOF + AE	0,1128	0,0779
IF + HBOS + AE	0,1101	0,0751
IF + LOF + HBOS	0,1092	0,0742
LOF + HBOS + AE	0,1030	0,0680
IF + LOF + HBOS + AE	0,1101	0,0751

Η ανάλυση των συνασπισμών παρουσιάζει ορισμένες αξιοσημείωτες διαφοροποιήσεις. Ο Autoencoder ως μεμονωμένο μοντέλο εμφανίζει την υψηλότερη AUPRC/AP, ίση με 0,1395. Ωστόσο, η αξιολόγηση μέσω τιμών Shapley δεν περιορίζεται στην ατομική απόδοση, αλλά λαμβάνει υπόψη τη μεταβολή που προκαλεί κάθε μοντέλο όταν προστίθεται σε διαφορετικούς συνασπισμούς. Τα αποτελέσματα της συγκεκριμένης ανάλυσης παρουσιάζονται στον Πίνακα 5.7.

Πίνακας 5.7. Shapley Values των τεσσάρων μοντέλων ως προς την AUPRC/AP.

Μοντέλο	Shapley Value	Σχετική συνεισφορά
Isolation Forest	0,029018	38,63%
Autoencoder	0,028263	37,63%
HBOS	0,018750	24,96%
LOF	-0,000918	-1,22%

Το Isolation Forest παρουσιάζει τη μεγαλύτερη τιμή Shapley, ίση με 0,029018, αντιπροσωπεύοντας περίπου το 38,63% της συνολικής μεταβολής της αξίας του πλήρους συνασπισμού. Το αποτέλεσμα υποδηλώνει ότι, υπό τη συγκεκριμένη συνάρτηση αξίας, το Isolation Forest παρέχει κατά μέσο όρο τη μεγαλύτερη θετική οριακή συνεισφορά όταν προστίθεται στους διαφορετικούς συνασπισμούς των υπόλοιπων μοντέλων.

Ο Autoencoder ακολουθεί με τιμή Shapley 0,028263, πολύ κοντά στην αντίστοιχη τιμή του Isolation Forest. Παρότι ο Autoencoder εμφανίζει την υψηλότερη ατομική AUPRC/AP, η ελαφρώς μεγαλύτερη τιμή Shapley του Isolation Forest υποδηλώνει ότι το δεύτερο παρουσιάζει κατά μέσο όρο μεγαλύτερη οριακή συνεισφορά κατά τη συμμετοχή του στους διαφορετικούς δυνατούς συνασπισμούς. Επομένως, η ατομική απόδοση ενός μοντέλου και η συνεισφορά του σε ένα συνεργατικό σύστημα δεν αποτελούν ταυτόσημες έννοιες.

Το HBOS παρουσιάζει επίσης θετική συνεισφορά, με τιμή Shapley 0,018750. Η τιμή αυτή υποδηλώνει μικρότερη, αλλά θετική, μέση οριακή συνεισφορά στην AUPRC/AP των εξεταζόμενων συνασπισμών σε σχέση με τα Isolation Forest και Autoencoder.

Ιδιαίτερο ενδιαφέρον παρουσιάζει το LOF, για το οποίο προέκυψε ελαφρώς αρνητική τιμή Shapley, ίση με $-0,000918$. Η τιμή αυτή δεν σημαίνει ότι το LOF δεν εντοπίζει πραγματικές περιπτώσεις απάτης ή ότι δεν παρουσιάζει διαφοροποίηση από τα υπόλοιπα μοντέλα. Αντίθετα, η ανάλυση συμπληρωματικότητας του προηγούμενου κεφαλαίου έδειξε ότι το LOF εντοπίζει και περιπτώσεις που δεν ανιχνεύονται από άλλους αλγόριθμους. Η αρνητική τιμή αφορά αποκλειστικά τη συγκεκριμένη συνάρτηση αξίας και υποδηλώνει ότι, όταν οι κανονικοποιημένες βαθμολογίες συνδυάζονται ισότιμα, η προσθήκη του LOF προκαλεί κατά μέσο όρο μικρή μείωση της AUPRC/AP των συνασπισμών στους οποίους συμμετέχει.

Το συγκεκριμένο εύρημα αναδεικνύει μία σημαντική διάκριση μεταξύ διαφοροποίησης και ωφέλιμης συνεισφοράς. Ένα μοντέλο μπορεί να εντοπίζει διαφορετικές περιπτώσεις σε σχέση με τους υπόλοιπους ανιχνευτές, χωρίς η προσθήκη της συνεχούς βαθμολογίας του να βελτιώνει απαραίτητα τη συνολική κατάταξη των συναλλαγών. Επομένως, η περιορισμένη επικάλυψη μεταξύ δύο μοντέλων δεν συνεπάγεται αυτομάτως μεγαλύτερη απόδοση του συνδυασμού τους.

Η ορθότητα της αποσύνθεσης των τιμών Shapley ελέγχθηκε μέσω της ιδιότητας αποδοτικότητας. Η AUPRC/AP του πλήρους συνασπισμού των τεσσάρων μοντέλων είναι $0,110102$, ενώ η τιμή αναφοράς είναι $0,034990$. Συνεπώς, η συνολική αύξηση της αξίας είναι $0,110102 - 0,034990 = 0,075112$. Το άθροισμα των τεσσάρων τιμών Shapley είναι επίσης $0,029018 + 0,028263 + 0,018750 - 0,000918 = 0,075112$, γεγονός που επαληθεύει την ιδιότητα αποδοτικότητας για τη συγκεκριμένη αποσύνθεση.

Συνολικά, η ανάλυση μέσω Θεωρίας Παιγνίων υποδηλώνει ότι, με βάση τη συγκεκριμένη συνάρτηση αξίας και τον ισότιμο συνδυασμό βαθμολογιών που χρησιμοποιήθηκε για τον υπολογισμό των τιμών Shapley, τα Isolation Forest και Autoencoder παρουσιάζουν τις μεγαλύτερες θετικές μέσες οριακές συνεισφορές. Το HBOS εμφανίζει μικρότερη αλλά θετική συνεισφορά, ενώ το LOF παρουσιάζει οριακά αρνητική μέση συνεισφορά ως προς την AUPRC/AP. Τα αποτελέσματα αυτά χρησιμοποιούνται αποκλειστικά ως συμπληρωματικό ερμηνευτικό στοιχείο και όχι ως άμεσος μηχανισμός καθορισμού των βαρών της τελικής τεχνικής συνδυασμού.

5.6 Επιλογή της Τελικής Τεχνικής Συνδυασμού

Η επιλογή της τελικής τεχνικής συνδυασμού βασίστηκε στη συνεκτίμηση των αποτελεσμάτων των προηγούμενων ενοτήτων και όχι στην αποκλειστική μεγιστοποίηση μίας μεμονωμένης μετρικής. Εξετάστηκαν η κάλυψη των πραγματικών περιπτώσεων απάτης, ο αριθμός των ψευδώς θετικών επισημάνσεων, ο συμβιβασμός μεταξύ Precision και Recall, η ποιότητα της συνολικής κατάταξης των συναλλαγών και η δυνατότητα παραγωγής συνεχούς βαθμολογίας ανωμαλίας για την εφαρμογή σε δεδομένα ERP.

Η τεχνική OR/Union παρουσίασε το υψηλότερο Recall, ίσο με $0,5519$, όμως η επίδοση αυτή συνοδεύτηκε από Precision μόλις $0,0840$ και 21.058 ψευδώς θετικές επισημάνσεις. Παρότι η συγκεκριμένη τεχνική μπορεί να είναι χρήσιμη όταν προτεραιότητα αποτελεί η μεγαλύτερη δυνατή κάλυψη πιθανών περιπτώσεων απάτης, ο μεγάλος αριθμός επισημάνσεων περιορίζει την πρακτική της καταλληλότητα σε ένα σύστημα όπου οι ύποπτες συναλλαγές πρόκειται να εξετάζονται μεταγενέστερα από ανθρώπινο ελεγκτή.

Αντίθετα, η διαμόρφωση Majority Voting ≥ 3 παρουσίασε το υψηλότερο Precision, ίσο με $0,1928$, και το υψηλότερο F1-score, ίσο με $0,2243$. Η αυστηρότερη απαίτηση συμφωνίας περιορίζει σημαντικά τις ψευδώς θετικές επισημάνσεις, αλλά μειώνει το Recall σε $0,2681$. Για τον στόχο της εφαρμογής στα δεδομένα ERP, η συγκεκριμένη διαμόρφωση δεν επιλέχθηκε ως τελική, καθώς η σημαντική μείωση του Recall θα περιόριζε την κάλυψη των συναλλαγών που αναδεικνύονται για περαιτέρω διερεύνηση.

Οι διαμορφώσεις του σταθμισμένου συνδυασμού βαθμολογιών παρουσίασαν περισσότερο ενδιάμεση συμπεριφορά και, παράλληλα, παρέχουν μία συνεχή συνολική βαθμολογία ανωμαλίας για κάθε συναλλαγή. Η ιδιότητα αυτή είναι ιδιαίτερα χρήσιμη για την εφαρμογή σε πραγματικό περιβάλλον ERP χωρίς διαθέσιμες ετικέτες, καθώς επιτρέπει την κατάταξη των συναλλαγών ανάλογα με τον εκτιμώμενο βαθμό ανωμαλίας και όχι αποκλειστικά τον δυαδικό χαρακτηρισμό τους.

Μεταξύ των τεσσάρων σταθμισμένων διαμορφώσεων, η Diversity Focused παρουσίασε την υψηλότερη AUPRC/AP, ίση με $0,1123$. Ωστόσο, η IF-Dominant παρουσίασε ελαφρώς υψηλότερες τιμές Precision, Recall και F1-score, καθώς και την υψηλότερη ROC-AUC, ίση με $0,7121$. Συγκεκριμένα, η IF-Dominant κατέγραψε Precision $0,1388$, Recall $0,3967$ και F1-score $0,2056$, ενώ η Diversity Focused παρουσίασε αντίστοιχα τιμές $0,1376$, $0,3933$ και $0,2039$. Η διαφορά στην AUPRC/AP μεταξύ των δύο

διαμορφώσεων είναι σχετικά περιορισμένη $0,112305 - 0,111762 = 0,000543$, ενώ η IF-Dominant παρουσιάζει καλύτερη επίδοση στις υπόλοιπες εξεταζόμενες μετρικές. Συμπληρωματικά, η ανάλυση μέσω τιμών Shapley έδειξε ότι, υπό τη συγκεκριμένη συνάρτηση αξίας, το Isolation Forest παρουσίασε τη μεγαλύτερη μέση οριακή συνεισφορά, ενώ το LOF παρουσίασε οριακά αρνητική μέση συνεισφορά ως προς την AUPRC/AP.

Η ανάλυση μέσω Θεωρίας Παιγνίων δεν χρησιμοποιήθηκε για τον άμεσο καθορισμό των βαρών και, επομένως, οι τιμές Shapley δεν μπορούν να θεωρηθούν ως επιβεβαίωση της συγκεκριμένης αριθμητικής κατανομής της διαμόρφωσης IF-Dominant. Παρέχουν, ωστόσο, μία επιπλέον ένδειξη ότι η αυξημένη συμμετοχή του Isolation Forest είναι συμβατή με τη σημαντική οριακή συνεισφορά που παρουσιάζει το συγκεκριμένο μοντέλο στους διαφορετικούς συνασπισμούς.

Με βάση τη συνολική συνεκτίμηση των παραπάνω αποτελεσμάτων, η διαμόρφωση IF-Dominant επιλέχθηκε ως η τελική τεχνική συνδυασμού για την εφαρμογή στα δεδομένα ERP, με βάρη

$$w_{IF}=0.50, w_{HBOS}=0.15, w_{LOF}=0.15, w_{AE}=0.20 \quad (5.3)$$

Η επιλογή δεν βασίζεται σε καθολική υπεροχή της συγκεκριμένης διαμόρφωσης σε όλες τις μετρικές, αλλά στη συνολική και σχετικά ισορροπημένη συμπεριφορά της ως προς Precision, Recall, F1-score και ROC-AUC, σε συνδυασμό με τη δυνατότητα παραγωγής συνεχούς βαθμολογίας ανωμαλίας. Παράλληλα, η επιλογή αφορά αποκλειστικά το συγκεκριμένο πειραματικό πλαίσιο και δεν υποδηλώνει ότι η συγκεκριμένη κατανομή βαρών αποτελεί γενικά βέλτιστη λύση για κάθε σύνολο οικονομικών δεδομένων.

Ο τελικός σταθμισμένος συνδυασμός βαθμολογιών παρέχει για κάθε συναλλαγή μία ενιαία συνεχή βαθμολογία ανωμαλίας, η οποία επιτρέπει την ιεράρχηση των συναλλαγών από περισσότερο προς λιγότερο ύποπτες. Η ιδιότητα αυτή, σε συνδυασμό με τη συνολική συμπεριφορά της διαμόρφωσης IF-Dominant, υποστηρίζει την επιλογή της για την εφαρμογή στο πραγματικό σύνολο δεδομένων ERP, το οποίο δεν διαθέτει γνωστές ετικέτες απάτης. Στο επόμενο κεφάλαιο εξετάζεται η εφαρμογή της τελικής τεχνικής στα δεδομένα ERP και η συμπληρωματική διερεύνηση των αποτελεσμάτων μέσω του Νόμου του Benford.

Κεφάλαιο 6ο: Εφαρμογή σε Πραγματικά Δεδομένα ERP και Ανάλυση Benford

Μετά την πειραματική αξιολόγηση των επιμέρους μοντέλων ανίχνευσης ανωμαλιών και των διαφορετικών τεχνικών συνδυασμού στο σύνολο δεδομένων IEEE-CIS, στο παρόν κεφάλαιο εξετάζεται η εφαρμογή του τελικού πλαισίου σε πραγματικά λογιστικά δεδομένα προερχόμενα από σύστημα ERP. Το στάδιο αυτό διαφοροποιείται ουσιαστικά από τα προηγούμενα πειράματα, καθώς δεν αποσκοπεί σε νέα επιλογή ή βελτιστοποίηση μοντέλων, αλλά στη διερεύνηση της δυνατότητας εφαρμογής της μεθοδολογίας σε πραγματικές συνθήκες.

Η βασική διαφοροποίηση του συνόλου ERP αφορά την απουσία πραγματικών ετικετών που να προσδιορίζουν εάν μία λογιστική εγγραφή αντιστοιχεί σε επιβεβαιωμένη περίπτωση απάτης. Για τον λόγο αυτό, δεν είναι δυνατός ο υπολογισμός μετρικών όπως Precision, Recall, F1-score ή AUPRC με τρόπο αντίστοιχο προς το IEEE-CIS. Η αξιολόγηση μετατοπίζεται επομένως από την πρόβλεψη μιας γνωστής κλάσης προς την παραγωγή βαθμολογιών κινδύνου και την ιεράρχηση των λογιστικών εγγραφών για πιθανό περαιτέρω έλεγχο. Η προσέγγιση αυτή είναι συμβατή με τη χρήση τεχνικών ανίχνευσης ανωμαλιών σε δεδομένα Γενικού Καθολικού, όπου μία ασυνήθιστη εγγραφή μπορεί να σχετίζεται τόσο με σφάλμα ή ασυνήθιστη επιχειρησιακή δραστηριότητα όσο και με περίπτωση που απαιτεί βαθύτερη διερεύνηση [20].

Στο πρώτο στάδιο εφαρμόζεται στα δεδομένα ERP το σταθμισμένο συνδυαστικό μοντέλο που επιλέχθηκε κατά την προηγούμενη πειραματική διαδικασία. Ακολουθεί ανάλυση των λογιστικών εγγραφών που έλαβαν τις υψηλότερες βαθμολογίες κινδύνου, με έμφαση στα χρηματικά ποσά, στους λογαριασμούς και στους τύπους λογιστικών κινήσεων. Στη συνέχεια εφαρμόζεται ο Νόμος του Benford στα κατάλληλα οικονομικά μεγέθη και εξετάζεται η κατανομή του πρώτου σημαντικού ψηφίου. Τέλος, τα αποτελέσματα των δύο προσεγγίσεων εξετάζονται συνδυαστικά, χωρίς η ανάλυση Benford να αντιμετωπίζεται ως δεύτερος ανεξάρτητος μηχανισμός ανίχνευσης απάτης.

6.1 Εφαρμογή και Ανάλυση των Αποτελεσμάτων στα Δεδομένα ERP

Μετά την ολοκλήρωση της διαδικασίας προεπεξεργασίας, το τελικό σύνολο δεδομένων ERP περιλάμβανε 65.479 λογιστικές εγγραφές και 15 χαρακτηριστικά που χρησιμοποιήθηκαν ως είσοδοι στα μοντέλα ανίχνευσης ανωμαλιών. Καθώς δεν υπήρχε πραγματική μεταβλητή κατάστασης που να διαχωρίζει τις κανονικές από τις δόλιες εγγραφές, η εφαρμογή επικεντρώθηκε στον υπολογισμό συνεχών βαθμολογιών ανωμαλίας και στη μετέπειτα σύνθεσή τους σε μία ενιαία βαθμολογία κινδύνου. Οι τέσσερις επιμέρους ανιχνευτές, Isolation Forest, LOF, HBOS και Autoencoder, προσαρμόστηκαν στα δεδομένα ERP χρησιμοποιώντας τις διαμορφώσεις που είχαν επιλεγεί κατά την πειραματική αξιολόγηση στο IEEE-CIS. Δεν πραγματοποιήθηκε νέα διαδικασία βελτιστοποίησης υπερπαραμέτρων στο ERP. Η επιλογή αυτή διατηρεί τον χαρακτήρα της εφαρμογής ως εξωτερικής διερεύνησης του ήδη διαμορφωμένου πλαισίου και αποφεύγει την εκ νέου προσαρμογή των ρυθμίσεων σε ένα σύνολο στο οποίο δεν υπάρχουν πραγματικές ετικέτες για αντικειμενική σύγκριση.

Για κάθε λογιστική εγγραφή παρήχθη μία βαθμολογία ανωμαλίας από κάθε μοντέλο. Οι τέσσερις βαθμολογίες κανονικοποιήθηκαν στο διάστημα $[0,1]$ και συνδυάστηκαν σύμφωνα με τη σταθμισμένη στρατηγική που είχε επιλεγεί στο Κεφάλαιο 5. Η τελική βαθμολογία κινδύνου υπολογίστηκε ως:

$$RiskScore = 0.50S_{IF} + 0.15S_{HBOS} + 0.15S_{LOF} + 0.20S_{AE} \quad (6.1)$$

όπου S_{IF} , S_{HBOS} , S_{LOF} και S_{AE} αντιστοιχούν στις κανονικοποιημένες βαθμολογίες των Isolation Forest, HBOS, LOF και Autoencoder αντίστοιχα. Τα βάρη διατηρήθηκαν αμετάβλητα από την πειραματική διαδικασία στο IEEE-CIS.

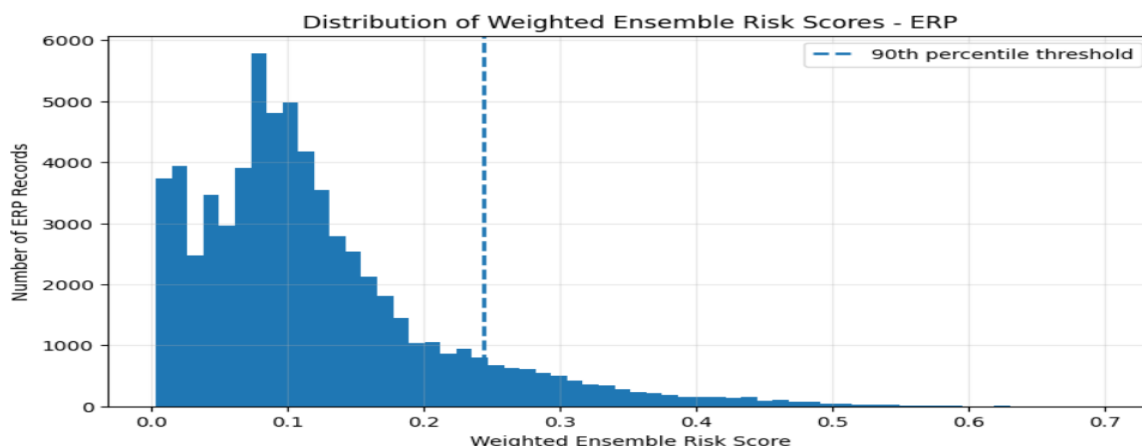
Η εφαρμογή του τελικού συνδυαστικού μοντέλου παρήγαγε μία βαθμολογία κινδύνου για καθεμία από τις 65.479 εγγραφές. Οι τιμές βαθμολογίας κινδύνου κυμάνθηκαν από 0,0034 έως 0,6997, με μέση τιμή 0,1204 και διάμεσο 0,1000. Η κατανομή χαρακτηρίζεται από συγκέντρωση των περισσότερων εγγραφών στις χαμηλότερες τιμές και από μικρότερο αριθμό εγγραφών με υψηλότερες βαθμολογίες.

Πίνακας 6.1. Βασικά χαρακτηριστικά της κατανομής της βαθμολογίας κινδύνου

Μέγεθος	Τιμή
Αριθμός εγγραφών	65.479
Ελάχιστη βαθμολογία	0,0034
Μέση βαθμολογία	0,1204
Διάμεσος	0,1000
90ό εκατοστημόριο	0,2435
95ό εκατοστημόριο	0,3077
99ό εκατοστημόριο	0,4387
Μέγιστη βαθμολογία	0,6997

Ως λειτουργικό κατώφλι επιλέχθηκε το 90ό εκατοστημόριο της κατανομής της βαθμολογίας κινδύνου(RiskScore), το οποίο αντιστοιχεί σε τιμή 0,2435. Με βάση το συγκεκριμένο κατώφλι επισημάνθηκαν 6.548 εγγραφές, δηλαδή περίπου το 10% του συνολικού συνόλου. Οι εγγραφές αυτές δεν αντιμετωπίζονται ως επιβεβαιωμένες περιπτώσεις απάτης, αλλά ως περιπτώσεις αυξημένης προτεραιότητας για περαιτέρω διερεύνηση.

Η συγκεκριμένη αναλογία δεν πρέπει να ερμηνεύεται ως εκτίμηση ότι το 10% των λογιστικών εγγραφών αποτελεί οικονομική απάτη. Το ποσοστό προκύπτει από το προκαθορισμένο λειτουργικό σημείο του συστήματος και εκφράζει το τμήμα του συνόλου που επιλέγεται για περαιτέρω έλεγχο. Η συνεχής βαθμολογία κινδύνου παρουσιάζει επομένως μεγαλύτερη χρησιμότητα από τη δυαδική επισήμανση, καθώς επιτρέπει την ιεράρχηση των εγγραφών ανάλογα με τη σχετική προτεραιότητά τους. Για λεπτομερέστερη διερεύνηση εξετάστηκε επίσης το ανώτερο 1% της κατανομής. Το αντίστοιχο κατώφλι ήταν 0,4387 και περιλάμβανε 655 λογιστικές εγγραφές. Η κατηγορία αυτή μπορεί να θεωρηθεί ως το σύνολο πολύ υψηλής προτεραιότητας, χωρίς όμως ο χαρακτηρισμός αυτός να συνεπάγεται επιβεβαιωμένη παρατυπία.



Σχήμα 6.1. Κατανομή των τελικών βαθμολογιών κινδύνου στις λογιστικές εγγραφές ERP και λειτουργικό κατώφλι του 90ού εκατοστημορίου.

Ενδιαφέρον παρουσιάζει και η αριθμητική συνεισφορά των τεσσάρων επιμέρους μοντέλων στις εγγραφές που υπερέβησαν το τελικό κατώφλι. Παρότι τα αρχικά βάρη του συνδυασμού ήταν 50% για το Isolation Forest, 15% για το HBOS και 15% για το LOF και 20% για το Autoencoder, η πραγματική μέση συνεισφορά στις επισημασμένες εγγραφές διαφοροποιήθηκε σημαντικά.

Πίνακας 6.2. Μέση πραγματική συνεισφορά των επιμέρους μοντέλων στις επισημασμένες εγγραφές

Μοντέλο	Ποσοστό μέσης συνεισφοράς
Isolation Forest	86,85%
HBOS	12,95%
Autoencoder	0,19%
LOF	0,01%

Η ανάλυση της σταθμισμένης συνεισφοράς στις 6.548 επισημασμένες εγγραφές έδειξε έντονη συγκέντρωση στο Isolation Forest, το οποίο αντιστοιχούσε κατά μέσο όρο στο 86,85% της τελικής βαθμολογίας. Το HBOS ακολούθησε με 12,95%, ενώ οι αντίστοιχες μέσες συνεισφορές του Autoencoder και του LOF ήταν σημαντικά μικρότερες. Επιπλέον, το Isolation Forest αποτέλεσε το μοντέλο με τη μεγαλύτερη επιμέρους σταθμισμένη συνεισφορά και στις 6.548 επισημασμένες εγγραφές.

Η παρατήρηση αυτή πρέπει να ερμηνευθεί σε συνάρτηση τόσο με τα βάρη της επιλεγμένης διαμόρφωσης όσο και με την κατανομή των κανονικοποιημένων βαθμολογιών των μοντέλων. Επομένως, δεν μπορεί να θεωρηθεί από μόνη της ένδειξη ότι το Isolation Forest είναι περισσότερο ακριβές στα ERP δεδομένα, καθώς δεν υπάρχουν πραγματικές ετικέτες για την επαλήθευση των επισημάνσεων.

6.2 Ανάλυση των Υποπτων Συναλλαγών

Η ανάλυση των 6.548 εγγραφών που βρέθηκαν πάνω από το λειτουργικό κατώφλι πραγματοποιήθηκε με στόχο την αναζήτηση κοινών λογιστικών και οικονομικών χαρακτηριστικών. Η απουσία πραγματικών ετικετών καθιστά αναγκαία την προσεκτική ερμηνεία των αποτελεσμάτων. Οι εγγραφές αυτές αντιμετωπίζονται ως περιπτώσεις αυξημένης προτεραιότητας και όχι ως επιβεβαιωμένες περιπτώσεις απάτης.

Ως προς τον τύπο λογιστικής εγγραφής, η μεγαλύτερη συγκέντρωση παρατηρήθηκε στις συμψηφιστικές εγγραφές. Συγκεκριμένα, 4.303 εγγραφές, δηλαδή περίπου 65,71% του επισημασμένου συνόλου, ανήκαν στην κατηγορία των συμψηφιστικών κινήσεων. Ακολούθησαν οι συμψηφιστικές εγγραφές

αγορών με 1.328 περιπτώσεις (20,28%), οι συμψηφιστικές εγγραφές πωλήσεων με 577 περιπτώσεις (8,81%) και οι ταμειακές εγγραφές με 263 περιπτώσεις (4,02%).

Πίνακας 6.3. Κατανομή των επισημασμένων εγγραφών ανά τύπο

Τύπος εγγραφής	Πλήθος	Ποσοστό επί των επισημασμένων
Συμψηφιστική	4.303	65,71%
Συμψηφιστική αγορών	1.328	20,28%
Συμψηφιστική πωλήσεων	577	8,81%
Ταμειακή	263	4,02%
Συμψηφισμός αξιογράφων	76	1,16%
Συμψηφιστική λιανικών πωλήσεων	1	0,02%

Η υψηλή παρουσία μιας κατηγορίας μεταξύ των επισημασμένων εγγραφών δεν συνεπάγεται από μόνη της ότι ο συγκεκριμένος τύπος είναι περισσότερο προβληματικός. Μέρος της συγκέντρωσης ενδέχεται να συνδέεται με τη συχνότητα με την οποία εμφανίζεται κάθε τύπος στο αρχικό σύνολο. Για τον λόγο αυτό, οι παραπάνω τιμές ερμηνεύονται κυρίως ως περιγραφικά χαρακτηριστικά των αποτελεσμάτων και όχι ως απόδειξη αυξημένου κινδύνου συγκεκριμένης λογιστικής κατηγορίας.

Ανάλογη συγκέντρωση παρατηρήθηκε και σε συγκεκριμένες κατηγορίες λογαριασμών. Οι λογαριασμοί πελατών εσωτερικού χονδρικής αντιστοιχούσαν σε 2.129 επισημασμένες εγγραφές, δηλαδή περίπου στο 32,51% του επισημασμένου συνόλου, ενώ ο λογαριασμός ταμείου εμφανίστηκε σε 1.595 περιπτώσεις (24,36%). Ακολούθησαν λογαριασμοί προμηθευτών εσωτερικού, τραπεζικών καταθέσεων, ΦΠΑ αγορών και πιστωτών. Η συγκέντρωση αυτή δεν συνεπάγεται από μόνη της αυξημένη πιθανότητα απάτης για τους συγκεκριμένους λογαριασμούς, καθώς μπορεί να επηρεάζεται και από τη συχνότητα εμφάνισής τους στο αρχικό σύνολο δεδομένων.

Η εξέταση των χρηματικών χαρακτηριστικών υποδηλώνει σημαντική διασπορά των ποσών. Στις επισημασμένες εγγραφές η μέγιστη τιμή χρέωσης και πίστωσης έφθασε τις 290.000, ενώ καταγράφηκαν και αρνητικές τιμές χρεώσεων και πιστώσεων. Οι αρνητικές τιμές πιθανώς σχετίζονται με αντιστοιχισμούς, διορθωτικές ή άλλες ειδικές λογιστικές κινήσεις, χωρίς ωστόσο να είναι δυνατή η ασφαλής εξακρίβωση της οικονομικής τους αιτίας μόνο από τα διαθέσιμα χαρακτηριστικά.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η διαφοροποίηση των χρηματικών μεγεθών μεταξύ των εγγραφών με τις υψηλότερες βαθμολογίες κινδύνου. Στις πρώτες θέσεις της κατάταξης εμφανίστηκαν κινήσεις ύψους 290.000, 172.997,82, 121.245,80 και 111.214,36, αλλά και σημαντικά χαμηλότερα ποσά, όπως 6.200 και 12.876,31. Η εικόνα αυτή υποδηλώνει ότι η τελική κατάταξη δεν καθορίζεται αποκλειστικά από το ύψος του χρηματικού ποσού, αλλά από τον συνδυασμό των χαρακτηριστικών που αξιοποιούνται από τους επιμέρους ανιχνευτές.

Οι εγγραφές υψηλότερης προτεραιότητας περιλάμβαναν κινήσεις που σχετίζονταν με τραπεζικές καταθέσεις, προμηθευτές, ταμειακές κινήσεις, ΦΠΑ και άλλους λογαριασμούς. Στην παρουσίαση των αποτελεσμάτων οι αναγνωριστικοί αριθμοί παραστατικών και οι πλήρεις τραπεζικοί λογαριασμοί ανωνυμοποιήθηκαν, ώστε να διατηρείται η απαραίτητη επιχειρησιακή πληροφορία χωρίς να αποκαλύπτονται μη αναγκαία στοιχεία του πραγματικού συστήματος ERP.

Συνολικά, η εφαρμογή του συνδυαστικού μοντέλου φαίνεται να παρέχει μία πρακτική διαδικασία περιορισμού του αρχικού όγκου των δεδομένων. Από τις 65.479 διαθέσιμες εγγραφές προκύπτει ένα σύνολο 6.548 εγγραφών αυξημένης προτεραιότητας και, σε αυστηρότερο επίπεδο, 655 εγγραφών πολύ υψηλής προτεραιότητας. Η χρησιμότητα αυτής της διαδικασίας βρίσκεται κυρίως στην ιεράρχηση του ελέγχου και όχι στην αυτόματη απόδοση χαρακτηρισμού απάτης [20].

6.3 Ανάλυση με τον Νόμο του Benford

Μετά την εφαρμογή του τελικού συνδυαστικού μοντέλου εξετάστηκε η αριθμητική δομή των οικονομικών ποσών σύμφωνα με τον Νόμο του Benford. Η ανάλυση βασίστηκε στο πρώτο σημαντικό ψηφίο των χρηματικών μεγεθών. Η θεωρητική κατανομή του πρώτου σημαντικού ψηφίου σύμφωνα με τον Benford είναι λογαριθμική και χρησιμοποιείται ως σημείο αναφοράς για τη σύγκριση παρατηρούμενων οικονομικών δεδομένων [44]–[49].

Για κάθε λογιστική εγγραφή χρησιμοποιήθηκε το απόλυτο χρηματικό ποσό της κίνησης, το οποίο προέκυψε από τη μη μηδενική τιμή μεταξύ χρέωσης και πίστωσης. Από τις 65.479 συνολικές εγγραφές, 65.435 διέθεταν θετικό χρηματικό ποσό και μπορούσαν να χρησιμοποιηθούν στην ανάλυση. Οι υπόλοιπες 44 εγγραφές είχαν μηδενική χρέωση και πίστωση και συνεπώς δεν διέθεταν πρώτο σημαντικό ψηφίο. Από αυτές, 18 είχαν επισημανθεί από το συνδυαστικό μοντέλο. Κατά συνέπεια, η ανάλυση Benford πραγματοποιήθηκε σε 6.530 επισημασμένες και 58.905 μη επισημασμένες επιλέξιμες εγγραφές.

Τα διαθέσιμα θετικά ποσά εκτείνονταν από 0,01 έως 290.000, καλύπτοντας περίπου 7,46 τάξεις μεγέθους. Η ευρεία διασπορά των χρηματικών τιμών δημιουργεί ένα κατάλληλο πλαίσιο για διερευνητική ανάλυση πρώτου σημαντικού ψηφίου, χωρίς αυτό να σημαίνει ότι η συμμόρφωση με τον Νόμο του Benford πρέπει να θεωρείται εκ των προτέρων δεδομένη.

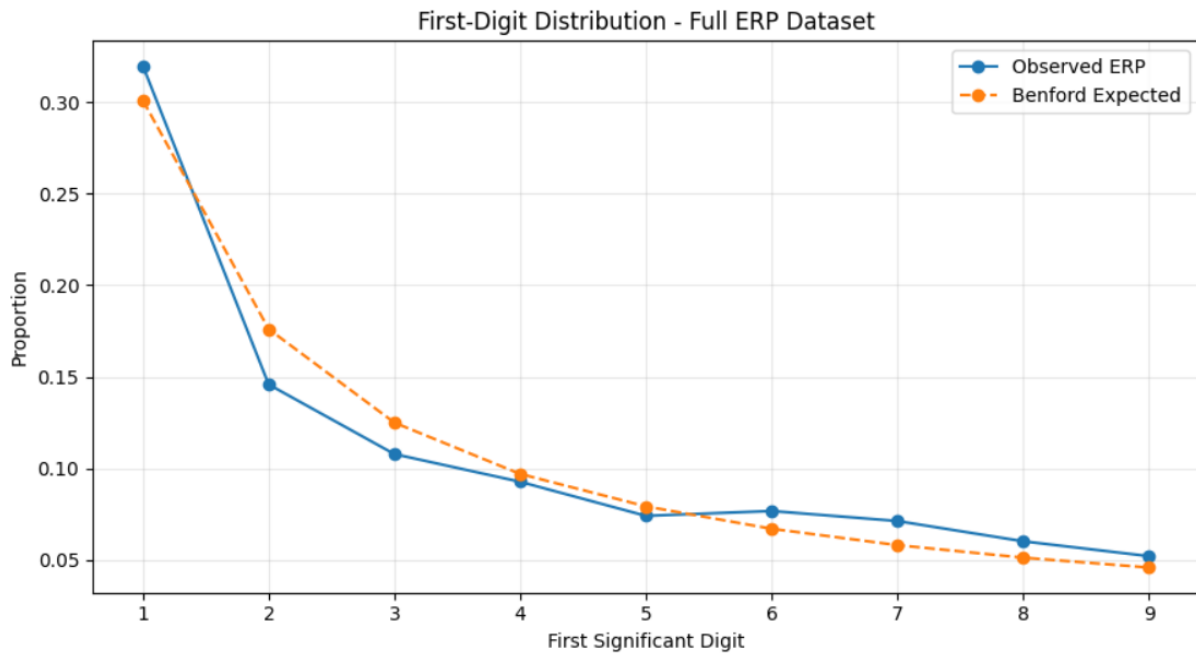
Η παρατηρούμενη κατανομή συγκρίθηκε με τη θεωρητική κατανομή τόσο για το συνολικό σύνολο ERP όσο και ξεχωριστά για τις εγγραφές που επισημάνθηκαν και δεν επισημάνθηκαν από το συνδυαστικό μοντέλο.

Πίνακας 6.4. Κατανομή πρώτου σημαντικού ψηφίου

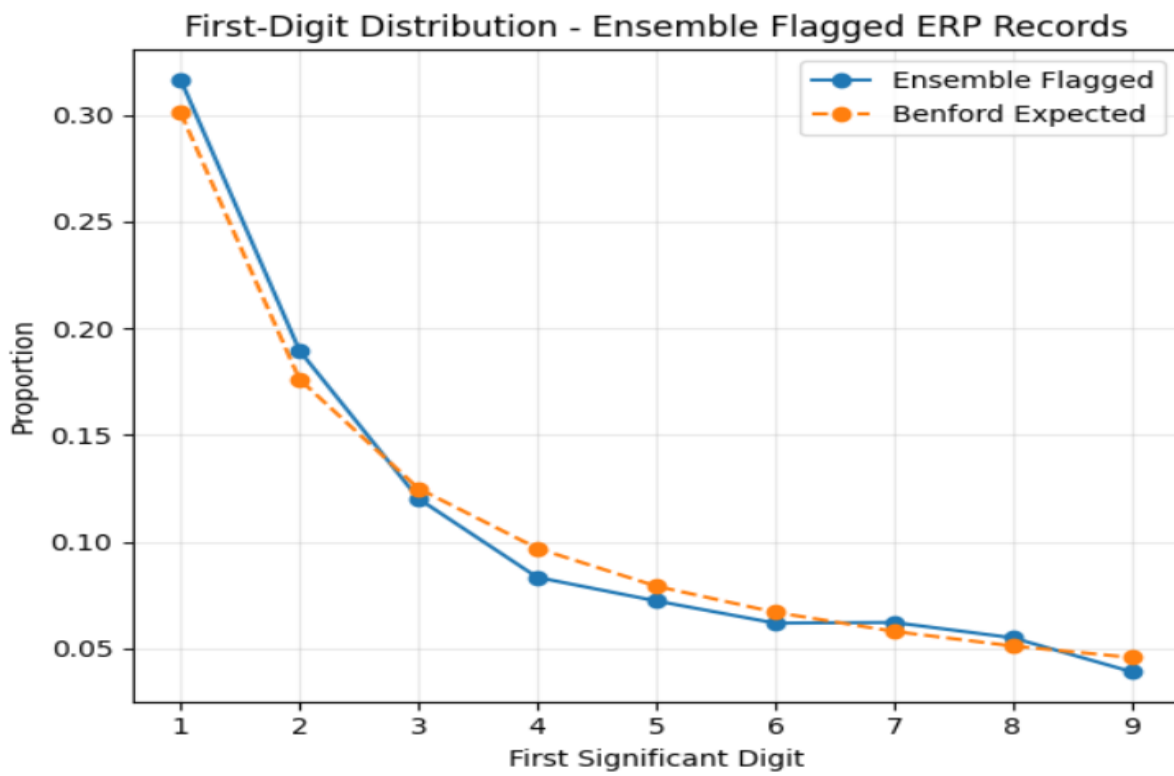
Ψηφίο	Αναμενόμενο Benford	Σύνολο ERP	Επισημασμένες	Μη επισημασμένες
1	30,10%	31,99%	31,65%	32,03%
2	17,61%	14,58%	18,96%	14,10%
3	12,49%	10,78%	12,02%	10,64%
4	9,69%	9,26%	8,33%	9,37%
5	7,92%	7,40%	7,23%	7,42%
6	6,69%	7,67%	6,19%	7,83%
7	5,80%	7,12%	6,22%	7,22%
8	5,12%	6,01%	5,50%	6,07%
9	4,58%	5,20%	3,91%	5,34%

Στο συνολικό σύνολο δεδομένων παρατηρείται μεγαλύτερη παρουσία του ψηφίου 1 από την αναμενόμενη, ενώ τα ψηφία 2 και 3 εμφανίζονται λιγότερο συχνά. Αντίθετα, στο σύνολο που επισημάνθηκε από το συνδυαστικό μοντέλο, η συχνότητα του ψηφίου 2 ανήλθε σε 18,99%, αρκετά κοντά αλλά ελαφρώς υψηλότερα από την αναμενόμενη τιμή 17,61%.

Στις επισημασμένες εγγραφές, οι μεγαλύτερες απόλυτες αποκλίσεις από τη θεωρητική κατανομή παρατηρήθηκαν στα ψηφία 1, 4 και 2. Το ψηφίο 1 παρουσίασε θετική απόκλιση περίπου +1,55 ποσοστιαίων μονάδων, το ψηφίο 4 αρνητική απόκλιση περίπου −1,36 ποσοστιαίων μονάδων και το ψηφίο 2 θετική απόκλιση περίπου +1,35 ποσοστιαίων μονάδων.



Σχήμα 6.2. Παρατηρούμενη και αναμενόμενη κατανομή πρώτου σημαντικού ψηφίου στο συνολικό σύνολο ERP.



Σχήμα 6.3. Παρατηρούμενη κατανομή πρώτου σημαντικού ψηφίου στις εγγραφές που επισημάνθηκαν από το συνδυαστικό μοντέλο και θεωρητική κατανομή Benford.

Για την ποσοτικοποίηση της συνολικής απόκλισης χρησιμοποιήθηκε η Μέση Απόλυτη Απόκλιση (MAD), ενώ συμπληρωματικά εφαρμόστηκε ο έλεγχος χ^2 . Ο τελευταίος αποτελεί καθιερωμένο τρόπο σύγκρισης παρατηρούμενων και αναμενόμενων συχνοτήτων σε εφαρμογές Benford [44].

Πίνακας 6.5. Αποτελέσματα ανάλυσης Benford

Σύνολο	Πλήθος	MAD	χ^2	p-value
Συνολικό ERP	65.435	0,01265	1.053,68	<0,001
Επισημασμένες εγγραφές	6.530	0,00822	42,30	<0,001
Μη επισημασμένες εγγραφές	58.905	0,01377	1.170,44	<0,001

Οι πολύ μικρές τιμές p υποδηλώνουν στατιστικά ανιχνεύσιμες αποκλίσεις από τη θεωρητική κατανομή και στα τρία σύνολα. Ωστόσο, λόγω του μεγάλου αριθμού παρατηρήσεων, η στατιστική σημαντικότητα δεν πρέπει να ταυτίζεται με μεγάλη πρακτική απόκλιση. Για τον λόγο αυτό, η MAD χρησιμοποιείται συμπληρωματικά για την αξιολόγηση του μεγέθους των διαφορών.

Η MAD των επισημασμένων εγγραφών διαμορφώθηκε σε 0,00822, χαμηλότερα τόσο από τη MAD του συνολικού συνόλου (0,01265) όσο και από εκείνη των μη επισημασμένων εγγραφών (0,01377). Επομένως, οι εγγραφές που χαρακτηρίστηκαν ως περισσότερο ασυνήθιστες από την τεχνική συνδυασμού δεν παρουσίασαν εντονότερη συνολική απόκλιση από τον Νόμο του Benford. Αντίθετα, ως σύνολο εμφάνισαν μεγαλύτερη εγγύτητα προς τη θεωρητική συνεισφορά του πρώτου σημαντικού ψηφίου.

Το αποτέλεσμα αυτό δεν φαίνεται να υποστηρίζει την υπόθεση ότι υψηλότερη βαθμολογία κινδύνου συνεπάγεται και μεγαλύτερη απόκλιση από τον Benford. Αντίθετα, είναι συμβατό με την ερμηνεία ότι οι δύο προσεγγίσεις εξετάζουν διαφορετικές διαστάσεις της δομής των δεδομένων.

6.4 Συνδυαστική Ερμηνεία του Μοντέλου και της Ανάλυσης Benford

Η συνδυαστική εξέταση των αποτελεσμάτων πραγματοποιήθηκε με σκοπό να διερευνηθεί κατά πόσο η στατιστική πληροφορία του Benford συμπίπτει ή συμπληρώνει την ιεράρχηση που παράγει το συνδυαστικό μοντέλο. Η συγκεκριμένη ανάλυση δεν μετατρέπει τον Benford σε πέμπτο ανιχνευτή ούτε μεταβάλλει την τελική βαθμολογία κινδύνου. Αντίθετα, χρησιμοποιείται ως πρόσθετο επίπεδο αριθμητικής ερμηνείας.

Η πρώτη παρατήρηση αφορά τη συνολική συμφωνία των δύο προσεγγίσεων. Εάν οι εγγραφές υψηλότερου κινδύνου παρουσίαζαν συστηματικά μεγαλύτερη απόκλιση από τον Benford, θα αναμενόταν η MAD του επισημασμένου συνόλου να είναι υψηλότερη. Στην παρούσα εφαρμογή παρατηρήθηκε το αντίθετο, καθώς η MAD των επισημασμένων εγγραφών ήταν χαμηλότερη από εκείνη των μη επισημασμένων. Συνεπώς, δεν προκύπτει ένδειξη γενικευμένης σύμπτωσης μεταξύ υψηλής βαθμολογίας κινδύνου και αυξημένης απόκλισης πρώτου ψηφίου.

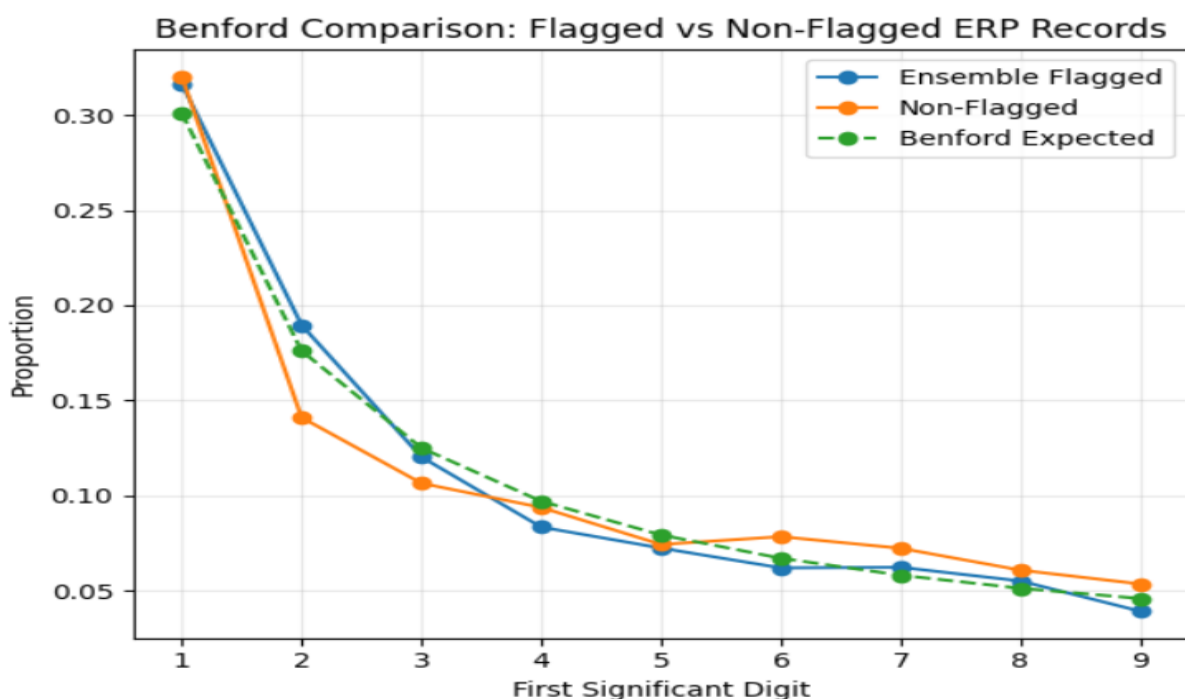
Παρά τη συγκεκριμένη εικόνα σε συνολικό επίπεδο, εμφανίζονται διαφοροποιήσεις όταν εξετάζεται κάθε ψηφίο ξεχωριστά. Το συνολικό ποσοστό επισήμανσης μεταξύ των εγγραφών που μπορούσαν να συμμετάσχουν στην ανάλυση Benford ήταν περίπου 9,98%. Για τις εγγραφές με πρώτο σημαντικό ψηφίο 2, το ποσοστό επισήμανσης αυξήθηκε σε 12,98%. Η σχετική αναλογία ως προς το συνολικό ποσοστό, δηλαδή ο λόγος αύξησης, ήταν 1,300. Επομένως, οι εγγραφές με πρώτο σημαντικό ψηφίο 2 παρουσίασαν ποσοστό επισήμανσης περίπου 30% υψηλότερο από το συνολικό ποσοστό επισήμανσης των εγγραφών που συμμετείχαν στην ανάλυση Benford.

Πίνακας 6.6. Ποσοστό επισήμανσης ανά πρώτο σημαντικό ψηφίο

Ψηφίο	Ποσοστό Επισήμανσης	Σχετικός Λόγος
1	9,87%	0,990
2	12,98%	1,300
3	11,13%	1,116
4	8,98%	0,899
5	9,75%	0,977
6	8,05%	0,807
7	8,72%	0,873
8	9,13%	0,915
9	7,50%	0,751

Το ψηφίο 3 παρουσίασε επίσης αυξημένη αναλογία επισήμανσης, με ποσοστό 11,13% και σχετικό λόγο 1,116. Αντίθετα, τα ψηφία 6, 7 και 9 εμφάνισαν χαμηλότερες αναλογίες επισήμανσης από τη συνολική τιμή αναφοράς. Οι διαφοροποιήσεις αυτές υποδηλώνουν κάποια σχέση μεταξύ της αριθμητικής δομής των ποσών και της κατάταξης του συνδυαστικού μοντέλου, χωρίς όμως να επαρκούν για την απόδοση αιτιώδους ή ελεγκτικής ερμηνείας.

Για τη συνολική διερεύνηση της σχέσης μεταξύ πρώτου σημαντικού ψηφίου και επισήμανσης εφαρμόστηκε έλεγχος χ^2 ανεξαρτησίας. Προέκυψε τιμή $\chi^2 = 168,59$ με οκτώ βαθμούς ελευθερίας και $p < 0,001$. Παρότι το αποτέλεσμα υποδηλώνει στατιστικά ανιχνεύσιμη σχέση, ο συντελεστής Cramér's V ήταν μόλις 0,0508, γεγονός που υποδηλώνει ότι το μέγεθος της σχέσης είναι ασθενές.



Σχήμα 6.4. Σύγκριση της κατανομής πρώτου σημαντικού ψηφίου μεταξύ επισημασμένων και μη επισημασμένων εγγραφών και της θεωρητικής κατανομής Benford.

Η ανάλυση αυτή απαντά και στο ερώτημα κατά πόσο ο Benford προσθέτει πληροφορία που δεν παρέχει το συνδυαστικό μοντέλο. Η διαφοροποίηση των συχνοτήτων ανά ψηφίο παρέχει ένα πρόσθετο στατιστικό πλαίσιο για την κατανόηση της αριθμητικής δομής των εγγραφών. Ωστόσο, η πληροφορία αυτή είναι διαφορετικής φύσης από τη βαθμολογία κινδύνου. Η βαθμολογία του συνδυαστικού μοντέλου υπολογίζεται σε επίπεδο μεμονωμένης εγγραφής, ενώ η απόκλιση Benford αποτελεί ιδιότητα μιας κατανομής ή ομάδας εγγραφών. Για τον λόγο αυτό, δεν είναι μεθοδολογικά ορθό να οριστεί μία συναλλαγή ως «ανωμαλία Benford» αποκλειστικά επειδή το πρώτο ψηφίο της ανήκει σε κατηγορία που παρουσιάζει συνολική απόκλιση.

Η διαφορά αυτή γίνεται εμφανής και στις 44 εγγραφές μηδενικού χρηματικού ποσού, για τις οποίες δεν μπορούσε να υπολογιστεί πρώτο σημαντικό ψηφίο. Από αυτές, 18 είχαν επισημανθεί από το συνδυαστικό μοντέλο. Το γεγονός αυτό υποδηλώνει ότι το μοντέλο μπορεί να αναδεικνύει περιπτώσεις με ασυνήθιστα χαρακτηριστικά ακόμη και όταν η ανάλυση Benford δεν μπορεί να εφαρμοστεί.

Από την άλλη πλευρά, μία ομάδα συναλλαγών μπορεί να συμβάλλει σε απόκλιση από τη θεωρητική κατανομή Benford χωρίς οι επιμέρους εγγραφές της να λαμβάνουν υψηλές βαθμολογίες κινδύνου. Συνεπώς, οι δύο προσεγγίσεις φαίνεται να λειτουργούν περισσότερο συμπληρωματικά παρά ως δύο μηχανισμοί που αναμένεται να καταλήγουν στα ίδια αποτελέσματα.

Η χρήση του Benford μπορεί επίσης να ενισχύσει την ερμηνευσιμότητα σε επίπεδο ομάδας. Για παράδειγμα, η αυξημένη αναλογία επισήμανσης στις εγγραφές με πρώτο ψηφίο 2 μπορεί να αποτελέσει ένα πρόσθετο στοιχείο για στοχευμένη εξέταση της συγκεκριμένης κατηγορίας. Δεν μπορεί όμως να υποστηριχθεί ότι η ύπαρξη πρώτου ψηφίου 2 αποτελεί ένδειξη απάτης από μόνη της.

Η συγκεκριμένη επιφύλαξη είναι ιδιαίτερα σημαντική. Η απόκλιση από τον Νόμο του Benford μπορεί να σχετίζεται με τη φυσική δομή των οικονομικών δεδομένων, με λογιστικούς κανόνες, περιορισμένα εύρη τιμών, στρογγυλοποιήσεις ή άλλες επιχειρησιακές συνθήκες. Αντίστοιχα, δεδομένα που συμμορφώνονται με τη θεωρητική κατανομή δεν μπορούν να θεωρηθούν αυτομάτως αξιόπιστα. Έχει επισημανθεί ακόμη και η δυνατότητα σκόπιμης δημιουργίας αριθμητικών τιμών που μιμούνται τη δομή Benford, γεγονός που περιορίζει τη χρήση απλών ελέγχων πρώτου ψηφίου ως ανεξάρτητων μηχανισμών ανίχνευσης [48].

Επομένως, στην παρούσα εργασία ο Benford λειτουργεί ως συμπληρωματικός μηχανισμός στατιστικής διερεύνησης. Η αξία του βρίσκεται κυρίως στην παροχή δεύτερης οπτικής πάνω στη δομή των χρηματικών δεδομένων και όχι στην επιβεβαίωση των προβλέψεων του συνδυαστικού μοντέλου.

6.5 Συζήτηση Αποτελεσμάτων

Η εφαρμογή του τελικού πλαισίου στα πραγματικά δεδομένα ERP αποτέλεσε διαφορετικό στάδιο από την προηγούμενη αξιολόγηση στο σύνολο δεδομένων IEEE-CIS Fraud Detection. Στο συγκεκριμένο σύνολο δεν υπήρχαν διαθέσιμες πραγματικές ετικέτες απάτης και, κατά συνέπεια, ο στόχος δεν ήταν η επανάληψη της συγκριτικής αξιολόγησης των μοντέλων, αλλά η διερεύνηση της τεχνικής δυνατότητας παραγωγής και ιεράρχησης βαθμολογιών κινδύνου για τις λογιστικές εγγραφές.

Ο τελικός σταθμισμένος συνδυασμός βαθμολογιών παρήγαγε μία συνεχή βαθμολογία κινδύνου για το σύνολο των 65.479 εγγραφών. Με βάση το προκαθορισμένο σημείο λειτουργίας, 6.548 εγγραφές κατατάχθηκαν στο ανώτερο 10% και χαρακτηρίστηκαν ως περιπτώσεις αυξημένης προτεραιότητας για περαιτέρω διερεύνηση. Επιπλέον, το ανώτερο 1% περιλάμβανε 655 εγγραφές, παρέχοντας ένα ακόμη περισσότερο περιορισμένο υποσύνολο για ενδεχόμενο στοχευμένο έλεγχο.

Η εξέταση των επισημασμένων εγγραφών ανέδειξε σημαντική παρουσία συμψηφιστικών κινήσεων, λογαριασμών πελατών, ταμείου, προμηθευτών και τραπεζικών καταθέσεων. Παράλληλα, στις εγγραφές υψηλότερης βαθμολογίας κινδύνου εμφανίστηκαν συναλλαγές με σημαντικές διαφοροποιήσεις ως προς το χρηματικό τους ύψος, από ποσά μερικών χιλιάδων έως ιδιαίτερα υψηλές χρηματικές αξίες. Η εικόνα αυτή υποδηλώνει ότι η τελική κατάταξη δεν καθορίζεται αποκλειστικά από το ύψος του ποσού, αλλά προκύπτει από τη συνδυαστική αξιοποίηση πολλαπλών χαρακτηριστικών από τα επιμέρους μοντέλα.

Ένα σημαντικό εύρημα αφορά την εσωτερική σύνθεση της τελικής βαθμολογίας κινδύνου. Η ανάλυση της σταθμισμένης συνεισφοράς στις επισημασμένες εγγραφές έδειξε ότι το Isolation Forest είχε τη μεγαλύτερη μέση σχετική συνεισφορά, ίση με 86,85%, ενώ το HBOS ακολούθησε με 12,95%. Οι αντίστοιχες μέσες συνεισφορές του Autoencoder και του LOF ήταν σημαντικά μικρότερες, περίπου 0,19% και 0,01% αντίστοιχα. Επιπλέον, το Isolation Forest αποτέλεσε το μοντέλο με τη μεγαλύτερη επιμέρους σταθμισμένη συνεισφορά και στις 6.548 επισημασμένες εγγραφές.

Η έντονη αυτή συγκέντρωση πρέπει να ερμηνευθεί σε συνάρτηση τόσο με την επιλεγμένη διαμόρφωση IF-Dominant, στην οποία το Isolation Forest λαμβάνει το μεγαλύτερο βάρος, όσο και με τις διαφορετικές κατανομές των κανονικοποιημένων βαθμολογιών ανωμαλίας στα πραγματικά δεδομένα ERP. Συνεπώς, το αποτέλεσμα περιγράφει τον τρόπο με τον οποίο διαμορφώθηκε η τελική βαθμολογία κινδύνου και δεν μπορεί να θεωρηθεί από μόνο του ένδειξη μεγαλύτερης πραγματικής ακρίβειας του Isolation Forest, καθώς στο ERP δεν υπάρχουν διαθέσιμες ετικέτες για την επαλήθευση των επισημάνσεων.

Η εφαρμογή του Νόμου του Benford παρείχε μία δεύτερη στατιστική οπτική. Παρότι παρατηρήθηκαν διαφοροποιήσεις στην κατανομή του πρώτου σημαντικού ψηφίου και στατιστικά ανιχνεύσιμη σχέση μεταξύ πρώτου ψηφίου και επισήμανσης, το μέγεθος της σχέσης ήταν ασθενές. Επιπλέον, το επισημασμένο σύνολο παρουσίασε μικρότερη Μέση Απόλυτη Απόκλιση (MAD) από τη θεωρητική κατανομή Benford σε σχέση με το μη επισημασμένο σύνολο. Συνεπώς, τα αποτελέσματα δεν υποστηρίζουν μία απλή σχέση σύμφωνα με την οποία οι εγγραφές υψηλότερης βαθμολογίας κινδύνου παρουσιάζουν ταυτόχρονα εντονότερη απόκλιση από τον Νόμο του Benford.

Η συγκεκριμένη διαφοροποίηση είναι συμβατή με τη συμπληρωματική χρήση των δύο προσεγγίσεων. Το συνδυαστικό μοντέλο εξετάζει πολυδιάστατα χαρακτηριστικά σε επίπεδο μεμονωμένης εγγραφής, ενώ ο Νόμος του Benford εξετάζει συλλογικές ιδιότητες της αριθμητικής κατανομής. Η μη πλήρης συμφωνία των δύο προσεγγίσεων δεν αποτελεί κατ' ανάγκη αρνητικό αποτέλεσμα, αλλά υποδηλώνει ότι εξετάζουν διαφορετικές πτυχές της ασυνήθιστης συμπεριφοράς των οικονομικών δεδομένων.

Οι παραπάνω παρατηρήσεις πρέπει να αξιολογηθούν υπό το πρίσμα των περιορισμών της εφαρμογής. Πρώτον, η απουσία πραγματικών ετικετών δεν επιτρέπει την επαλήθευση του κατά πόσο οι εγγραφές υψηλής βαθμολογίας κινδύνου αντιστοιχούν σε πραγματικές περιπτώσεις απάτης, λογιστικές ανακρίβειες ή νόμιμες αλλά ασυνήθιστες συναλλαγές. Δεύτερον, το ποσοστό 10% αποτελεί λειτουργικό σημείο επιλογής και όχι εκτίμηση της πραγματικής συχνότητας απάτης. Τρίτον, η κατανομή των βαθμολογιών των επιμέρους μοντέλων επηρεάζει την πραγματική συμμετοχή τους στον σταθμισμένο συνδυασμό. Τέλος, η απόκλιση ή η συμμόρφωση με τον Νόμο του Benford δεν αποτελεί από μόνη της επαρκή ένδειξη δόλιας συμπεριφοράς [44], [48].

Παρά τους περιορισμούς αυτούς, η εφαρμογή υποστηρίζει την τεχνική δυνατότητα χρήσης του πλαισίου ως μηχανισμού ιεράρχησης λογιστικών εγγραφών προς περαιτέρω έλεγχο. Αντί να αποδίδει αυτόματα τον χαρακτηρισμό της απάτης, το σύστημα περιορίζει τον όγκο των εγγραφών που απαιτούν άμεση προσοχή και παρέχει συνεχή βαθμολογία κινδύνου. Η ανάλυση Benford προσφέρει ένα συμπληρωματικό επίπεδο στατιστικής διερεύνησης, χωρίς να λειτουργεί ως μηχανισμός επιβεβαίωσης των αποτελεσμάτων του συνδυαστικού μοντέλου.

Υπό αυτή την έννοια, η εφαρμογή στο ERP και η ανάλυση Benford αποτελούν το τελικό στάδιο εμπειρικής διερεύνησης της εργασίας. Τα αποτελέσματα υποστηρίζουν την τεχνική δυνατότητα μεταφοράς του πλαισίου από ένα ελεγχόμενο σύνολο δεδομένων αναφοράς σε πραγματικά μη επισημασμένα λογιστικά δεδομένα με σκοπό την παραγωγή βαθμολογιών κινδύνου και την υποστήριξη της διαδικασίας ελέγχου. Δεν επιτρέπουν, ωστόσο, την εξαγωγή συμπερασμάτων σχετικά με την πραγματική ικανότητα ανίχνευσης απάτης στο συγκεκριμένο ERP. Για τον λόγο αυτό, η ανθρώπινη αξιολόγηση των υψηλότερα καταταγμένων εγγραφών παραμένει απαραίτητη πριν από οποιονδήποτε οριστικό χαρακτηρισμό τους ως λογιστικής παρατυπίας ή οικονομικής απάτης.

Κεφάλαιο 7ο: Συζήτηση και Συμπεράσματα

Το παρόν κεφάλαιο συνοψίζει και συνθέτει τα βασικά ευρήματα της εργασίας, συνδέοντας την αξιολόγηση των επιμέρους μοντέλων ανίχνευσης ανωμαλιών, τη διερεύνηση της μεταξύ τους συμπληρωματικότητας, την εφαρμογή διαφορετικών στρατηγικών συνδυασμού και τη μεταφορά του τελικού πλαισίου σε πραγματικά δεδομένα ERP. Παράλληλα, εξετάζεται ο ρόλος του Νόμου του Benford ως συμπληρωματικού στατιστικού μηχανισμού και διατυπώνονται οι βασικοί περιορισμοί και οι κατευθύνσεις για περαιτέρω έρευνα.

7.1 Συνολική Αποτίμηση

Η παρούσα εργασία διερεύνησε την ανίχνευση οικονομικής απάτης και λογιστικών ανωμαλιών μέσω ετερογενών μοντέλων ανίχνευσης ανωμαλιών, με βασικό στόχο όχι μόνο τη σύγκριση της ατομικής τους απόδοσης αλλά και την εξέταση του κατά πόσο οι διαφορετικοί μηχανισμοί ανίχνευσης μπορούν να αξιοποιηθούν συνδυαστικά. Η πειραματική διαδικασία πραγματοποιήθηκε αρχικά στο σύνολο δεδομένων IEEE-CIS Fraud Detection, όπου οι διαθέσιμες ετικέτες επέτρεψαν ποσοτική αξιολόγηση, και στη συνέχεια η επιλεγμένη μεθοδολογία εφαρμόστηκε σε πραγματικές λογιστικές εγγραφές ERP χωρίς διαθέσιμες γνωστές πραγματικές ετικέτες. Η συγκεκριμένη διάκριση μεταξύ του συνόλου δεδομένων αναφοράς και πραγματικού επιχειρησιακού περιβάλλοντος αποτέλεσε βασικό στοιχείο του σχεδιασμού της εργασίας.

Η αξιολόγηση των Isolation Forest, LOF, HBOS και Autoencoder ανέδειξε ότι δεν υπήρξε ένα μοντέλο που να υπερέχει ταυτόχρονα σε όλες τις μετρικές. Στο κοινό δείγμα των 100.000 συναλλαγών, το Isolation Forest παρουσίασε την ισχυρότερη συνολική συμπεριφορά στο επιλεγμένο λειτουργικό σημείο, καταγράφοντας τις υψηλότερες τιμές Precision, Recall και F1-score, καθώς και την υψηλότερη ROC-AUC. Αντίθετα, ο Autoencoder εμφάνισε την υψηλότερη AUPRC/AP, γεγονός που υποδηλώνει καλύτερη συνολική ικανότητα κατάταξης των πραγματικών περιπτώσεων απάτης σε ένα έντονα ανισόρροπο περιβάλλον. Το αποτέλεσμα αυτό αναδεικνύει τη σημασία της αξιολόγησης με περισσότερες από μία μετρικές και υποστηρίζει ότι η έννοια του «καλύτερου μοντέλου» εξαρτάται από το λειτουργικό κριτήριο που εξετάζεται.

Παράλληλα, η ανάλυση συμπληρωματικότητας έδειξε ότι τα τέσσερα μοντέλα δεν αποτελούν ισοδύναμους μηχανισμούς ανίχνευσης. Οι διαφορές στις κοινές και αποκλειστικές πραγματικές θετικές (TP) ανιχνεύσεις, στις τιμές Jaccard, στις συσχετίσεις των βαθμολογιών ανωμαλίας και στα αποτελέσματα του ελέγχου McNemar υποδηλώνουν διαφορετικές συμπεριφορές μεταξύ των μοντέλων. Για παράδειγμα, το Isolation Forest και το HBOS παρουσίασαν σχετικά υψηλή επικάλυψη στις πραγματικές περιπτώσεις απάτης, ενώ άλλα ζεύγη, όπως LOF-HBOS, εμφάνισαν αισθητά χαμηλότερη επικάλυψη. Επιπλέον, ο έλεγχος McNemar μετά από διόρθωση Holm ανέδειξε στατιστικά σημαντική ασυμμετρία στις αποκλειστικές ορθές ανιχνεύσεις για πέντε από τα έξι ζεύγη. Επομένως, η διαφορετικότητα μεταξύ των μοντέλων δεν περιορίστηκε στη θεωρητικά διαφορετική αρχή λειτουργίας τους αλλά αποτυπώθηκε και στην πειραματική τους συμπεριφορά.

Η ύπαρξη συμπληρωματικότητας, ωστόσο, δεν σήμαινε ότι κάθε συνδυασμός των μοντέλων οδηγούσε αυτομάτως σε καλύτερη συνολική απόδοση. Τα αποτελέσματα των τεχνικών συνδυασμού ανέδειξαν κυρίως διαφορετικά λειτουργικά προφίλ. Το OR/Union αύξησε σημαντικά το Recall, επιτρέποντας τον εντοπισμό μεγαλύτερου αριθμού πραγματικών περιπτώσεων απάτης, αλλά ταυτόχρονα αύξησε σημαντικά τις ψευδώς θετικές επισημάνσεις (FP) και περιορίσε το Precision. Αντίθετα, ο αυστηρότερος κανόνας πλειοψηφικής βαθμολογίας με τουλάχιστον τρεις θετικές ψήφους παρουσίασε το υψηλότερο Precision και F1-score μεταξύ των τεχνικών συνδυασμού, με αντίστοιχη όμως μείωση του Recall. Οι διαμορφώσεις σταθμισμένης βαθμολογίας κινήθηκαν σε ενδιάμεσο επίπεδο, προσφέροντας παράλληλα το πλεονέκτημα μιας συνεχούς βαθμολογίας κινδύνου.

Ένα από τα σημαντικότερα ευρήματα της εργασίας είναι επομένως ότι η χρήση τεχνικής συνδυασμού δεν μπορεί να θεωρηθεί γενικά ανώτερη από τη χρήση του καλύτερου μεμονωμένου μοντέλου. Κανένα από τα εξεταζόμενα υβριδικά μοντέλα δεν ξεπέρασε την AUPRC/AP του Autoencoder, ενώ το Isolation Forest διατήρησε υψηλότερη ROC-AUC από τις διαμορφώσεις βαρών. Αντίθετα, συγκεκριμένες στρατηγικές συνδυασμού βελτίωσαν συγκεκριμένες διαστάσεις της λειτουργικής συμπεριφοράς, όπως

το Recall ή το Precision. Το αποτέλεσμα αυτό υποστηρίζει ότι η αξία ενός υβριδικού μοντέλου σε προβλήματα ανίχνευσης απάτης πρέπει να εξετάζεται με βάση τον επιδιωκόμενο συμβιβασμό μεταξύ της κάλυψης πραγματικών περιστατικών και του αριθμού των ψευδών θετικών(FP), και όχι μέσω της γενικής υπόθεσης ότι ο συνδυασμός πολλών μοντέλων συνεπάγεται υποχρεωτικά καλύτερη απόδοση. Στο πλαίσιο των συνδυασμών σταθμισμένης ψηφοφορίας, η διαμόρφωση IF Dominant επιλέχθηκε ως τελική στρατηγική μετά από συνδυαστική συνεκτίμηση των μετρικών. Η Diversity Focused διαμόρφωση παρουσίασε οριακά υψηλότερη AUPRC/AP, όμως η IF Dominant εμφάνισε υψηλότερα Precision, Recall και F1-score και την υψηλότερη ROC-AUC μεταξύ των σταθμισμένων διαμορφώσεων. Η επιλογή αυτή είναι συνεπής με τη μεθοδολογική λογική της εργασίας, όπου η AUPRC αποκτά ιδιαίτερη σημασία λόγω της ανισορροπίας, χωρίς όμως η τελική επιλογή να εξαρτάται αποκλειστικά από μία μεμονωμένη μετρική.

Η συμπληρωματική ανάλυση μέσω Θεωρίας Παιγνίων προσέφερε διαφορετική οπτική στη συνεισφορά των τεσσάρων μοντέλων. Οι τιμές Shapley ανέδειξαν το Isolation Forest ως το μοντέλο με τη μεγαλύτερη μέση οριακή συνεισφορά, ακολουθούμενο σε πολύ μικρή απόσταση από τον Autoencoder, ενώ το HBOS διατήρησε θετική συνεισφορά. Αντίθετα, το LOF παρουσίασε οριακά αρνητική τιμή Shapley υπό τη συγκεκριμένη συνάρτηση αξίας και τον ισότιμο συνδυασμό των βαθμολογιών στους συνασπισμούς. Το εύρημα αυτό δεν υποδηλώνει ότι το LOF δεν εντοπίζει χρήσιμες ή διαφορετικές περιπτώσεις. Αναδεικνύει όμως ότι η διαφορετικότητα ενός ανιχνευτή δεν συνεπάγεται απαραίτητα θετική μέση επίδραση στην ποιότητα της τελικής κατάταξης όταν αυτός ενσωματώνεται σε έναν συνδυασμό.

Η εφαρμογή του τελικού πλαισίου στα δεδομένα ERP διαφοροποιήθηκε ουσιαστικά από την αξιολόγηση του συνόλου δεδομένων αναφοράς. Εφόσον δεν υπήρχαν επιβεβαιωμένες ετικέτες απάτης, δεν ήταν δυνατό να υπολογιστούν Precision, Recall ή AUPRC και, επομένως, το σύστημα χρησιμοποιήθηκε ως μηχανισμός παραγωγής και ιεράρχησης βαθμολογίας κινδύνου. Από τις 65.479 λογιστικές εγγραφές, το ανώτερο 10% της κατανομής χαρακτηρίστηκε ως σύνολο αυξημένης προτεραιότητας για περαιτέρω διερεύνηση. Η προσέγγιση αυτή είναι συνεπής με τη βασική θέση της εργασίας ότι ένα σύστημα ανίχνευσης ανωμαλιών δεν πρέπει να αντιμετωπίζεται ως μηχανισμός αυτόματης απόδειξης απάτης, αλλά ως εργαλείο περιορισμού και προτεραιοποίησης του ελεγκτικού φόρτου.

Η εφαρμογή στο ERP ανέδειξε επίσης ότι τα προκαθορισμένα βάρη δεν μεταφράζονται απαραίτητα σε αντίστοιχα ποσοστά πραγματικής αριθμητικής συμμετοχής στη τελική βαθμολογία κινδύνου. Στις επισημασμένες εγγραφές, το Isolation Forest κυριάρχησε στη διαμόρφωση της τελικής βαθμολογίας, με το HBOS να ακολουθεί και τις συνεισφορές του Autoencoder και του LOF να είναι αισθητά χαμηλότερες. Το εύρημα αυτό αναδεικνύει τη σημασία όχι μόνο της επιλογής των βαρών αλλά και της βαθμονόμησης και της κατανομής των βαθμολογιών ανωμαλίας όταν ένα υβριδικό μοντέλο μεταφέρεται σε διαφορετικό σύνολο δεδομένων.

Τέλος, η ανάλυση με τον Νόμο του Benford δεν παρουσίασε απλή αντιστοιχία με τις επισημάνσεις του υβριδικού μοντέλου. Οι επισημασμένες εγγραφές παρουσίασαν μικρότερη MAD από τις μη επισημασμένες, ενώ η στατιστικά ανιχνεύσιμη σχέση μεταξύ πρώτου σημαντικού ψηφίου και επισημάνση συνδυαστικού μοντέλου συνοδεύτηκε από χαμηλή τιμή Cramér's V, υποδηλώνοντας ασθενές μέγεθος σχέσης. Επομένως, τα αποτελέσματα δεν φαίνεται να υποστηρίζουν την υπόθεση ότι οι εγγραφές με υψηλότερη βαθμολογία κινδύνου πρέπει να παρουσιάζουν ταυτόχρονα μεγαλύτερη απόκλιση από τον Benford. Αντίθετα, φαίνεται περισσότερο εύλογο να θεωρηθούν οι δύο προσεγγίσεις ως διαφορετικοί και συμπληρωματικοί μηχανισμοί διερεύνησης, καθώς το υβριδικό μοντέλο αξιολογεί πολυδιάστατη συμπεριφορά σε επίπεδο εγγραφής, ενώ ο Benford εξετάζει συλλογικές ιδιότητες της αριθμητικής κατανομής. Η βιβλιογραφία επίσης επισημαίνει ότι η μη συμμόρφωση με τον Benford δεν αποτελεί από μόνη της απόδειξη χειραγώγησης και ότι τα αποτελέσματα πρέπει να ερμηνεύονται με προσοχή [44], [45], [48].

7.2 Απαντήσεις στα Ερευνητικά Ερωτήματα

RQ1 – Ποιο από τα εξεταζόμενα μοντέλα ανίχνευσης ανωμαλιών παρουσιάζει την ισχυρότερη συνολική συμπεριφορά για την ανίχνευση οικονομικής απάτης στο σύνολο δεδομένων αναφοράς;

Η αξιολόγηση δεν ανέδειξε ένα μοντέλο που να υπερέχει σε όλες τις εξεταζόμενες μετρικές. Το Isolation Forest παρουσίασε την ισχυρότερη συνολική συμπεριφορά στο επιλεγμένο σημείο λειτουργίας,

επιτυγχάνοντας τις υψηλότερες τιμές Precision, Recall, F1-score και ROC-AUC. Ο Autoencoder, ωστόσο, παρουσίασε την υψηλότερη AUPRC/AP. Συνεπώς, όταν η αξιολόγηση επικεντρώνεται στις δυαδικές αποφάσεις του συγκεκριμένου σημείου λειτουργίας, το Isolation Forest παρουσίασε την ισχυρότερη συμπεριφορά μεταξύ των εξεταζόμενων μοντέλων. Αντίθετα, όταν το ενδιαφέρον επικεντρώνεται στην ποιότητα κατάταξης των συναλλαγών σε όλο το εύρος των πιθανών σημείων λειτουργίας, ο Autoencoder παρουσίασε πλεονέκτημα ως προς την AUPRC/AP.

RQ2 – Σε ποιο βαθμό τα διαφορετικά μοντέλα παρουσιάζουν συμπληρωματικότητα ως προς τις περιπτώσεις απάτης που εντοπίζουν και τις βαθμολογίες ανωμαλίας που παράγουν;

Τα αποτελέσματα υποδηλώνουν ουσιαστική αλλά ανομοιομορφή συμπληρωματικότητα μεταξύ των μοντέλων. Κανένα ζεύγος δεν παρουσίασε πλήρη ταύτιση των πραγματικά θετικών ανιχνεύσεων, ενώ οι τιμές του δείκτη ομοιότητας Jaccard διαφοροποιήθηκαν αισθητά μεταξύ των ζευγών. Παράλληλα, η ανάλυση των βαθμολογιών ανωμαλίας παρουσίασε διαφορετικά επίπεδα συσχέτισης, ενώ ο έλεγχος McNemar ανέδειξε στατιστικά σημαντικές ασυμμετρίες για τα περισσότερα ζεύγη. Επομένως, οι διαφορετικές αρχές λειτουργίας των μοντέλων φαίνεται να συνδέονται με διαφορετική πρακτική συμπεριφορά ως προς τις συναλλαγές που αναδεικνύουν ως ύποπτες. Η συμπληρωματικότητα, ωστόσο, δεν πρέπει να ταυτίζεται με ισότιμη χρησιμότητα, καθώς ένα μοντέλο μπορεί να εντοπίζει διαφορετικές περιπτώσεις χωρίς η ενσωμάτωση της βαθμολογίας του να βελτιώνει απαραίτητα τη συνολική απόδοση ενός συνδυαστικού μοντέλου.

RQ3 – Μπορούν οι στρατηγικές συνδυασμού να βελτιώσουν επιμέρους χαρακτηριστικά της συνολικής συμπεριφοράς του συστήματος σε σχέση με τη χρήση μεμονωμένων ανιχνευτών;

Τα αποτελέσματα υποδηλώνουν ότι οι στρατηγικές συνδυασμού μπορούν να βελτιώσουν επιμέρους χαρακτηριστικά της λειτουργίας του συστήματος, χωρίς όμως να υπερέχουν καθολικά των μεμονωμένων μοντέλων. Η τεχνική OR/Union παρουσίασε το υψηλότερο Recall, ενώ η διαμόρφωση Majority Voting ≥ 3 παρουσίασε υψηλότερα Precision και F1-score από οποιοδήποτε μεμονωμένο μοντέλο στο συγκεκριμένο σημείο λειτουργίας. Ωστόσο, καμία συνδυαστική διαμόρφωση δεν βελτίωσε ταυτόχρονα όλες τις εξεταζόμενες μετρικές και καμία δεν ξεπέρασε την AUPRC/AP του Autoencoder. Συνεπώς, τα αποτελέσματα δεν υποστηρίζουν γενική υπεροχή των τεχνικών συνδυασμού, αλλά υποδηλώνουν τη δυνατότητά τους να διαμορφώνουν διαφορετικά προφίλ ανίχνευσης ανάλογα με τον επιχειρησιακό στόχο.

RQ4 – Πώς επηρεάζεται ο συμβιβασμός μεταξύ Recall και Precision από τον τρόπο συνδυασμού των μοντέλων και από το επιλεγμένο σημείο λειτουργίας;

Ο συμβιβασμός μεταξύ Recall και Precision επηρεάστηκε άμεσα από το επίπεδο συναίνεσης που απαιτεί κάθε στρατηγική συνδυασμού. Στην τεχνική OR/Union, όπου αρκεί η θετική απόφαση ενός μόνο ανιχνευτή, αυξήθηκε το Recall, αλλά παράλληλα αυξήθηκε σημαντικά ο αριθμός των ψευδώς θετικών επισημάνσεων. Η αύξηση του απαιτούμενου αριθμού θετικών ψήφων περιόρισε σταδιακά τον αριθμό των επισημάνσεων και αύξησε το Precision, με αντίστοιχη μείωση του Recall. Οι προσεγγίσεις σταθμισμένου συνδυασμού βαθμολογιών παρουσίασαν ενδιάμεση συμπεριφορά και επέτρεψαν επιπλέον την παραγωγή συνεχούς βαθμολογίας ανωμαλίας. Επομένως, δεν προκύπτει ένα μοναδικό σημείο λειτουργίας κατάλληλο για κάθε εφαρμογή. Η επιλογή του εξαρτάται από το σχετικό κόστος των ψευδώς θετικών και ψευδώς αρνητικών περιπτώσεων, καθώς και από τη διαθέσιμη δυνατότητα ανθρώπινου ελέγχου.

RQ5 – Μπορεί ο Νόμος του Benford να λειτουργήσει συμπληρωματικά προς την ανίχνευση μέσω συνδυασμού μοντέλων ως μηχανισμός στατιστικής διερεύνησης των ύποπτων οικονομικών συναλλαγών;

Τα αποτελέσματα είναι συμβατά με τη χρήση του Νόμου του Benford ως συμπληρωματικού μηχανισμού στατιστικής διερεύνησης και όχι ως μεθόδου επιβεβαίωσης των επισημάνσεων του συνδυαστικού μοντέλου. Η κατανομή του πρώτου σημαντικού ψηφίου παρουσίασε διαφοροποιήσεις μεταξύ επισημασμένων και μη επισημασμένων εγγραφών, όμως το μέγεθος της σχέσης ήταν ασθενές. Επιπλέον, οι επισημασμένες εγγραφές παρουσίασαν μικρότερη MAD από τις μη επισημασμένες. Συνεπώς, οι συναλλαγές που λαμβάνουν υψηλότερη βαθμολογία κινδύνου από το συνδυαστικό μοντέλο δεν είναι απαραίτητο να παρουσιάζουν ταυτόχρονα μεγαλύτερη απόκλιση από την αναμενόμενη κατανομή Benford. Η χρησιμότητα του νόμου στην παρούσα εργασία εντοπίζεται επομένως στην παροχή μιας δεύτερης, διαφορετικής στατιστικής οπτικής στα οικονομικά δεδομένα και όχι στην επικύρωση των αποτελεσμάτων του συνδυαστικού μοντέλου.

RQ6 – Πώς συμπεριφέρεται το προτεινόμενο πλαίσιο όταν εφαρμόζεται σε πραγματικά δεδομένα ERP χωρίς διαθέσιμες ετικέτες απάτης;

Το προτεινόμενο πλαίσιο εφαρμόστηκε στο πραγματικό σύνολο δεδομένων ERP και παρήγαγε συνεχείς βαθμολογίες κινδύνου για το σύνολο των 65.479 λογιστικών εγγραφών. Η χρήση κατωφλίου βάσει εκατοστημορίου επέτρεψε την επιλογή ενός περιορισμένου υποσυνόλου εγγραφών αυξημένης προτεραιότητας και την περαιτέρω διερεύνησή τους με βάση λογιστικά χαρακτηριστικά και τον Νόμο του Benford. Ωστόσο, λόγω της απουσίας επιβεβαιωμένων πραγματικών ετικετών, δεν είναι δυνατό να εξαχθεί συμπέρασμα σχετικά με την πραγματική Precision ή το Recall του συστήματος στα δεδομένα ERP. Η εφαρμογή υποστηρίζει επομένως την τεχνική και λειτουργική δυνατότητα χρήσης του πλαισίου ως μηχανισμού ιεράρχησης ύποπτων εγγραφών και όχι την αποδεδειγμένη ικανότητά του να αναγνωρίζει πραγματικές περιπτώσεις απάτης.

7.3 Περιορισμοί της Εργασίας

Παρά τα ευρήματα που προέκυψαν, η παρούσα εργασία υπόκειται σε ορισμένους περιορισμούς, οι οποίοι πρέπει να λαμβάνονται υπόψη κατά την ερμηνεία των αποτελεσμάτων.

Ένας πρώτος περιορισμός αφορά τη φύση του συνόλου δεδομένων IEEE-CIS Fraud Detection. Παρότι αποτελεί μεγάλο και σύνθετο σύνολο δεδομένων αναφοράς για την ανίχνευση απάτης, δεν αντιπροσωπεύει πλήρως τη δομή ενός λογιστικού Γενικού Καθολικού (General Ledger – GL) ή ενός συστήματος ERP. Τα χαρακτηριστικά, η διαδικασία δημιουργίας των συναλλαγών και οι μορφές πιθανής απάτης διαφοροποιούνται από εκείνες ενός πραγματικού λογιστικού περιβάλλοντος. Για τον λόγο αυτό, η καλή συμπεριφορά ενός μοντέλου στο IEEE-CIS δεν συνεπάγεται αντίστοιχη αποτελεσματικότητα όταν αυτό εφαρμόζεται σε δεδομένα ERP.

Παράλληλα, λόγω των υπολογιστικών απαιτήσεων ορισμένων αλγορίθμων, ιδιαίτερα του LOF και του Autoencoder, μέρος της πειραματικής διαδικασίας πραγματοποιήθηκε σε κοινό δείγμα 100.000 συναλλαγών και όχι στο πλήρες σύνολο IEEE-CIS. Η επιλογή κοινού δείγματος εξασφάλισε άμεση συγκρισιμότητα μεταξύ των μοντέλων, περιορίζει όμως τη δυνατότητα να υποστηριχθεί ότι η ίδια ακριβώς συμπεριφορά θα εμφανιζόταν στο σύνολο των διαθέσιμων συναλλαγών.

Ένας επιπλέον περιορισμός σχετίζεται με τη χρήση της παραμέτρου contamination και των κατωφλίων βάσει εκατοστημορίων. Η τιμή contamination χρησιμοποιήθηκε ως παράμετρος του σημείου λειτουργίας για τον καθορισμό του αριθμού των παρατηρήσεων που επισημαίνονται ως ανωμαλίες και δεν πρέπει να ερμηνεύεται ως πραγματική εκτίμηση της συχνότητας οικονομικής απάτης. Διαφορετική επιλογή του κατωφλίου θα μετέβαλλε άμεσα τις δυαδικές προβλέψεις και, κατά συνέπεια, τις τιμές των Precision, Recall και F1-score.

Παρότι η εκπαίδευση των βασικών ανιχνευτών ανωμαλιών δεν βασίστηκε στη μεταβλητή isFraud ως στόχο επιβλεπόμενης μάθησης, οι πραγματικές ετικέτες του IEEE-CIS χρησιμοποιήθηκαν για την αξιολόγηση των αποτελεσμάτων και την επιλογή των τελικών πειραματικών διαμορφώσεων. Συνεπώς, η διαδικασία επιλογής παραμέτρων πραγματοποιήθηκε σε ελεγχόμενο σύνολο δεδομένων αναφοράς με γνωστές πραγματικές ετικέτες και δεν αντιστοιχεί πλήρως σε μία κατάσταση στην οποία δεν υπάρχει καμία διαθέσιμη πληροφορία για την πραγματική κλάση. Επιπλέον, η εκπαίδευση του τελικού Autoencoder σε σύνολο εκπαίδευσης που περιλαμβάνει μόνο κανονικές παρατηρήσεις προϋποθέτει τη γνώση της πραγματικής κλάσης κατά τον σχηματισμό του συνόλου εκπαίδευσης. Για τον λόγο αυτό, η συγκεκριμένη παραλλαγή είναι ακριβέστερο να αντιμετωπίζεται ως στρατηγική μίας κλάσης (one-class) και όχι ως αυστηρά πλήρως μη επιβλεπόμενη διαδικασία.

Περιορισμούς παρουσιάζει και ο σταθμισμένος συνδυασμός βαθμολογιών. Παρότι οι βαθμολογίες ανωμαλίας μεταφέρθηκαν σε κοινή κλίμακα μέσω κανονικοποίησης Min–Max, η διαδικασία αυτή δεν εξαλείφει τις διαφορές ως προς τη μορφή και την κατανομή των βαθμολογιών των επιμέρους μοντέλων. Το γεγονός αυτό έγινε ιδιαίτερα εμφανές κατά την εφαρμογή στα δεδομένα ERP, όπου η πραγματική αριθμητική συμμετοχή των μοντέλων στην τελική βαθμολογία κινδύνου διαφοροποιήθηκε σημαντικά από τα προκαθορισμένα βάρη. Συνεπώς, ένα σύνολο βαρών που παρουσιάζει ικανοποιητική συμπεριφορά σε ένα σύνολο δεδομένων αναφοράς δεν είναι βέβαιο ότι θα διατηρεί την ίδια λειτουργική ισορροπία όταν εφαρμόζεται σε διαφορετική κατανομή δεδομένων.

Ο σημαντικότερος περιορισμός της εφαρμογής στα δεδομένα ERP είναι η απουσία επιβεβαιωμένων πραγματικών ετικετών. Δεν είναι γνωστό εάν οι εγγραφές που έλαβαν τις υψηλότερες βαθμολογίες κινδύνου αντιστοιχούν σε πραγματικές περιπτώσεις απάτης, λογιστικά σφάλματα, ασυνήθιστες αλλά

νόμιμες κινήσεις ή απλώς σε παρατηρήσεις που αποκλίνουν από τη συνηθισμένη δομή των δεδομένων. Επομένως, η ανάλυση στα δεδομένα ERP επιτρέπει την αξιολόγηση της τεχνικής δυνατότητας ιεράρχησης και της εσωτερικής συμπεριφοράς του συστήματος, αλλά όχι της πραγματικής ικανότητάς του να αναγνωρίζει περιπτώσεις απάτης.

Επιπλέον, η εφαρμογή στα δεδομένα ERP πραγματοποιήθηκε σε ένα συγκεκριμένο σύνολο δεδομένων και σε περιορισμένη χρονική περίοδο. Διαφορετικές επιχειρήσεις μπορούν να χρησιμοποιούν διαφορετικά λογιστικά σχέδια, τύπους εγγραφών, συστήματα ERP και επιχειρησιακές διαδικασίες. Συνεπώς, τα ευρήματα δεν μπορούν να γενικευθούν αυτομάτως σε κάθε περιβάλλον ERP χωρίς πρόσθετη εξωτερική αξιολόγηση.

Τέλος, περιορισμοί ισχύουν και για την εφαρμογή του Νόμου του Benford. Η αναμενόμενη κατανομή του πρώτου σημαντικού ψηφίου δεν αποτελεί καθολικό χαρακτηριστικό κάθε οικονομικού συνόλου δεδομένων, ενώ το μέγεθος του δείγματος και η διαδικασία δημιουργίας των αριθμητικών τιμών μπορούν να επηρεάσουν τα αποτελέσματα των στατιστικών ελέγχων. Επιπλέον, η συμμόρφωση ή η απόκλιση από τον Νόμο του Benford δεν μπορεί από μόνη της να χρησιμοποιηθεί για τον χαρακτηρισμό μιας συναλλαγής ως δόλιας [44], [45], [48]. Για τον λόγο αυτό, στην παρούσα εργασία ο Νόμος του Benford χρησιμοποιήθηκε ως συμπληρωματικό στατιστικό εργαλείο διερεύνησης.

7.4 Μελλοντική Έρευνα

Τα αποτελέσματα της εργασίας δημιουργούν αρκετές δυνατότητες περαιτέρω ανάπτυξης του προτεινόμενου πλαισίου. Μία από τις σημαντικότερες αφορά την ενίσχυση της ερμηνευσιμότητας. Η παρούσα ανάλυση μέσω τιμών Shapley εξετάζει τη μέση οριακή συνεισφορά των ίδιων των ανιχνευτών στο συνδυαστικό μοντέλο. Μία επόμενη κατεύθυνση θα μπορούσε να επεκτείνει την ερμηνεία σε επίπεδο μεμονωμένης συναλλαγής, ώστε για κάθε υψηλή βαθμολογία κινδύνου να προσδιορίζονται τα συγκεκριμένα χαρακτηριστικά που συνέβαλαν περισσότερο στην επισήμανση. Με τον τρόπο αυτό, το σύστημα δεν θα παρέχει μόνο μία αριθμητική βαθμολογία κινδύνου, αλλά και πληροφορία που θα μπορούσε να διευκολύνει τον ελεγκτή στην κατανόηση των παραγόντων που οδήγησαν στην επισήμανση.

Ιδιαίτερο ενδιαφέρον παρουσιάζει επίσης η αντικατάσταση των στατικών βαρών από δυναμικούς μηχανισμούς στάθμισης ή μηχανισμούς που βασίζονται στα δεδομένα. Τα αποτελέσματα στο ERP υποδηλώνουν ότι η πραγματική αριθμητική επίδραση ενός μοντέλου μπορεί να διαφέρει σημαντικά από το ονομαστικό του βάρος. Μελλοντική έρευνα θα μπορούσε επομένως να εξετάσει τη βαθμονόμηση των βαθμολογιών ανωμαλίας πριν από τον συνδυασμό τους, την αυτόματη προσαρμογή των βαρών ανάλογα με τη σταθερότητα ή την αξιοπιστία κάθε ανιχνευτή ή ακόμη και τη διαφοροποίηση των βαρών ανά κατηγορία συναλλαγής, λογαριασμό ή χρονική περίοδο. Εάν είναι διαθέσιμο ένα μικρό σύνολο συναλλαγών που έχει αξιολογηθεί από ελεγκτές, η πληροφορία αυτή θα μπορούσε να αξιοποιείται για την προοδευτική προσαρμογή του συνδυασμού.

Μία ακόμη επέκταση αφορά την αξιοποίηση προσεγγίσεων βασισμένων σε γράφους. Στην παρούσα εργασία κάθε εγγραφή αναλύθηκε κυρίως ως μεμονωμένη παρατήρηση με ένα σύνολο χαρακτηριστικών. Ωστόσο, οι οικονομικές συναλλαγές συνδέονται μεταξύ τους μέσω λογαριασμών, παραστατικών, χρηστών, προμηθευτών και άλλων επιχειρησιακών οντοτήτων. Η αναπαράσταση αυτών των σχέσεων με τη μορφή γράφου θα μπορούσε να επιτρέψει την αναγνώριση ύποπτων δομών ή αλληλουχιών που δεν είναι εύκολο να εντοπιστούν όταν κάθε εγγραφή εξετάζεται ανεξάρτητα. Η ανίχνευση απάτης με προσεγγίσεις βασισμένες σε γράφους έχει ήδη αναδειχθεί ως σχετική ερευνητική κατεύθυνση για την αξιοποίηση σύνθετων σχέσεων μεταξύ οικονομικών οντοτήτων [17].

Η αξιολόγηση σε μεγαλύτερο αριθμό συνόλων δεδομένων ERP αποτελεί επίσης σημαντική προϋπόθεση για την εξαγωγή ισχυρότερων συμπερασμάτων. Η εφαρμογή της ίδιας μεθοδολογίας σε διαφορετικούς οργανισμούς, οικονομικούς κλάδους, λογιστικά σχέδια και μεγαλύτερες χρονικές περιόδους θα επέτρεπε να εξεταστούν η σταθερότητα των μοντέλων, η δυνατότητα μεταφοράς των βαρών και η επίδραση διαφορετικών επιχειρησιακών κατανομών στις βαθμολογίες ανωμαλίας. Ιδιαίτερη αξία θα είχε η δημιουργία συνόλου δεδομένων στο οποίο μέρος των επισημασμένων εγγραφών θα είχε αξιολογηθεί από έμπειρους ελεγκτές, παρέχοντας έτσι ένα περιορισμένο αλλά πραγματικό σύνολο επιβεβαιωμένων ετικετών αναφοράς.

Σε αυτό το πλαίσιο, ημι-επιβλεπόμενες ή υβριδικές προσεγγίσεις αποτελούν φυσική συνέχεια της εργασίας. Ένα πραγματικό σύστημα θα μπορούσε αρχικά να λειτουργεί μη επιβλεπόμενα και να

προτείνει εγγραφές προς διερεύνηση. Οι αποφάσεις των ελεγκτών σχετικά με τις συγκεκριμένες εγγραφές θα μπορούσαν στη συνέχεια να λειτουργούν ως περιορισμένες αξιόπιστες ετικέτες, οι οποίες θα επανατροφοδοτούν το σύστημα. Με αυτόν τον τρόπο, η μεγάλη ποσότητα μη επισημασμένων δεδομένων ERP θα εξακολουθούσε να αξιοποιείται, αλλά το πλαίσιο θα μπορούσε σταδιακά να προσαρμόζεται σε πραγματικά επιβεβαιωμένα μοτίβα του συγκεκριμένου οργανισμού [18].

Η χρονική διάσταση αποτελεί επίσης σημαντικό πεδίο περαιτέρω διερεύνησης. Η παρούσα εργασία εφαρμόζει τα μοντέλα σε στατικό σύνολο συναλλαγών, ενώ σε πραγματικές συνθήκες οι λογιστικές εγγραφές παράγονται συνεχώς και η φυσιολογική επιχειρησιακή συμπεριφορά μπορεί να μεταβάλλεται με την πάροδο του χρόνου. Ένα σύστημα συνεχούς ανίχνευσης ανωμαλιών σε ροές δεδομένων θα μπορούσε να ενημερώνεται περιοδικά ή σταδιακά, ώστε να προσαρμόζεται στις μεταβολές της φυσιολογικής συμπεριφοράς και να λαμβάνει υπόψη πιθανές μεταβολές των προτύπων των δεδομένων (concept drift). Με τον τρόπο αυτό, τα κατώφλια ανίχνευσης και οι κατανομές των βαθμολογιών ανωμαλίας δεν θα θεωρούνταν σταθερά επ' αόριστον, αλλά θα μπορούσαν να επανεκτιμώνται καθώς μεταβάλλεται η δομή των δεδομένων.

Τέλος, μία περισσότερο διευρυμένη κατεύθυνση μελλοντικής έρευνας αφορά την ανίχνευση απάτης με ομοσπονδιακή μάθηση (federated learning). Η προσέγγιση αυτή συνδέεται άμεσα με το πρόβλημα της περιορισμένης πρόσβασης σε πραγματικά οικονομικά δεδομένα. Διαφορετικές επιχειρήσεις, οργανισμοί ή ακόμη και επιμέρους μονάδες ενός μεγάλου οργανισμού διαθέτουν συναλλαγές που ενδέχεται να περιέχουν χρήσιμα πρότυπα, όμως η συγκέντρωσή τους σε ένα κοινό σύνολο δεδομένων μπορεί να είναι δύσκολη λόγω απαιτήσεων εμπιστευτικότητας και προστασίας δεδομένων. Μία μελλοντική επέκταση θα μπορούσε να διατηρεί την εκπαίδευση των μοντέλων τοπικά σε κάθε οργανισμό και να συνδυάζει μόνο κατάλληλες ενημερώσεις ή παραμέτρους των μοντέλων, χωρίς να απαιτείται η άμεση μεταφορά των πρωτογενών οικονομικών συναλλαγών.

Στο πλαίσιο της παρούσας εργασίας, μια τέτοια προσέγγιση θα μπορούσε να διερευνήσει κατά πόσο τα Isolation Forest, HBOS, LOF και Autoencoder μπορούν να εφαρμοστούν σε κατανεμημένα περιβάλλοντα ERP και κατά πόσο η αξιοποίηση πληροφορίας από πολλαπλές πηγές μπορεί να συμβάλει στη βελτίωση της ανίχνευσης σε σχέση με την αποκλειστικά τοπική εκπαίδευση. Ιδιαίτερη ερευνητική πρόκληση θα αποτελούσε η διαφορετική κατανομή των δεδομένων μεταξύ οργανισμών, καθώς διαφορετικές επιχειρήσεις ενδέχεται να παρουσιάζουν σημαντικά διαφορετικά λογιστικά πρότυπα. Επομένως, ένα πλαίσιο ομοσπονδιακής μάθησης δεν θα αποτελούσε απλώς μια διαφορετική τεχνική υλοποίηση, αλλά μια επέκταση του βασικού προβλήματος που εξετάστηκε στην παρούσα εργασία: πώς μπορούν να συνδυαστούν ετερογενείς πηγές πληροφορίας, διατηρώντας παράλληλα τα ιδιαίτερα χαρακτηριστικά κάθε επιμέρους περιβάλλοντος.

7.5 Συμπεράσματα

Η παρούσα διπλωματική εργασία ανέπτυξε και αξιολόγησε ένα πλαίσιο ανίχνευσης οικονομικής απάτης και λογιστικών ανωμαλιών, το οποίο βασίζεται σε ετερογενή μοντέλα ανίχνευσης ανωμαλιών και διαφορετικές στρατηγικές συνδυασμού. Η αξιολόγηση στο σύνολο δεδομένων IEEE-CIS Fraud Detection ανέδειξε διαφοροποιήσεις μεταξύ των Isolation Forest, LOF, HBOS και Autoencoder, τόσο ως προς τις συνολικές μετρικές απόδοσης όσο και ως προς τις συγκεκριμένες περιπτώσεις που εντοπίζει κάθε μοντέλο.

Τα αποτελέσματα υποδηλώνουν ουσιαστική αλλά ανομοιομορφή συμπληρωματικότητα μεταξύ των μοντέλων, χωρίς όμως ο συνδυασμός τους να οδηγεί αυτομάτως σε καθολική βελτίωση της απόδοσης. Οι διαφορετικές στρατηγικές συνδυασμού μεταβάλλουν το λειτουργικό προφίλ του συστήματος, διαμορφώνοντας διαφορετικούς συμβιβασμούς μεταξύ Recall και Precision. Η τελική επιλογή της διαμόρφωσης IF-Dominant του σταθμισμένου συνδυασμού βαθμολογιών βασίστηκε στη συνολική συνεκτίμηση των εξεταζόμενων μετρικών, του λειτουργικού προφίλ της και της δυνατότητας παραγωγής συνεχούς βαθμολογίας ανωμαλίας για την εφαρμογή σε δεδομένα ERP χωρίς διαθέσιμες ετικέτες απάτης. Η επιλογή αυτή αφορά το συγκεκριμένο πειραματικό πλαίσιο και δεν υποδηλώνει καθολική υπεροχή της διαμόρφωσης έναντι όλων των εναλλακτικών προσεγγίσεων.

Η εφαρμογή στο πραγματικό σύνολο δεδομένων ERP κατέδειξε την τεχνική δυνατότητα του πλαισίου να παράγει και να ιεραρχεί βαθμολογίες κινδύνου ακόμη και όταν δεν είναι διαθέσιμες επιβεβαιωμένες ετικέτες απάτης. Η δυνατότητα αυτή παρουσιάζει πρακτικό ενδιαφέρον για περιβάλλοντα εσωτερικού ελέγχου, όπου η αναλυτική εξέταση του συνόλου των λογιστικών εγγραφών μπορεί να είναι ιδιαίτερα

απαιτητική. Το προτεινόμενο πλαίσιο μπορεί επομένως να λειτουργήσει ως μηχανισμός προτεραιοποίησης, συμβάλλοντας στον περιορισμό του αριθμού των εγγραφών που απαιτούν άμεση ανθρώπινη διερεύνηση. Η εφαρμογή αυτή, ωστόσο, δεν επιτρέπει την εξαγωγή συμπερασμάτων σχετικά με την πραγματική ικανότητα του συστήματος να αναγνωρίζει απάτη στο συγκεκριμένο ERP, λόγω της απουσίας επιβεβαιωμένων πραγματικών ετικετών.

Η συμπληρωματική εφαρμογή του Νόμου του Benford υποδήλωσε ότι η απόκλιση από την αναμενόμενη ψηφιακή κατανομή και η πολυδιάστατη ανωμαλία που εντοπίζεται από τα μοντέλα δεν αποτελούν ταυτόσημες έννοιες. Για τον λόγο αυτό, η χρησιμότητα του Benford εντοπίζεται κυρίως στην παροχή ενός πρόσθετου επιπέδου στατιστικής διερεύνησης και όχι στην επιβεβαίωση των αποτελεσμάτων του συνδυαστικού μοντέλου.

Συνολικά, τα ευρήματα της εργασίας φαίνεται να υποστηρίζουν ότι τα συστήματα συνδυασμού μοντέλων ανίχνευσης ανωμαλιών μπορούν να αποτελέσουν χρήσιμα εργαλεία υποστήριξης του οικονομικού και λογιστικού ελέγχου, ιδιαίτερα σε περιβάλλοντα όπου οι αξιόπιστες ετικέτες είναι περιορισμένες ή απουσιάζουν. Η χρησιμότητά τους εντοπίζεται κυρίως στον εντοπισμό, την κατάταξη και την προτεραιοποίηση ύποπτων περιπτώσεων και όχι στην αυτόματη απόδοση του χαρακτηρισμού της απάτης. Η τελική αξιολόγηση μιας επισημασμένης συναλλαγής εξακολουθεί να απαιτεί ανθρώπινη διερεύνηση, γνώση του επιχειρησιακού πλαισίου και ελεγκτική κρίση.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] M. N. Ashtiani and B. Raahemi, “Intelligent fraud detection in financial statements using machine learning and data mining: A systematic literature review,” *IEEE Access*, vol. 10, pp. 72504–72525, 2022.
- [2] A. Bakumenko, K. Hlaváčková-Schindler, C. Plant, and N. C. Hubig, “Advancing anomaly detection: Non-semantic financial data encoding with large language models,” *IEEE Access*, vol. 13, 2025.
- [3] S. Nielsen, “Management accounting and the concepts of exploratory data analysis and unsupervised machine learning: A literature study and future directions,” *Journal of Accounting & Organizational Change*, vol. 18, no. 5, pp. 811–853, 2022.
- [4] L. Soni and D. Mangala, “Preventing and detecting internal fraud risk: A critical analysis of the anti-fraud tools used by the Indian banking industry,” *Journal of Money Laundering Control*, vol. 28, no. 2, pp. 424–441, 2025.
- [5] P. Craja, A. Kim, and S. Lessmann, “Deep learning for detecting financial statement fraud,” *Decision Support Systems*, vol. 139, Art. no. 113421, 2020.
- [6] W. Hilal, S. A. Gadsden, and J. Yawney, “Financial fraud: A review of anomaly detection techniques and recent advances,” *Expert Systems with Applications*, vol. 193, Art. no. 116429, 2022.
- [7] Y. Chen, C. Zhao, Y. Xu, C. Nie, and Y. Zhang, “Deep learning in financial fraud detection: Innovations, challenges, and applications,” *Data Science and Management*, vol. 9, Art. no. 100162, 2026.
- [8] J. Bertomeu, E. Cheynel, E. Floyd, and W. Pan, “Using machine learning to detect misstatements,” *Review of Accounting Studies*, vol. 26, pp. 468–519, 2021.
- [9] J. Jin and Y. Zhang, “The analysis of fraud detection in financial market under machine learning,” *Scientific Reports*, vol. 15, Art. no. 29959, 2025.
- [10] L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, “Financial fraud detection through the application of machine learning techniques: A literature review,” *Humanities and Social Sciences Communications*, vol. 11, Art. no. 1130, 2024.
- [11] G. Alessi and M. Fugini, “Adaptive real-time financial fraud detection with explainable AI tools,” *Digital Threats: Research and Practice*, vol. 7, no. 1, Art. no. 7, Mar. 2026.
- [12] A. Somasundaram and S. Reddy, “Parallel and incremental credit card fraud detection model to handle concept drift and data imbalance,” *Neural Computing and Applications*, vol. 31, Suppl. 1, pp. S3–S14, 2019.
- [13] M. Kim and K.-B. Hwang, “An empirical evaluation of sampling methods for the classification of imbalanced data,” *PLOS ONE*, vol. 17, no. 7, Art. no. e0271260, Jul. 2022.
- [14] J. T. Hancock, T. M. Khoshgoftaar, and J. M. Johnson, “Evaluating classifier performance with highly imbalanced Big Data,” *Journal of Big Data*, vol. 10, Art. no. 42, 2023.
- [15] N. Seliya, A. A. Zadeh, and T. M. Khoshgoftaar, “A literature review on one-class classification and its potential applications in big data,” *Journal of Big Data*, vol. 8, Art. no. 122, 2021.
- [16] J. Adejoh, N. Owoh, M. Ashawa, S. Hosseinzadeh, A. Shahrabi, and S. Mohamed, “An adaptive unsupervised learning approach for credit card fraud detection,” *Big Data and Cognitive Computing*, vol. 9, Art. no. 217, 2025.
- [17] S. Khodabandehlou and A. Hashemi Golpayegani, “FiFrauD: Unsupervised financial fraud detection in dynamic graph streams,” *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 5, Art. no. 111, Feb. 2024.
- [18] Z. Liu, H. Yu, and X. Luo, “A semi-supervised fraud detection model based on fuzzy graph neural networks,” *Applied Soft Computing*, vol. 202, Art. no. 115913, Jul. 2026.
- [19] T. R. Bandaragoda, K. M. Ting, D. Albrecht, F. T. Liu, Y. Zhu, and J. R. Wells, “Isolation-based anomaly detection using nearest-neighbor ensembles,” *Computational Intelligence*, vol. 34, no. 4, pp. 968–998, 2018.
- [20] A. Bakumenko and A. Elragal, “Detecting anomalies in financial data using machine learning algorithms,” *Systems*, vol. 10, no. 5, Art. no. 130, 2022.

- [21] H. Xu, G. Pang, Y. Wang, and Y. Wang, "Deep isolation forest for anomaly detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 12, pp. 12591–12604, Dec. 2023.
- [22] J. Lesouple, C. Baudoin, M. Spigai, and J.-Y. Tourneret, "Generalized isolation forest for anomaly detection," *Pattern Recognition Letters*, vol. 149, pp. 109–119, Sep. 2021.
- [23] V. Yepmo, G. Smits, M.-J. Lesot, and O. Pivert, "Leveraging an isolation forest to anomaly detection and data clustering," *Data & Knowledge Engineering*, vol. 151, Art. no. 102302, May 2024.
- [24] O. Alghushairy, R. Alsini, T. Soule, and X. Ma, "A review of Local Outlier Factor algorithms for outlier detection in big data streams," *Big Data and Cognitive Computing*, vol. 5, no. 1, Art. no. 1, 2021.
- [25] S. Shahrestani and I. Sanislav, "How does dimensionality influence outlier detection effectiveness in multivariate geochemical data? Insights from LOF and IF methods," *Earth Science Informatics*, vol. 18, no. 1, Art. no. 27, 2025.
- [26] Z. Zhang, Y. Hou, Y. Jia, and R. Zhang, "An outlier detection algorithm based on local density feedback," *Knowledge and Information Systems*, vol. 67, no. 4, pp. 3599–3629, 2025.
- [27] L. Hosseini, F. Movahedi Sobhani, and J. Kazemitabar, "Unsupervised fraud detection in banking transactions using fuzzy clustering and outlier detection techniques," *Journal of Fuzzy Extension and Applications*, vol. 7, no. 2, pp. 530–551, 2026.
- [28] J. Zhang, Y. Qin, H. Zhang, W. Hu, and X. Bai, "Application of a local outlier detection algorithm based on high-dimensional subspaces in near-infrared spectroscopy," *Journal of Analytical Atomic Spectrometry*, vol. 40, no. 9, pp. 2548–2555, 2025.
- [29] A. Smiti, "A critical overview of outlier detection methods," *Computer Science Review*, vol. 38, Art. no. 100306, 2020.
- [30] Y. Zhou, H. Xia, D. Yu, J. Cheng, and J. Li, "Outlier detection method based on high-density iteration," *Information Sciences*, vol. 662, Art. no. 120286, 2024.
- [31] R. Bouman, Z. Bukhsh, and T. Heskes, "Unsupervised anomaly detection algorithms on real-world data: How many do we need?," *Journal of Machine Learning Research*, vol. 25, no. 105, pp. 1–34, 2024.
- [32] E. A. Hinojosa-Palafox, O. M. Rodríguez-Elías, J. H. Pacheco-Ramírez, J. A. Hoyo-Montaño, M. Pérez-Patricio, and D. F. Espejel-Blanco, "A novel unsupervised anomaly detection framework for early fault detection in complex industrial settings," *IEEE Access*, vol. 12, pp. 181823–181845, 2024.
- [33] K. Demestichas, N. Peppes, T. Alexakis, and E. Adamopoulou, "An advanced abnormal behavior detection engine embedding autoencoders for the investigation of financial transactions," *Information*, vol. 12, no. 1, Art. no. 34, Jan. 2021.
- [34] D.-M. Tsai and P.-H. Jen, "Autoencoder-based anomaly detection for surface defect inspection," *Advanced Engineering Informatics*, vol. 48, Art. no. 101272, Apr. 2021.
- [35] H. Torabi, S. L. Mirtaheri, and S. Greco, "Practical autoencoder based anomaly detection by using vector reconstruction error," *Cybersecurity*, vol. 6, Art. no. 1, Jan. 2023.
- [36] Z. Cheng, S. Wang, P. Zhang, S. Wang, X. Liu, and E. Zhu, "Improved autoencoder for unsupervised anomaly detection," *International Journal of Intelligent Systems*, vol. 36, no. 12, pp. 7103–7125, Dec. 2021.
- [37] M. Tayebi and S. El Kafhali, "Combining autoencoders and deep learning for effective fraud detection in credit card transactions," *Operations Research Forum*, vol. 6, Art. no. 8, Jan. 2025.
- [38] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, and G. Obaido, "A neural network ensemble with feature engineering for improved credit card fraud detection," *IEEE Access*, vol. 10, pp. 16400–16407, 2022.
- [39] A. R. Khalid, N. Owoh, O. Uthmani, M. Ashawa, J. Osamor, and J. Adejoh, "Enhancing credit card fraud detection: An ensemble machine learning approach," *Big Data and Cognitive Computing*, vol. 8, no. 1, Art. no. 6, 2024.
- [40] S. Kandanaarachchi, "Unsupervised anomaly detection ensembles using item response theory," *Information Sciences*, vol. 587, pp. 142–163, 2022.
- [41] N. Baghalizadeh Moghadam, F. Cuppens, and N. Boulahia-Cuppens, "Explainable multi-modal unsupervised learning for insider threat detection in enterprise environments," *Journal of Information Security and Applications*, vol. 102, Art. no. 104581, 2026.
- [42] J. Forough and S. Momtazi, "Ensemble of deep sequential models for credit card fraud detection," *Applied Soft Computing*, vol. 99, Art. no. 106883, Feb. 2021.

- [43] F. Craighero, F. Angaroni, F. Stella, C. Damiani, M. Antoniotti, and A. Graudenzi, "Unity is strength: Improving the detection of adversarial examples with ensemble approaches," *Neurocomputing*, vol. 554, Art. no. 126576, 2023.
- [44] E. G. Harb, N. Nasrallah, R. El Khoury, and K. Hussainey, "Applying Benford's law to detect accounting data manipulation in the pre- and post-financial engineering periods," *Journal of Applied Accounting Research*, vol. 24, no. 4, pp. 745–768, 2023.
- [45] E. Druică, B. Oancea, and C. Vălsan, "Benford's law and the limits of digit analysis," *International Journal of Accounting Information Systems*, vol. 31, pp. 75–82, 2018.
- [46] E. Badal-Valero, J. A. Alvarez-Jareño, and J. M. Pavía, "Combining Benford's law and machine learning to detect money laundering. An actual Spanish court case," *Forensic Science International*, vol. 282, pp. 24–34, 2018.
- [47] L. Barabesi, A. Cerasa, A. Cerioli, and D. Perrotta, "On characterizations and tests of Benford's law," *Journal of the American Statistical Association*, vol. 117, no. 540, pp. 1887–1903, 2022.
- [48] L. Barabesi, A. Cerioli, A. Cerasa, and D. Perrotta, "Statistical detection of adversarial compliance with Benford's law," *Statistical Analysis and Data Mining: An ASA Data Science Journal*, vol. 19, Art. no. e70081, 2026.
- [49] H. R. Morales, M. Porporato, and N. Epelbaum, "Benford's law for integrity tests of high-volume databases: A case study of internal audit in a state-owned enterprise," *Journal of Economics, Finance and Administrative Science*, vol. 27, no. 53, pp. 154–174, 2022.
- [50] Wang, J., Liu, J., Zheng, W., & Ge, Y., Temporal Heterogeneous Graph Contrastive Learning for Fraud Detection in Credit Card Transactions, *IEEE Access*, 2025.
- [51] Qazi, A., & Jadoon, A. A. K., Explainable Ensemble Learning Using SHAP for ERP Anomaly Detection, *Scientific Reports*, 2026.
- [52] J. Miao and W. Zhu, "Precision–recall curve (PRC) classification trees," *Evolutionary Intelligence*, vol. 15, pp. 1545–1569, 2022, doi: 10.1007/s12065-021-00565-2.
- [53] F. Movahedi, R. Padman, and J. F. Antaki, "Limitations of receiver operating characteristic curve on imbalanced data: Assist device mortality risk scores," *The Journal of Thoracic and Cardiovascular Surgery*, vol. 165, no. 4, pp. 1433–1442, 2023.
- [54] E. Birihanu and I. Lendák, "Explainable correlation-based anomaly detection for Industrial Control Systems," *Frontiers in Artificial Intelligence*, vol. 7, Art. no. 1508821, 2025, doi: 10.3389/frai.2024.1508821.
- [55] J. Chen, Q. Chen, F. Jiang, X. Guo, K. Sha, and Y. Wang, "SCN_GNN: A GNN-based fraud detection algorithm combining strong node and graph topology information," *Expert Systems with Applications*, vol. 237, Art. no. 121643, 2024, doi: 10.1016/j.eswa.2023.121643.
- [56] T. G. Dietterich, "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms," *Neural Computation*, vol. 10, no. 7, pp. 1895–1923, 1998, doi: 10.1162/089976698300017197.
- [57] S. Holm, "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979, doi: 10.2307/4615733.
- [58] L. S. Shapley, "A Value for n-Person Games," in *Contributions to the Theory of Games, Volume II*, H. W. Kuhn and A. W. Tucker, Eds. Princeton, NJ, USA: Princeton University Press, 1953, pp. 307–318.
- [59] H. Sahlaoui, E. A. Abdellaoui Alaoui, A. Nayyar, S. Agoujl, and M. M. Jaber, "Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations," *IEEE Access*, vol. 9, pp. 152688–152703, 2021, doi: 10.1109/ACCESS.2021.3124270.
- [60] K. Lin and Y. Gao, "Model interpretability of financial fraud detection by group SHAP," *Expert Systems with Applications*, vol. 210, Art. no. 118354, 2022, doi: 10.1016/j.eswa.2022.118354.

Παράρτημα Α: Αποτελέσματα Παραμετροποίησης μοντέλων

Πίνακας Α.1. Συγκεντρωτικά αποτελέσματα παραμετροποίησης του Isolation Forest για τους 108 εξεταζόμενους συνδυασμούς υπερπαραμέτρων.

n_estimators	max_samples	max_features	contamination	Precision	Recall	F1-score	AUPRC	ROC-AUC	TN	FP	FN	TP
100	128	0,4	0,035	0,1685	0,1686	0,1685	0,1031	0,7154	552691	17186	17180	3483
100	128	0,4	0,06	0,1538	0,2637	0,1942	0,1031	0,7154	539892	29985	15215	5448
100	128	0,4	0,08	0,1443	0,3298	0,2007	0,1031	0,7154	529448	40429	13848	6815
100	128	0,4	0,1	0,134	0,3829	0,1985	0,1031	0,7154	518734	51143	12752	7911
100	256	0,4	0,035	0,1717	0,1717	0,1717	0,1045	0,7221	552757	17120	17115	3548
100	256	0,4	0,06	0,1545	0,265	0,1952	0,1045	0,7221	539920	29957	15187	5476
100	256	0,4	0,08	0,1407	0,3218	0,1958	0,1045	0,7221	529282	40595	14014	6649
100	256	0,4	0,1	0,1301	0,3719	0,1928	0,1045	0,7221	518508	51369	12978	7685
100	512	0,4	0,035	0,1808	0,1808	0,1808	0,1064	0,7097	552944	16933	16927	3736
100	512	0,4	0,06	0,1574	0,2699	0,1988	0,1064	0,7097	540020	29857	15087	5576
100	512	0,4	0,08	0,1431	0,3272	0,1991	0,1064	0,7097	529393	40484	13903	6760
100	512	0,4	0,1	0,1297	0,3708	0,1922	0,1064	0,7097	518484	51393	13002	7661
100	128	0,6	0,035	0,1784	0,1784	0,1784	0,1074	0,7156	552895	16982	16976	3687
100	128	0,6	0,06	0,1631	0,2797	0,206	0,1074	0,7156	540223	29654	14884	5779
100	128	0,6	0,08	0,1475	0,3372	0,2052	0,1074	0,7156	529601	40276	13695	6968
100	128	0,6	0,1	0,1351	0,3862	0,2002	0,1074	0,7156	518803	51074	12683	7980
100	256	0,6	0,035	0,1907	0,1907	0,1907	0,1121	0,7183	553149	16728	16722	3941
100	256	0,6	0,06	0,1642	0,2815	0,2074	0,1121	0,7183	540261	29616	14846	5817
100	256	0,6	0,08	0,1467	0,3354	0,2041	0,1121	0,7183	529564	40313	13732	6931
100	256	0,6	0,1	0,1334	0,3813	0,1977	0,1121	0,7183	518702	51175	12784	7879
100	512	0,6	0,035	0,1932	0,1933	0,1933	0,1123	0,7174	553202	16675	16669	3994
100	512	0,6	0,06	0,1629	0,2793	0,2058	0,1123	0,7174	540216	29661	14892	5771
100	512	0,6	0,08	0,1451	0,3318	0,2019	0,1123	0,7174	529488	40389	13808	6855

100	512	0,6	0,1	0,1307	0,3735	0,1936	0,1123	0,7174	518541	51336	12945	7718
100	128	0,8	0,035	0,1623	0,1624	0,1623	0,1009	0,7181	552563	17314	17308	3355
100	128	0,8	0,06	0,1482	0,2541	0,1872	0,1009	0,7181	539694	30183	15413	5250
100	128	0,8	0,08	0,1385	0,3167	0,1927	0,1009	0,7181	529176	40701	14120	6543
100	128	0,8	0,1	0,1304	0,3727	0,1932	0,1009	0,7181	518525	51352	12961	7702
100	256	0,8	0,035	0,1815	0,1815	0,1815	0,1085	0,7219	552959	16918	16912	3751
100	256	0,8	0,06	0,1579	0,2708	0,1995	0,1085	0,7219	540040	29837	15067	5596
100	256	0,8	0,08	0,1439	0,329	0,2002	0,1085	0,7219	529431	40446	13865	6798
100	256	0,8	0,1	0,1332	0,3806	0,1973	0,1085	0,7219	518688	51189	12798	7865
100	512	0,8	0,035	0,1728	0,1728	0,1728	0,1028	0,7129	552779	17098	17092	3571
100	512	0,8	0,06	0,1545	0,2649	0,1951	0,1028	0,7129	539917	29960	15190	5473
100	512	0,8	0,08	0,14	0,3202	0,1949	0,1028	0,7129	529249	40628	14047	6616
100	512	0,8	0,1	0,1289	0,3685	0,1911	0,1028	0,7129	518438	51439	13048	7615
200	128	0,4	0,035	0,1958	0,1959	0,1958	0,1155	0,7293	553255	16622	16616	4047
200	128	0,4	0,06	0,1696	0,2908	0,2142	0,1155	0,7293	540453	29424	14654	6009
200	128	0,4	0,08	0,1561	0,357	0,2173	0,1155	0,7293	530010	39867	13286	7377
200	128	0,4	0,1	0,1419	0,4058	0,2104	0,1115	0,7235	519203	50674	12283	8380
200	128	0,6	0,035	0,194	0,194	0,194	0,1134	0,721	553217	16660	16654	4009
200	128	0,6	0,06	0,17	0,2915	0,2148	0,1134	0,721	540468	29409	14639	6024
200	128	0,6	0,08	0,1517	0,3469	0,2111	0,1134	0,721	529802	40075	13494	7169
200	128	0,6	0,1	0,1389	0,3969	0,2058	0,1134	0,721	519025	50852	12461	8202
200	128	0,8	0,035	0,1861	0,1861	0,1861	0,1115	0,7254	553054	16823	16817	3846
200	128	0,8	0,06	0,1656	0,284	0,2092	0,1115	0,7254	540313	29564	14794	5869
200	128	0,8	0,08	0,153	0,3499	0,2129	0,1115	0,7254	529862	40015	13434	7229
200	128	0,8	0,1	0,1402	0,4007	0,2077	0,1115	0,7254	519103	50774	12383	8280
200	256	0,4	0,035	0,1895	0,1895	0,1895	0,1132	0,7306	553124	16753	16747	3916
200	256	0,4	0,06	0,1641	0,2814	0,2073	0,1132	0,7306	540258	29619	14849	5814
200	256	0,4	0,08	0,1479	0,3381	0,2058	0,1132	0,7306	529620	40257	13676	6987
200	256	0,4	0,1	0,1364	0,3898	0,2021	0,1132	0,7306	518877	51000	12609	8054
200	256	0,6	0,035	0,1915	0,1916	0,1915	0,1127	0,7193	553166	16711	16705	3958

200	256	0,6	0,06	0,1666	0,2857	0,2105	0,1127	0,7193	540347	29530	14760	5903
200	256	0,6	0,08	0,1498	0,3425	0,2084	0,1127	0,7193	529710	40167	13586	7077
200	256	0,6	0,1	0,1354	0,387	0,2006	0,1127	0,7193	518819	51058	12667	7996
200	256	0,8	0,035	0,1944	0,1945	0,1945	0,1149	0,7213	553227	16650	16644	4019
200	256	0,8	0,06	0,1672	0,2867	0,2112	0,1149	0,7213	540368	29509	14739	5924
200	256	0,8	0,08	0,1512	0,3456	0,2103	0,1149	0,7213	529775	40102	13521	7142
200	256	0,8	0,1	0,1366	0,3905	0,2024	0,1149	0,7213	518892	50985	12594	8069
200	512	0,4	0,035	0,1931	0,1932	0,1932	0,1139	0,7261	553200	16677	16671	3992
200	512	0,4	0,06	0,1669	0,2863	0,2109	0,1139	0,7261	540359	29518	14748	5915
200	512	0,4	0,08	0,1494	0,3416	0,2079	0,1139	0,7261	529692	40185	13604	7059
200	512	0,4	0,1	0,1352	0,3863	0,2003	0,1139	0,7261	518806	51071	12680	7983
200	512	0,6	0,035	0,1934	0,1934	0,1934	0,1119	0,7127	553205	16672	16666	3997
200	512	0,6	0,06	0,1639	0,2811	0,2071	0,1119	0,7127	540253	29624	14854	5809
200	512	0,6	0,08	0,1464	0,3348	0,2037	0,1119	0,7127	529550	40327	13746	6917
200	512	0,6	0,1	0,1323	0,3781	0,196	0,1119	0,7127	518635	51242	12851	7812
200	512	0,8	0,035	0,1901	0,1902	0,1902	0,1103	0,7139	553138	16739	16733	3930
200	512	0,8	0,06	0,1637	0,2807	0,2068	0,1103	0,7139	540244	29633	14863	5800
200	512	0,8	0,08	0,1471	0,3363	0,2046	0,1103	0,7139	529581	40296	13715	6948
200	512	0,8	0,1	0,1329	0,3797	0,1968	0,1103	0,7139	518669	51208	12817	7846
300	128	0,4	0,035	0,1913	0,1913	0,1913	0,1139	0,7263	553161	16716	16710	3953
300	128	0,4	0,06	0,1667	0,2859	0,2106	0,1139	0,7263	540352	29525	14755	5908
300	128	0,4	0,08	0,1549	0,3541	0,2155	0,1139	0,7263	529950	39927	13346	7317
300	128	0,4	0,1	0,1418	0,4053	0,2101	0,1139	0,7263	519197	50680	12289	8374
300	256	0,4	0,035	0,188	0,1881	0,188	0,1122	0,7243	553094	16783	16777	3886
300	256	0,4	0,06	0,1632	0,2798	0,2061	0,1122	0,7243	540225	29652	14882	5781
300	256	0,4	0,08	0,1475	0,3372	0,2052	0,1122	0,7243	529601	40276	13695	6968
300	256	0,4	0,1	0,1349	0,3857	0,1999	0,1122	0,7243	518792	51085	12694	7969
300	512	0,4	0,035	0,1889	0,189	0,189	0,1114	0,7214	553113	16764	16758	3905
300	512	0,4	0,06	0,1633	0,28	0,2063	0,1114	0,7214	540229	29648	14878	5785
300	512	0,4	0,08	0,1478	0,3379	0,2056	0,1114	0,7214	529614	40263	13682	6981

300	512	0,4	0,1	0,1336	0,3818	0,198	0,1114	0,7214	518713	51164	12773	7890
300	128	0,6	0,035	0,1878	0,1879	0,1878	0,1126	0,7218	553090	16787	16781	3882
300	128	0,6	0,06	0,1672	0,2868	0,2113	0,1126	0,7218	540370	29507	14737	5926
300	128	0,6	0,08	0,1509	0,345	0,21	0,1126	0,7218	529762	40115	13534	7129
300	128	0,6	0,1	0,1388	0,3968	0,2057	0,1126	0,7218	519022	50855	12464	8199
300	256	0,6	0,035	0,1871	0,1871	0,1871	0,1125	0,7218	553075	16802	16796	3867
300	256	0,6	0,06	0,1658	0,2843	0,2095	0,1125	0,7218	540319	29558	14788	5875
300	256	0,6	0,08	0,1497	0,3423	0,2083	0,1125	0,7218	529705	40172	13591	7072
300	256	0,6	0,1	0,1358	0,3882	0,2013	0,1125	0,7218	518845	51032	12641	8022
300	512	0,6	0,035	0,1905	0,1906	0,1906	0,1117	0,7119	553146	16731	16727	3936
300	512	0,6	0,06	0,1633	0,28	0,2063	0,1117	0,7119	540229	29648	14878	5785
300	512	0,6	0,08	0,1456	0,3329	0,2026	0,1117	0,7119	529512	40365	13784	6879
300	512	0,6	0,1	0,1319	0,3771	0,1955	0,1117	0,7119	518614	51263	12872	7791
300	128	0,8	0,035	0,187	0,187	0,187	0,1127	0,7283	553073	16804	16798	3865
300	128	0,8	0,06	0,1648	0,2827	0,2083	0,1127	0,7283	540285	29592	14822	5841
300	128	0,8	0,08	0,1531	0,35	0,213	0,1127	0,7283	529866	40011	13430	7233
300	128	0,8	0,1	0,1406	0,4019	0,2083	0,1127	0,7283	519127	50750	12359	8304
300	256	0,8	0,035	0,1922	0,1923	0,1922	0,1139	0,7232	553181	16696	16690	3973
300	256	0,8	0,06	0,1667	0,2858	0,2105	0,1139	0,7232	540349	29528	14758	5905
300	256	0,8	0,08	0,1502	0,3434	0,209	0,1139	0,7232	529729	40148	13567	7096
300	256	0,8	0,1	0,1365	0,3902	0,2023	0,1139	0,7232	518886	50991	12600	8063
300	512	0,8	0,035	0,1826	0,1827	0,1827	0,1079	0,7113	552983	16894	16888	3775
300	512	0,8	0,06	0,1607	0,2756	0,203	0,1079	0,7113	540139	29738	14968	5695
300	512	0,8	0,08	0,1439	0,3291	0,2003	0,1079	0,7113	529433	40444	13863	6800
300	512	0,8	0,1	0,1315	0,3757	0,1948	0,1079	0,7113	518586	51291	12900	7763

Πίνακας Α.2. Συγκεντρωτικά αποτελέσματα παραμετροποίησης του LOF για τους 16 εξεταζόμενους συνδυασμούς υπερπαραμέτρων.

n_neighbors	contamination	TN	FP	FN	TP	Precision	Recall	F1-score	AUPRC	ROC_AUC
50	0.035	93280	3221	3220	279	0.079714	0.079737	0.079726	0.054949	0.610246
50	0.060	90931	5570	3069	430	0.071667	0.122892	0.090536	0.054949	0.610246
50	0.080	89066	7435	2934	565	0.070625	0.161475	0.098269	0.054949	0.610246
50	0.100	87185	9316	2815	684	0.071400	0.204184	0.105841	0.054949	0.605746
20	0.035	93244	3257	3256	243	0.069429	0.069448	0.069438	0.053215	0.608560
20	0.060	90916	5585	3084	415	0.069167	0.118605	0.087378	0.053215	0.608560
20	0.080	89001	7500	2999	500	0.062500	0.142898	0.086964	0.053215	0.608560
20	0.100	87131	9370	2869	630	0.063000	0.180051	0.093340	0.053215	0.608560
10	0.035	93257	3244	3243	256	0.073143	0.073164	0.073153	0.052269	0.599406

10	0.060	90930	5571	3070	429	0.071500	0.122606	0.090325	0.052269	0.599406
10	0.080	89046	7455	2954	545	0.068125	0.155759	0.094791	0.052269	0.599406
10	0.100	87140	9361	2860	639	0.063900	0.182624	0.094674	0.052269	0.599406
5	0.035	93237	3264	3263	236	0.067429	0.067448	0.067438	0.048920	0.578545
5	0.060	90881	5620	3119	380	0.063333	0.108602	0.080008	0.048920	0.578545
5	0.080	88989	7512	3011	488	0.061000	0.139468	0.084877	0.048920	0.578545
5	0.100	87098	9403	2902	597	0.059700	0.170620	0.088451	0.048920	0.578545

Πίνακας Α.3. Συγκεντρωτικά αποτελέσματα παραμετροποίησης του HBOS για τους 16 εξεταζόμενους συνδυασμούς υπερπαραμέτρων.

contamination	n_bins	TN	FP	FN	TP	Precision	Recall	F1-score	AUPRC	ROC_AUC
0.035	50	553020	16857	16851	3812	0.184431	0.184484	0.184458	0.100906	0.666846
0.060	50	540133	29744	14974	5689	0.160557	0.275323	0.202831	0.100906	0.666846
0.100	50	517747	52130	13743	6920	0.119689	0.341838	0.177223	0.100906	0.655746
0.080	50	529029	40848	14267	6396	0.135382	0.309539	0.188375	0.100906	0.666846
0.080	30	528798	41079	14507	6156	0.130327	0.297924	0.181331	0.099159	0.660085
0.100	30	517563	52314	13923	6740	0.114133	0.326187	0.169098	0.099159	0.660085
0.060	30	539985	29892	15122	5541	0.156380	0.268160	0.197554	0.099159	0.660085
0.035	30	553078	16799	16793	3870	0.187237	0.187291	0.187264	0.099159	0.660085
0.100	20	517557	52320	13935	6728	0.113941	0.325606	0.168810	0.097745	0.656841
0.080	20	528778	41099	14548	6115	0.129517	0.295940	0.180179	0.097745	0.656841
0.035	20	553046	16831	16825	3838	0.185689	0.185743	0.185716	0.097745	0.656841
0.060	20	539875	30002	15232	5431	0.153275	0.262837	0.193632	0.097745	0.656841
0.060	10	539713	30164	15394	5269	0.148703	0.254997	0.187857	0.096286	0.619742
0.035	10	553138	16739	16733	3930	0.190140	0.190195	0.190167	0.096286	0.619742

0.080	10	528534	41343	14762	5901	0.124905	0.285583	0.173797	0.096286	0.619742
0.100	10	517485	52392	14002	6661	0.112797	0.322364	0.167118	0.096286	0.619742

Πίνακας Α.4. Συγκεντρωτικά αποτελέσματα παραμετροποίησης του Autoencoder για τους 36 εξεταζόμενους συνδυασμούς υπερπαραμέτρων.

latent_dim	epochs	contamination	TN	FP	FN	TP	Precision	Recall	F1-score	AUPRC	ROC_AUC
8	20	0.035	93710	2791	2790	709	0.202571	0.202629	0.202600	0.130396	0.667868
8	20	0.060	91413	5088	2587	912	0.152000	0.260646	0.192020	0.130396	0.667868
8	20	0.080	89557	6944	2443	1056	0.132000	0.301801	0.183668	0.130396	0.667868

8	20	0.100	87679	8822	2321	1178	0.117800	0.336668	0.174531	0.130396	0.667868
8	50	0.035	93697	2804	2803	696	0.198857	0.198914	0.198886	0.133326	0.668698
8	50	0.060	91408	5093	2592	907	0.151167	0.259217	0.190967	0.133326	0.668698
8	50	0.080	89531	6970	2469	1030	0.128750	0.294370	0.179146	0.133326	0.668698
8	50	0.100	87646	8855	2354	1145	0.114500	0.327236	0.169642	0.133326	0.668698
8	100	0.035	93705	2796	2795	704	0.201143	0.201200	0.201172	0.140699	0.671805
8	100	0.060	91398	5103	2602	897	0.149500	0.256359	0.188862	0.140699	0.671805
8	100	0.080	89526	6975	2474	1025	0.128125	0.292941	0.178276	0.140699	0.671805
8	100	0.100	87656	8845	2344	1155	0.115500	0.330094	0.171124	0.140699	0.671805
16	20	0.035	93718	2783	2782	717	0.204857	0.204916	0.204886	0.133286	0.670850
16	20	0.060	91413	5088	2587	912	0.152000	0.260646	0.192020	0.133286	0.670850
16	20	0.080	89559	6942	2441	1058	0.132250	0.302372	0.184016	0.133286	0.670850
16	20	0.100	87687	8814	2313	1186	0.118600	0.338954	0.175717	0.133286	0.670850
16	50	0.035	93716	2785	2784	715	0.204286	0.204344	0.204315	0.137436	0.681214
16	50	0.060	91447	5054	2553	946	0.157667	0.270363	0.199179	0.137436	0.681214
16	50	0.080	89590	6911	2410	1089	0.136125	0.311232	0.189408	0.137436	0.681214

16	50	0.100	87702	8799	2298	1201	0.120100	0.343241	0.177939	0.137436	0.681214
16	100	0.035	93708	2793	2792	707	0.202000	0.202058	0.202029	0.141083	0.683877
16	100	0.060	91432	5069	2568	931	0.155167	0.266076	0.196021	0.141083	0.683877
16	100	0.080	89571	6930	2429	1070	0.133750	0.305802	0.186103	0.141083	0.683877
16	100	0.100	87696	8805	2304	1195	0.119500	0.341526	0.177050	0.141083	0.683877
32	20	0.035	93716	2785	2784	715	0.204286	0.204344	0.204315	0.136170	0.682246
32	20	0.060	91441	5060	2559	940	0.156667	0.268648	0.197916	0.136170	0.682246
32	20	0.080	89600	6901	2400	1099	0.137375	0.314090	0.191147	0.136170	0.682246
32	20	0.100	87730	8771	2270	1229	0.122900	0.351243	0.182088	0.136170	0.682246
32	50	0.035	93723	2778	2777	722	0.206286	0.206345	0.206315	0.140630	0.684773
32	50	0.060	91447	5054	2553	946	0.157667	0.270363	0.199179	0.140630	0.684773
32	50	0.080	89588	6913	2412	1087	0.135875	0.310660	0.189060	0.140630	0.684773
32	50	0.100	87706	8795	2294	1205	0.120500	0.344384	0.178532	0.140630	0.684773
32	100	0.035	93732	2769	2768	731	0.208857	0.208917	0.208887	0.141883	0.685315
32	100	0.060	91458	5043	2542	957	0.159500	0.273507	0.201495	0.141883	0.685315
32	100	0.080	89594	6907	2406	1093	0.136625	0.312375	0.190103	0.141883	0.685315

32	100	0.100	87731	8770	2269	1230	0.124500	0.355829	0.184536	0.139523	0.690515
----	-----	-------	-------	------	------	------	----------	----------	----------	----------	----------

