

2026-02

$$\begin{aligned} & \text{p}\ddot{\text{y}} \text{ — } \S \acute{\text{A}} \textcircled{\text{R}} \tilde{\text{A}} \cdot \ddot{\text{A}} \cdot \hat{\text{A}} \propto \mu \zeta^{1/2} \cdot \ddot{\text{A}} \textcircled{\text{R}} \hat{\text{A}} \cdot \zeta \cdot \frac{1}{4} \zeta \\ & \text{p}\ddot{\text{y}} \S \acute{\text{A}} \cdot \frac{1}{4} \pm \ddot{\text{A}} \zeta \zeta^{1^0} \zeta^{1/2} \zeta^{1/4} 1^0 - \hat{\text{A}} \cdot \acute{\text{A}}^3 \pm \tilde{\text{A}}^- \mu \end{aligned}$$

$$\text{p}\ddot{\text{y}} \propto \P \mu \cdot \acute{\text{A}} \pm \frac{1}{2} \text{ — } \zeta \text{ Å} , \quad \text{£}^2 \mu \ddot{\text{A}} \gg \neg \frac{1}{2} \pm$$

$$\begin{aligned} & \text{p}\ddot{\text{y}} \propto \mu \ddot{\text{A}} \pm \ddot{\text{A}} \text{Å} \zeta^{1 \pm 0} \text{Ì} \quad \acute{\text{A}} \text{Ì}^3 \acute{\text{A}} \pm \frac{1}{4} \frac{1}{4} \pm \tilde{\text{A}} \ddot{\text{A}} \cdot \frac{1}{2} \quad \frac{1}{2} \neg \gg \text{Å} \tilde{\text{A}} \cdot \quad " \mu' \zeta^{1/4} - \frac{1}{2} \acute{\text{E}} \frac{1}{2} \quad 0 \pm 1 \quad \S \acute{\text{A}} \cdot \frac{1}{4} \pm \ddot{\text{A}} \zeta \zeta^{1^0} \zeta^{1/2} \zeta \\ & \text{p}\ddot{\text{y}} \text{£} \zeta \gg \textcircled{\text{R}} \quad \ddot{\text{Y}}^{1^0} \zeta^{1/2} \zeta^{1/4} 1^0 \hat{1} \frac{1}{2} \cdot \text{À}^1 \tilde{\text{A}} \ddot{\text{A}} \cdot \frac{1}{4} \hat{1} \frac{1}{2} \quad 0 \pm 1 \quad " 1 \zeta^{-0} \cdot \tilde{\text{A}} \cdot \hat{\text{A}} , \quad \pm \frac{1}{2} \mu \text{À}^1 \tilde{\text{A}} \textcircled{\text{R}} \frac{1}{4} 1 \zeta \cdot \mu \neg \text{À} \zeta \gg 1 \hat{\text{A}} \end{aligned}$$

<http://hdl.handle.net/11728/13545>

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository



**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΤΕΧΝΟΛΟΓΙΑ**

Πανεπιστήμιο Νεάπολις Πάφος

**«Η ΧΡΗΣΗ ΤΗΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ ΣΕ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΕΣ ΕΡΓΑΣΙΕΣ»**

**Σβεtlάνα Τζεϊρανίδου
Φεβρουάριος, 2026**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΤΕΧΝΟΛΟΓΙΑ**

**«Η ΧΡΗΣΗ ΤΗΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ ΣΕ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΕΣ ΕΡΓΑΣΙΕΣ»**

**Διπλωματική Εργασία η οποία υποβλήθηκε προς απόκτηση εξ
αποστάσεως μεταπτυχιακού τίτλου σπουδών στην Ανάλυση
Δεδομένων και Χρηματοοικονομική Τεχνολογία στο
Πανεπιστήμιο Νεάπολις Πάφος.**

**Σβετλάνα Τζεϊρανίδου
Φεβρουάριος, 2026**

Πνευματικά δικαιώματα

Copyright © Σβεtlάνα Τζεϊρανίδου, 2026

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της παρούσας Διπλωματικής Εργασίας από το Πανεπιστημίου Νεάπολιν δεν υποδηλώνει απαραιτήτως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Πανεπιστημίου.

Η ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Η Σβετλάνα Τζεϊρανίδου, γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα ότι η παρούσα εργασία με τίτλο « Η Χρήση της Τεχνητής Νοημοσύνης σε Χρηματοοικονομικές Εργασίες», αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές που έχω χρησιμοποιήσει, έχουν δηλωθεί κατάλληλα στις βιβλιογραφικές παραπομπές και αναφορές. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

Η Δηλούσα

Σβετλάνα Τζεϊρανίδου

Πρόλογος και Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο του Προγράμματος Μεταπτυχιακών Σπουδών στην Ανάλυση Δεδομένων και Χρηματοοικονομική Τεχνολογία του Πανεπιστημίου Νεάπολις Πάφος και αποτελεί το επιστέγασμα της ακαδημαϊκής μου πορείας στο συγκεκριμένο πρόγραμμα. Η επιλογή του θέματος προέκυψε από το έντονο ενδιαφέρον μου για τις σύγχρονες τεχνολογικές εξελίξεις και τον τρόπο με τον οποίο η Τεχνητή Νοημοσύνη και η Μηχανική Μάθηση επηρεάζουν και μετασχηματίζουν τον χρηματοοικονομικό τομέα. Ιδιαίτερα, η εφαρμογή των αλγοριθμικών μοντέλων στη διαχείριση πιστωτικού κινδύνου αποτέλεσε ένα πεδίο που συνδυάζει θεωρητική πρόκληση και πρακτική σημασία.

Κατά τη διάρκεια της εκπόνησης της εργασίας, ήρθα αντιμέτωπη με ποικίλες προκλήσεις, τόσο σε επίπεδο κατανόησης και ανάλυσης σύνθετων θεωρητικών εννοιών όσο και σε πρακτικό επίπεδο, κυρίως ως προς τη διαχείριση και επεξεργασία πραγματικών δεδομένων και την αξιολόγηση των αποτελεσμάτων των μοντέλων. Η διαδικασία αυτή, ωστόσο, υπήρξε ιδιαίτερα δημιουργική και εκπαιδευτική, καθώς συνέβαλε ουσιαστικά στην εμβάθυνση των γνώσεών μου, στην ανάπτυξη αναλυτικής σκέψης και στην απόκτηση πολύτιμων ερευνητικών δεξιοτήτων.

Ιδιαίτερες ευχαριστίες οφείλω στον επιβλέποντα καθηγητή μου, Δρ. Κωνσταντίνο Παναγιωτάκη, για την επιστημονική καθοδήγηση, τις ουσιαστικές παρατηρήσεις και τη συνεχή υποστήριξή του καθ' όλη τη διάρκεια της εκπόνησης της εργασίας. Η συμβολή του υπήρξε καθοριστική τόσο στη διαμόρφωση της ερευνητικής προσέγγισης όσο και στη βελτίωση της ποιότητας του τελικού αποτελέσματος.

Θα ήθελα, επίσης, να ευχαριστήσω τα μέλη της Εξεταστικής Επιτροπής για τον χρόνο και την προσοχή που αφιέρωσαν στη μελέτη της εργασίας μου, καθώς και το ακαδημαϊκό και διοικητικό προσωπικό του Πανεπιστημίου Νεάπολις Πάφος για την υποστήριξη που παρείχαν κατά τη διάρκεια των σπουδών μου.

Τέλος, θερμές ευχαριστίες απευθύνω στην οικογένειά μου και στον σύντροφό μου, οι οποίοι στάθηκαν δίπλα μου σε όλη αυτή τη διαδρομή, προσφέροντάς μου ηθική στήριξη, υπομονή και ενθάρρυνση. Η συμβολή τους υπήρξε καθοριστική για την ολοκλήρωση της παρούσας διπλωματικής εργασίας.

Πίνακας Περιεχομένων

Πρόλογος και Ευχαριστίες	5
Περίληψη	8
Abstract	9
ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ.....	11
1.1 Εισαγωγή.....	11
1.2 Στόχοι.....	11
ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ.....	12
2.1 Ορισμοί και Βασικές Έννοιες	12
2.1.1 Τεχνητή Νοημοσύνη	13
2.1.2 Μηχανική Μάθηση	13
2.1.3 Χρηματοοικονομική Τεχνολογία	14
2.1.4 Εννοιολογική Σύνθεση.....	14
2.2 Ιστορική Εξέλιξη της ΤΝ στον Χρηματοπιστωτικό Τομέα.....	15
2.3 Κύριες Εφαρμογές της ΤΝ στα Χρηματοοικονομικά.....	17
2.3.1 Πιστωτικός Κίνδυνος και Πιθανότητα Αθέτησης.....	17
2.3.2 Τεχνητή Νοημοσύνη στο Trading και στις Επενδυτικές Αποφάσεις	18
2.3.3 Αλγοριθμική Διαχείριση Χρηματοοικονομικού Κινδύνου	19
2.4 Ηθικά και Ρυθμιστικά Ζητήματα στην ΤΝ και τη FinTech	19
2.4.1 Ζητήματα Ισότητας, Αλγοριθμικής Δικαιοσύνης και Μεροληψίας.....	20
2.4.2 Ρυθμιστικό Πλαίσιο και Εποπτεία.....	20
2.4.3 Ηθικά Ζητήματα ΑΙ στη Φορολογία & Συμμόρφωση	20
2.5 Συνολική Αξιολόγηση και Ερευνητικά Κενά.....	21
2.5.1 Αξιολόγηση της Βιβλιογραφίας και Ανάγκη Ολιστικής Προσέγγισης.....	21
2.5.2 Περιορισμοί και Κατακερματισμός της Υπάρχουσας Βιβλιογραφίας.....	22
2.5.3 Ερευνητικά Κενά στις Χρηματοοικονομικές Εργασίες	22
2.5.4 Θέση και Συμβολή της Παρούσας Διπλωματικής Εργασίας	23
ΚΕΦΑΛΑΙΟ 3: ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΚΑΙ ΨΗΦΙΟΠΟΙΗΣΗ.....	23
3.1 Τράπεζες στο Σύγχρονο Χρηματοοικονομικό Σύστημα	24
3.2 Ο Ρόλος των Χρηματοοικονομικών Αγορών στη Σύγχρονη Οικονομία	25
3.3 Το Οικοσύστημα FinTech στο Σύγχρονο Χρηματοοικονομικό Σύστημα.....	26
3.4 Ο Ψηφιακός Μετασχηματισμός του Χρηματοοικονομικού Τομέα.....	27
ΚΕΦΑΛΑΙΟ 4: ΜΕΘΟΔΟΛΟΓΙΑ.....	28
4.1 Σχεδιασμός της Έρευνας.....	28
4.2 Δεδομένα και Πηγή Δεδομένων	29
4.3 Προετοιμασία Δεδομένων	30
4.3.1 Εξερευνητική Ανάλυση Δεδομένων	30
4.3.2 Προεπεξεργασία Δεδομένων και Μετασχηματισμοί	35
4.4 Μοντέλα Ταξινόμησης.....	35
4.4.1 Λογιστική Παλινδρόμηση (Logistic Regression).....	36
4.4.2 Δέντρο Απόφασης (Decision Tree)	36
4.4.3 Συγκριτική Θεώρηση των δύο Μοντέλων.....	37
4.5 Αξιολόγηση Μοντέλων.....	37
ΚΕΦΑΛΑΙΟ 5: ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ.....	39

5.1 Λογιστική Παλινδρόμηση	39
5.1.1 Περιγραφή και Εκπαίδευση Λογιστικής Παλινδρόμησης.....	39
5.1.2 Ποσοτικά Αποτελέσματα Λογιστικής Παλινδρόμησης	39
5.1.3 Ερμηνεία και Σχολιασμός Αποτελεσμάτων.....	41
5.2 Εκπαίδευση και Αξιολόγηση Δέντρου Απόφασης.....	42
5.2.1 Περιγραφή και Εκπαίδευση Δέντρου Απόφασης	42
5.2.2 Ποσοτικά Αποτελέσματα Δέντρου Απόφασης.....	43
5.2.3 Ερμηνεία και Σχολιασμός Αποτελεσμάτων	44
5.3 Σύγκριση Μοντέλων	45
5.3.1 Ποσοτική Σύγκριση.....	45
5.3.2 Ποιοτική Σύγκριση	46
5.3.3 Ανάλυση Σφαλμάτων: Κόστος False Positives και False Negatives.....	47
5.3.4 Περιορισμοί και Πηγές Μεροληψίας	48
5.4 Καταλληλότητα των Μοντέλων για την Πρόβλεψη Αθέτησης.....	49
5.5 Προτάσεις για Βελτιώσεις και Μελλοντική Έρευνα	49
ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ.....	50
6.1 Ανακεφαλαίωση	50
6.2 Μελλοντική Έρευνα	51
Βιβλιογραφία	53
Παράρτημα Α: Πλήρης κώδικας σε Python.....	56

Πίνακας Εικόνων

Εικόνα 1 Κατανομή μεταβλητής-στόχου (Default).....	33
Εικόνα 2 Πίνακας συσχέτισης Pearson σε μορφή θερμικού χάρτη (heatmap).....	34
Εικόνα 3 Σύγκριση καμπυλών ROC.....	38
Εικόνα 4 Πίνακας σύγχυσης Λογιστικής Παλινδρόμησης.....	40
Εικόνα 5 Πίνακας σύγχυσης Δέντρου Απόφασης	44

Πίνακας Πινάκων

Πίνακας 1 Κατανομή κατηγορικών μεταβλητών	32
Πίνακας 2 Μεταβλητή-στόχος Default.....	32
Πίνακας 3 Μετρικές απόδοσης Λογιστικής Παλινδρόμησης ανά κλάση.....	41
Πίνακας 4 Συνολικές μετρικές απόδοσης Λογιστικής Παλινδρόμησης.....	41
Πίνακας 5 Ποσοτικά Αποτελέσματα Δέντρου Απόφασης	43

Τίτλος Διπλωματικής Εργασίας: Η Χρήση της Τεχνητής Νοημοσύνης σε Χρηματοοικονομικές Εργασίες

Η παρούσα Διπλωματική Εργασία εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση εξ αποστάσεως μεταπτυχιακού τίτλου στο Πανεπιστήμιο Νεάπολις και εγκρίθηκε στις από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή:

Πρώτος επιβλέπων (Πανεπιστήμιο Νεάπολις Πάφος): Δρ. Κωνσταντίνος Παναγιωτάκης, Καθηγητής

Μέλος Εξεταστικής Επιτροπής: Χρήστος Λεμονάκης, Αναπληρωτής Καθηγητής

Μέλος Εξεταστικής Επιτροπής: Γεώργιος Μαστοράκης, Αναπληρωτής Καθηγητής

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει τη χρήση τεχνικών Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης σε χρηματοοικονομικές εργασίες, με έμφαση στην πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας. Στόχος της εργασίας είναι η αξιολόγηση της

αποτελεσματικότητας διαφορετικών αλγοριθμικών μοντέλων στην υποστήριξη αποφάσεων διαχείρισης πιστωτικού κινδύνου, καθώς και η ανάδειξη της σημασίας της σωστής επιλογής μετρικών αξιολόγησης σε προβλήματα ανισορροπίας κλάσεων.

Για την εμπειρική ανάλυση χρησιμοποιήθηκε πραγματικό σύνολο δεδομένων πιστωτικών καρτών, το οποίο υποβλήθηκε σε διαδικασίες προεπεξεργασίας και εξερευνητικής ανάλυσης. Το πρόβλημα διατυπώθηκε ως πρόβλημα επιτηρούμενης μάθησης δυαδικής ταξινόμησης και εφαρμόστηκαν δύο διαφορετικά μοντέλα Μηχανικής Μάθησης: η Λογιστική Παλινδρόμηση και το Δέντρο Απόφασης. Η αξιολόγηση των μοντέλων πραγματοποιήθηκε με τη χρήση πολλαπλών μετρικών, όπως accuracy, precision, recall, F1-score και ROC-AUC, ώστε να εξαχθούν αξιόπιστα και συγκρίσιμα αποτελέσματα.

Τα αποτελέσματα έδειξαν ότι η Λογιστική Παλινδρόμηση παρουσίασε υψηλή συνολική ακρίβεια, ωστόσο υστέρησε σημαντικά στην ανίχνευση της αθέτησης πληρωμής, εμφανίζοντας χαμηλό recall και αυξημένο αριθμό false negatives. Αντίθετα, το Δέντρο Απόφασης παρουσίασε βελτιωμένη απόδοση στην αναγνώριση επισφαλών πελατών, επιτυγχάνοντας υψηλότερες τιμές recall και F1-score, γεγονός που το καθιστά καταλληλότερο σε περιβάλλοντα όπου το κόστος της λανθασμένης ταξινόμησης είναι υψηλό.

Συμπερασματικά, η εργασία καταδεικνύει ότι οι τεχνικές Μηχανικής Μάθησης μπορούν να αποτελέσουν αποτελεσματικά εργαλεία υποστήριξης αποφάσεων στον χρηματοοικονομικό τομέα, υπό την προϋπόθεση της ορθής μεθοδολογικής εφαρμογής και της προσαρμογής τους στις επιχειρησιακές ανάγκες. Παράλληλα, αναδεικνύεται η ανάγκη για περαιτέρω έρευνα σε πιο προηγμένα και ερμηνεύσιμα μοντέλα Τεχνητής Νοημοσύνης, με στόχο τη βιώσιμη και υπεύθυνη ενσωμάτωσή τους στις χρηματοοικονομικές διαδικασίες.

Λέξεις-Κλειδιά: Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Πιστωτικός Κίνδυνος, Αθέτηση Πληρωμής, Λογιστική Παλινδρόμηση, Δέντρο Απόφασης

Abstract

This master's thesis examines the use of Artificial Intelligence and Machine Learning techniques in financial applications, with an emphasis on predicting credit card default. The aim of the study is to evaluate the effectiveness of different algorithmic models in supporting

credit risk management decisions, as well as to highlight the importance of selecting appropriate evaluation metrics in classification problems with class imbalance.

For the empirical analysis, a real-world credit card dataset was used, which underwent preprocessing and exploratory data analysis. The problem was formulated as a supervised binary classification task, and two different Machine Learning models were applied: Logistic Regression and Decision Tree. Model evaluation was conducted using multiple performance metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, in order to obtain reliable and comparable results.

The results showed that Logistic Regression achieved high overall accuracy but performed poorly in detecting credit default cases, exhibiting low recall and a high number of false negatives. In contrast, the Decision Tree model demonstrated improved performance in identifying high-risk customers, achieving higher recall and F1-score values, making it more suitable in environments where the cost of misclassification is high.

In conclusion, the study demonstrates that Machine Learning techniques can serve as effective decision-support tools in the financial sector, provided that they are applied using appropriate methodological practices and aligned with business requirements. At the same time, the need for further research on more advanced and interpretable Artificial Intelligence models is highlighted, aiming at their sustainable and responsible integration into financial decision-making processes.

Keywords: Artificial Intelligence, Machine Learning, Credit Risk, Default Prediction, Logistic Regression, Decision Tree

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 Εισαγωγή

Η ραγδαία ανάπτυξη της Τεχνητής Νοημοσύνης και της Μηχανικής Μάθησης τα τελευταία χρόνια έχει επιφέρει σημαντικές αλλαγές στον τρόπο λειτουργίας των χρηματοοικονομικών οργανισμών. Η αυξανόμενη διαθεσιμότητα δεδομένων, σε συνδυασμό με την πρόοδο των υπολογιστικών συστημάτων, έχει καταστήσει δυνατή την αξιοποίηση προηγμένων αλγοριθμικών τεχνικών για την υποστήριξη κρίσιμων οικονομικών αποφάσεων. Ιδιαίτερα στον τραπεζικό τομέα, η ανάγκη για αποτελεσματική διαχείριση κινδύνου και βελτιστοποίηση των διαδικασιών αξιολόγησης πελατών καθιστά τις τεχνολογίες Τεχνητής Νοημοσύνης ιδιαίτερα επίκαιρες και σημαντικές.

Ένας από τους βασικότερους τομείς εφαρμογής της Μηχανικής Μάθησης στις χρηματοοικονομικές εργασίες είναι η διαχείριση πιστωτικού κινδύνου. Η πρόβλεψη της αθέτησης πληρωμής αποτελεί κρίσιμο παράγοντα για τη βιωσιμότητα των χρηματοπιστωτικών οργανισμών, καθώς επηρεάζει άμεσα την κερδοφορία και τη σταθερότητα του χρηματοοικονομικού συστήματος. Οι παραδοσιακές στατιστικές μέθοδοι αξιολόγησης πιστοληπτικής ικανότητας, αν και ευρέως χρησιμοποιούμενες, συχνά αδυνατούν να αποτυπώσουν τις σύνθετες και μη γραμμικές σχέσεις που χαρακτηρίζουν τη χρηματοοικονομική συμπεριφορά των πελατών.

Στο πλαίσιο αυτό, οι τεχνικές Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης προσφέρουν νέες δυνατότητες ανάλυσης και πρόβλεψης, επιτρέποντας την αξιοποίηση μεγάλων και ετερογενών συνόλων δεδομένων. Μέσω της εφαρμογής αλγορίθμων ταξινόμησης, είναι εφικτή η ανάπτυξη μοντέλων που μπορούν να εντοπίζουν έγκαιρα πελάτες με αυξημένη πιθανότητα αθέτησης, συμβάλλοντας στη βελτίωση των διαδικασιών λήψης αποφάσεων και στη μείωση του χρηματοοικονομικού κινδύνου.

Η παρούσα διπλωματική εργασία εστιάζει στη μελέτη και σύγκριση βασικών μοντέλων Μηχανικής Μάθησης για την πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας. Μέσα από θεωρητική ανάλυση και εμπειρική εφαρμογή σε πραγματικά δεδομένα, επιχειρείται η αξιολόγηση της απόδοσης διαφορετικών αλγοριθμικών προσεγγίσεων, καθώς και η ανάδειξη της σημασίας της επιλογής κατάλληλων μετρικών αξιολόγησης σε προβλήματα ανισορροπίας κλάσεων. Στόχος της εργασίας είναι να συμβάλει στην κατανόηση του ρόλου της Τεχνητής Νοημοσύνης στις σύγχρονες χρηματοοικονομικές εργασίες και να αναδείξει τις δυνατότητες αλλά και τους περιορισμούς της εφαρμογής της στον τομέα της διαχείρισης πιστωτικού κινδύνου.

1.2 Στόχοι

Ο βασικός στόχος της παρούσας διπλωματικής εργασίας είναι η διερεύνηση και αξιολόγηση της αποτελεσματικότητας τεχνικών Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης στη διαχείριση πιστωτικού κινδύνου, με έμφαση στην πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας. Η εργασία επιδιώκει να αναδείξει τον ρόλο των αλγοριθμικών μοντέλων ως εργαλείων υποστήριξης χρηματοοικονομικών αποφάσεων, καθώς και να εξετάσει τον βαθμό στον οποίο μπορούν να συμβάλουν στη μείωση του κινδύνου για τους χρηματοπιστωτικούς οργανισμούς.

Στο πλαίσιο αυτό, η εργασία στοχεύει στην παρουσίαση του θεωρητικού υποβάθρου της Τεχνητής Νοημοσύνης και της Μηχανικής Μάθησης και στη διερεύνηση των εφαρμογών τους στις χρηματοοικονομικές εργασίες. Παράλληλα, εξετάζεται το πρόβλημα της πρόβλεψης αθέτησης πληρωμής ως πρόβλημα επιτηρούμενης μάθησης δυαδικής ταξινόμησης, με τη χρήση πραγματικών χρηματοοικονομικών δεδομένων. Ιδιαίτερη έμφαση δίνεται στη διαδικασία προετοιμασίας και ανάλυσης των δεδομένων, καθώς και στην αντιμετώπιση της ανισορροπίας κλάσεων, η οποία αποτελεί κρίσιμο ζήτημα σε προβλήματα πιστωτικού κινδύνου.

Επιπλέον, η εργασία αποσκοπεί στην ανάπτυξη, εκπαίδευση και αξιολόγηση διαφορετικών μοντέλων Μηχανικής Μάθησης, συγκεκριμένα της Λογιστικής Παλινδρόμησης και του Δέντρου Απόφασης, με στόχο τη συγκριτική ανάλυση της απόδοσής τους. Η αξιολόγηση των μοντέλων πραγματοποιείται με τη χρήση πολλαπλών μετρικών, ώστε να εξαχθούν αξιόπιστα συμπεράσματα σχετικά με την καταλληλότητά τους σε πραγματικές συνθήκες εφαρμογής. Τέλος, μέσα από την ανάλυση των αποτελεσμάτων, επιδιώκεται η εξαγωγή συμπερασμάτων σχετικά με τα πλεονεκτήματα, τους περιορισμούς και τις προοπτικές αξιοποίησης των τεχνικών Τεχνητής Νοημοσύνης στη διαχείριση πιστωτικού κινδύνου.

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

2.1 Ορισμοί και Βασικές Έννοιες

Η Τεχνητή Νοημοσύνη αναφέρεται σε υπολογιστικά συστήματα που εκτελούν λειτουργίες συνδεδεμένες με την ανθρώπινη νοημοσύνη, όπως η μάθηση, η συλλογιστική και η λήψη αποφάσεων. Στον χρηματοοικονομικό τομέα, η σημασία της είναι αυξανόμενη, καθώς υποστηρίζει την αυτοματοποίηση διαδικασιών και την αξιοποίηση δεδομένων μεγάλης κλίμακας, ενισχύοντας την αποδοτικότητα και τη διαχείριση κινδύνου (Μάργαρης, 2025). Σύμφωνα με την Ευρωπαϊκή Επιτροπή, η ΤΝ περιλαμβάνει συστήματα που αναλύουν το περιβάλλον τους και ενεργούν με κάποιο βαθμό αυτονομίας προς την επίτευξη στόχων (Μάργαρης, 2025). Παρά την ευρεία χρήση του όρου, δεν υπάρχει ενιαίος ορισμός της Τεχνητής Νοημοσύνης. Ο John McCarthy συνέδεσε την ΤΝ με τη δημιουργία ευφύων μηχανών, ενώ οι Stuart Russell και Peter Norvig την ορίζουν ως τη μελέτη συστημάτων που αντιλαμβάνονται το περιβάλλον τους και δρουν ορθολογικά (McCarthy, 1956; Russell & Norvig, 2021).

Κύριο χαρακτηριστικό που διαφοροποιεί τα συστήματα ΤΝ από το παραδοσιακό λογισμικό είναι η ικανότητα μάθησης από δεδομένα χωρίς ρητό επαναπρογραμματισμό, γεγονός που τα καθιστά ιδιαίτερα κατάλληλα για σύνθετα περιβάλλοντα όπως ο χρηματοοικονομικός τομέας. Στο πλαίσιο του Banking 4.0, η ΤΝ αποτελεί βασικό άξονα της ψηφιοποίησης και της ανάλυσης δεδομένων σε πραγματικό χρόνο (Μπαλοδήμος, 2024). Σε εννοιολογικό επίπεδο, διακρίνεται κυρίως η στενή ΤΝ (Narrow AI), η οποία αφορά συγκεκριμένες εφαρμογές και αποτελεί τη βάση των σημερινών χρηματοοικονομικών συστημάτων, ενώ η γενική ΤΝ παραμένει αντικείμενο έρευνας. Στην πράξη, οι περισσότερες εφαρμογές βασίζονται σε συστήματα που μαθαίνουν από δεδομένα, δηλαδή σε τεχνικές Μηχανικής Μάθησης και Βαθιάς Μάθησης, οι οποίες χρησιμοποιούνται ευρέως σε πιστωτική αξιολόγηση, ανίχνευση απάτης και διαχείριση κινδύνου.

2.1.1 Τεχνητή Νοημοσύνη

Η Τεχνητή Νοημοσύνη δεν αποτελεί μια ενιαία και ομοιογενή τεχνολογία, αλλά ένα σύνολο επιπέδων και κατηγοριών που διαφοροποιούνται ως προς τις δυνατότητες, τον βαθμό αυτονομίας και την πολυπλοκότητα των εφαρμογών τους. Η κατηγοριοποίηση αυτή συμβάλλει στην καλύτερη κατανόηση της λειτουργίας των συστημάτων ΤΝ και της σχέσης τους με επιμέρους τεχνολογίες, όπως η Μηχανική Μάθηση και η Βαθιά Μάθηση. Σύμφωνα με τη διεθνή βιβλιογραφία, η ΤΝ διακρίνεται συχνά με βάση το επίπεδο «νοημοσύνης» που επιδεικνύουν τα συστήματα. Η λεγόμενη στενή ή αδύναμη ΤΝ (Narrow AI) αφορά συστήματα σχεδιασμένα για την εκτέλεση συγκεκριμένων και καλά ορισμένων εργασιών, όπως η ανάλυση δεδομένων, η πρόβλεψη και η ταξινόμηση πληροφοριών, και αποτελεί τη βάση των περισσότερων σύγχρονων εφαρμογών στον χρηματοοικονομικό τομέα. Αντίθετα, η γενική ΤΝ (General AI), η οποία θα μπορούσε να εκτελεί ευρύ φάσμα γνωστικών λειτουργιών σε επίπεδο συγκρίσιμο με αυτό του ανθρώπου, παραμένει αντικείμενο θεωρητικής έρευνας και δεν έχει ακόμη υλοποιηθεί στην πράξη.

Παράλληλα, μια εναλλακτική προσέγγιση κατηγοριοποίησης αφορά τη λειτουργική αρχιτεκτονική των συστημάτων ΤΝ. Στο πλαίσιο αυτό, διακρίνονται τα συστήματα βασισμένα σε κανόνες (rule-based systems), τα οποία λειτουργούν με προκαθορισμένη λογική «αν-τότε», και τα συστήματα που βασίζονται στη μάθηση από δεδομένα. Η δεύτερη κατηγορία περιλαμβάνει τις τεχνικές Μηχανικής Μάθησης, οι οποίες επιτρέπουν τη δυναμική προσαρμογή και τη βελτίωση της απόδοσης των συστημάτων με βάση την εμπειρία. Η κατανόηση των επιπέδων και των κατηγοριών της ΤΝ είναι ιδιαίτερα σημαντική για τον χρηματοοικονομικό τομέα, καθώς κάθε προσέγγιση συνεπάγεται διαφορετικά επίπεδα αυτοματοποίησης, διαφάνειας και ελέγχου. Στο πλαίσιο της παρούσας εργασίας, η ανάλυση εστιάζει κυρίως σε συστήματα στενής ΤΝ και σε τεχνολογίες που βασίζονται στη Μηχανική Μάθηση, οι οποίες αποτελούν τον πυρήνα των σύγχρονων εφαρμογών ΤΝ στις χρηματοοικονομικές εργασίες.

2.1.2 Μηχανική Μάθηση

Η Μηχανική Μάθηση αποτελεί βασικό υποπεδίο της Τεχνητής Νοημοσύνης και αναφέρεται σε αλγοριθμικά συστήματα που έχουν τη δυνατότητα να μαθαίνουν από δεδομένα και να βελτιώνουν την απόδοσή τους χωρίς να απαιτείται ρητός και προκαθορισμένος προγραμματισμός. Στη διεθνή βιβλιογραφία διακρίνονται τέσσερις κύριες κατηγορίες μηχανικής μάθησης: η επιτηρούμενη μάθηση, η μη επιτηρούμενη μάθηση, η ημι-επιτηρούμενη μάθηση και η ενισχυτική μάθηση, οι οποίες διαφοροποιούνται ως προς τον τρόπο εκπαίδευσης των αλγορίθμων και το είδος της διαθέσιμης ανατροφοδότησης (Μάργαρης, 2025).

Σε αντίθεση με τα παραδοσιακά συστήματα που βασίζονται σε προκαθορισμένους κανόνες, οι μέθοδοι Μηχανικής Μάθησης στηρίζονται στη στατιστική επεξεργασία μεγάλων όγκων δεδομένων και στη σταδιακή προσαρμογή των μοντέλων μέσω της εμπειρίας. Η προσέγγιση αυτή καθιστά τη Μηχανική Μάθηση ιδιαίτερα κατάλληλη για προβλήματα πρόβλεψης και υποστήριξης αποφάσεων σε περιβάλλοντα υψηλής πολυπλοκότητας και αβεβαιότητας, όπως αυτά που χαρακτηρίζουν τον χρηματοπιστωτικό τομέα (Μάργαρης, 2025; Μπαλοδήμος, 2024).

Η Μηχανική Μάθηση έχει εξελιχθεί σε κρίσιμο εργαλείο για τα χρηματοπιστωτικά ιδρύματα, καθώς επιτρέπει την ανάπτυξη ακριβέστερων υποδειγμάτων πιστοληπτικής αξιολόγησης, την ανάλυση κινδύνου σε πραγματικό χρόνο, την αυτόματη ανίχνευση απάτης, καθώς και τη

διαμόρφωση προσωποποιημένων επενδυτικών προτάσεων μέσω συστημάτων robo-advisors. Οι εφαρμογές αυτές αξιοποιούν την ικανότητα των αλγορίθμων να εντοπίζουν πολύπλοκα πρότυπα σε μεγάλα σύνολα δεδομένων, τα οποία είναι δύσκολο ή αδύνατο να ανιχνευθούν με παραδοσιακές αναλυτικές μεθόδους (Μπαλοδήμος, 2024).

Η στενή σχέση μεταξύ Μηχανικής Μάθησης και Big Data αποτελεί βασικό παράγοντα της ευρείας υιοθέτησης των τεχνολογιών αυτών στον χρηματοοικονομικό τομέα, καθώς η αυξημένη διαθεσιμότητα δεδομένων συμβάλλει στη σημαντική βελτίωση της ακρίβειας, της αξιοπιστίας και της προσαρμοστικότητας των προγνωστικών μοντέλων (Μπαλοδήμος, 2024).

2.1.3 Χρηματοοικονομική Τεχνολογία

Ο όρος Χρηματοοικονομική Τεχνολογία αναφέρεται στη διασταύρωση της τεχνολογίας με τις χρηματοοικονομικές υπηρεσίες και περιλαμβάνει καινοτομίες που βελτιώνουν, αυτοματοποιούν ή αναδιαμορφώνουν διαδικασίες και προϊόντα στον χρηματοοικονομικό κλάδο. Στη διεθνή βιβλιογραφία, το FinTech ορίζεται ως το σύνολο των τεχνολογικά υποστηριζόμενων λύσεων που καλύπτουν κάθε τομέα των χρηματοοικονομικών υπηρεσιών, από τις πληρωμές και τη χρηματοδότηση έως τις επενδύσεις και τη διαχείριση κινδύνου (Μελισσάρη, 2025).

Σύγχρονες ελληνικές μελέτες επισημαίνουν ότι οι FinTech επιχειρήσεις και οι ψηφιακές τράπεζες (neobanks) έχουν συμβάλει καθοριστικά στον μετασχηματισμό του χρηματοπιστωτικού τοπίου, προσφέροντας ταχύτερες, ασφαλέστερες και περισσότερο προσβάσιμες χρηματοοικονομικές υπηρεσίες. Ενδεικτικά παραδείγματα αποτελούν οι ψηφιακές πληρωμές, οι αυτοματοποιημένες επενδυτικές υπηρεσίες και τα διαδικτυακά εργαλεία χρηματοδότησης, τα οποία ενισχύουν τη χρηματοοικονομική συμπερίληψη και βελτιώνουν την εμπειρία των χρηστών (Μελισσάρη, 2025).

Οι βασικοί κλάδοι εφαρμογής της FinTech περιλαμβάνουν τις ψηφιακές πληρωμές, τις επενδυτικές πλατφόρμες και τα συστήματα robo-advisors, το crowdfunding και το peer-to-peer lending, τη RegTech (κανονιστική τεχνολογία), την InsurTech, καθώς και το Open Banking. Οι τομείς αυτοί βασίζονται σε προηγμένες ψηφιακές υποδομές και στη συστηματική αξιοποίηση δεδομένων, γεγονός που δημιουργεί πρόσφορο έδαφος για την ενσωμάτωση τεχνολογιών Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης (Μελισσάρη, 2025).

Η ανάπτυξη των ψηφιακών τεχνολογιών και των FinTech λύσεων έχει εντείνει τον ανταγωνισμό προς τις παραδοσιακές τράπεζες, οι οποίες καλούνται να προσαρμόσουν τα επιχειρησιακά τους μοντέλα προκειμένου να διατηρήσουν την ανταγωνιστικότητά τους. Στο πλαίσιο αυτό, η FinTech μπορεί να λειτουργήσει είτε ως παράγοντας ενίσχυσης των τραπεζών μέσω συνεργασιών και τεχνολογικών συμπράξεων είτε ως δυνητική απειλή για τον παραδοσιακό τους ρόλο, εφόσον δεν προσαρμοστούν έγκαιρα στις εξελίξεις του ψηφιακού χρηματοπιστωτικού περιβάλλοντος (Σκαρλατάκης, 2022).

2.1.4 Εννοιολογική Σύνθεση

Η Τεχνητή Νοημοσύνη, η Μηχανική Μάθηση και η Χρηματοοικονομική Τεχνολογία δεν αποτελούν ανεξάρτητα και απομονωμένα πεδία, αλλά αλληλένδετους και συμπληρωματικούς άξονες του σύγχρονου ψηφιακού μετασχηματισμού του χρηματοοικονομικού συστήματος. Η συνδυαστική αξιοποίησή τους έχει οδηγήσει σε βαθιές αλλαγές στον τρόπο με τον οποίο

σχεδιάζονται, προσφέρονται και διαχειρίζονται οι χρηματοοικονομικές υπηρεσίες (Σκαρλατάκης, 2022; Μελισσάρη, 2025; Μπαλοδήμος, 2024).

Το FinTech λειτουργεί ως το εφαρμοσμένο οικοσύστημα εντός του οποίου ενσωματώνονται και αξιοποιούνται οι τεχνολογίες αυτές, παρέχοντας τις ψηφιακές πλατφόρμες και τις υποδομές για την ανάπτυξη και διάχυση καινοτόμων χρηματοοικονομικών προϊόντων. Η Μηχανική Μάθηση, από την πλευρά της, παρέχει τα αλγοριθμικά εργαλεία που επιτρέπουν την ανάλυση μεγάλων όγκων δεδομένων, την εξαγωγή προτύπων και τη δημιουργία προγνωστικών μοντέλων, τα οποία αποτελούν τη βάση για πολλές εφαρμογές στον χρηματοπιστωτικό τομέα. Η Τεχνητή Νοημοσύνη συνθέτει τα παραπάνω, επιτρέποντας την αυτοματοποίηση σύνθετων διαδικασιών και τη λήψη αποφάσεων με ελάχιστη ή μηδενική ανθρώπινη παρέμβαση (Μπαλοδήμος, 2024).

Η αλληλεπίδραση των τριών τεχνολογιών οδηγεί στη δημιουργία νέων επιχειρησιακών μοντέλων, όπως οι ψηφιακές τράπεζες (neobanks), οι αυτοματοποιημένες επενδυτικές υπηρεσίες μέσω robo-advisors και οι προηγμένες διαδικασίες αξιολόγησης και διαχείρισης κινδύνου. Παράλληλα, ενισχύεται η δυνατότητα παροχής προσωποποιημένων υπηρεσιών, βελτιώνεται η αποδοτικότητα των οργανισμών και μειώνεται το λειτουργικό κόστος (Μελισσάρη, 2025).

Σύμφωνα με τη σύγχρονη βιβλιογραφία, η υιοθέτηση και η ολοκληρωμένη ενσωμάτωση της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης και της FinTech αποτελεί βασική προϋπόθεση για τα χρηματοπιστωτικά ιδρύματα που επιθυμούν να παραμείνουν ανταγωνιστικά στο περιβάλλον του Banking 4.0, το οποίο χαρακτηρίζεται από εκτεταμένη ψηφιοποίηση, αυτοματοποίηση και αξιοποίηση δεδομένων σε πραγματικό χρόνο (Μπαλοδήμος, 2024).

2.2 Ιστορική Εξέλιξη της TN στον Χρηματοπιστωτικό Τομέα

Η ιστορική εξέλιξη της τεχνητής νοημοσύνης (TN) στον χρηματοπιστωτικό τομέα μπορεί να ιδωθεί ως μια αλληλουχία «κυμάτων» τεχνολογικής υιοθέτησης, όπου κάθε κύμα μετασχηματίζει σταδιακά τον τρόπο με τον οποίο οργανώνονται, εκτελούνται και ελέγχονται οι χρηματοοικονομικές λειτουργίες. Με βάση τη διπλωματική εργασία του Tarenzi, τη μελέτη της Golić για τον ρόλο της TN στον ψηφιακό μετασχηματισμό, καθώς και τη συστηματική βιβλιομετρική ανασκόπηση των Vuković et al., μπορεί να χαρτογραφηθεί μια σχετικά σαφής ιστορική πορεία από τα πρώτα στατιστικά υποδείγματα έως τις σύγχρονες, ολοκληρωμένες ψηφιακές αρχιτεκτονικές (Tarenzi, 2024; Golić, 2019; Vuković et al., 2025).

Οι πρώτες θεωρητικές διατυπώσεις της Τεχνητής Νοημοσύνης από τον Turing και τον McCarthy δεν βρήκαν άμεση εφαρμογή στον χρηματοπιστωτικό τομέα. Ωστόσο, ήδη από τη δεκαετία του 1960, οι χρηματοοικονομικοί οργανισμοί συγκαταλέγονταν μεταξύ των πρωτοπόρων στη χρήση της πληροφορικής για την αυτοματοποίηση βασικών λειτουργιών, όπως ο διακανονισμός πληρωμών, η εκκαθάριση συναλλαγών και η τήρηση λογαριασμών μέσω κεντρικών υπολογιστικών συστημάτων (mainframes). Την ίδια περίοδο αναπτύχθηκαν τα πρώτα συστήματα πιστοληπτικής αξιολόγησης που βασίζονταν σε απλά στατιστικά μοντέλα, όπως η γραμμική παλινδρόμηση και η διακριτική ανάλυση, εισάγοντας τη λογική της τυποποιημένης και δεδομενοκεντρικής λήψης αποφάσεων στον χρηματοπιστωτικό τομέα (Tarenzi, 2024).

Κατά τις δεκαετίες του 1980 και του 1990, η εξέλιξη των υπολογιστικών συστημάτων και η διάδοση των expert systems επέτρεψαν τη μεταφορά των κανόνα-βασισμένων προσεγγίσεων σε εφαρμογές όπως η πιστοδότηση και το underwriting. Παράλληλα, εμφανίζονται οι πρώτες πειραματικές χρήσεις νευρωνικών δικτύων και συστημάτων υποστήριξης επενδυτικών αποφάσεων για την πρόβλεψη τιμών και τάσεων στις αγορές. Συνολικά, η περίοδος αυτή μπορεί να θεωρηθεί ως μεταβατική φάση από την απλή αυτοματοποίηση και τα κλασικά στατιστικά υποδείγματα προς πιο ευέλικτες και «ευφυείς» μεθόδους ανάλυσης, ενώ η σταδιακή ψηφιοποίηση των αγορών και η εμφάνιση των πρώτων ηλεκτρονικών συστημάτων συναλλαγών στα τέλη της δεκαετίας του 1990 έθεσαν τις βάσεις για την επόμενη φάση (Tarenzi, 2024; Golíc, 2019).

Από τις αρχές της δεκαετίας του 2000, η εφαρμογή της TN στον χρηματοπιστωτικό τομέα εισέρχεται σε φάση ταχείας επιτάχυνσης, λόγω της αύξησης της υπολογιστικής ισχύος, της πλήρους ψηφιοποίησης των αγορών και της ραγδαίας διεύρυνσης του όγκου των διαθέσιμων δεδομένων. Η εξάπλωση των ηλεκτρονικών συστημάτων συναλλαγών και η δυνατότητα εκτέλεσης εντολών σε εξαιρετικά μικρούς χρόνους συνέβαλαν καθοριστικά στην ανάπτυξη του algorithmic και του high-frequency trading, μεταβάλλοντας ουσιαστικά τη μικροδομή των αγορών. Την ίδια περίοδο, οι τράπεζες και οι επενδυτικοί οργανισμοί άρχισαν να αξιοποιούν συστηματικότερα τεχνικές μηχανικής μάθησης και data mining για τη διαχείριση κινδύνου, την πρόβλεψη μεταβλητότητας, τη διαμόρφωση τιμών και την ανίχνευση απάτης (Golíc, 2019).

Μετά το 2010, στο πλαίσιο της ώριμης πλέον FinTech εποχής, παρατηρείται σταδιακή μετατόπιση της χρήσης της TN από εφαρμογές που αφορούν κυρίως το back-office προς το front-office των χρηματοπιστωτικών οργανισμών (Tarenzi, 2024). Η διάδοση των smartphones, του cloud computing και των APIs οδήγησε σε εκρηκτική αύξηση του όγκου και της ταχύτητας παραγωγής δεδομένων, επιτρέποντας την ανάπτυξη ψηφιακών πλατφορμών με δυνατότητες ανάλυσης της συμπεριφοράς των χρηστών σε πραγματικό χρόνο. Στο πλαίσιο αυτό, η TN αξιοποιήθηκε εκτεταμένα για την παροχή προσωποποιημένων υπηρεσιών, όπως chatbots και robo-advisors, ενώ αναπτύχθηκαν πιο σύνθετα μοντέλα πιστοληπτικής αξιολόγησης που ενσωματώνουν εναλλακτικά και συμπεριφορικά δεδομένα. Παράλληλα, η ανίχνευση απάτης και η διαχείριση πιστωτικού κινδύνου καθιερώνονται ως κεντρικά πεδία εφαρμογής της TN. Οι ρυθμιστικές εξελίξεις, όπως το open banking και η οδηγία PSD2, επιτάχυναν τη συνεργασία μεταξύ τραπεζών και FinTech επιχειρήσεων και εδραίωσαν την TN ως κρίσιμη υποδομή του σύγχρονου χρηματοπιστωτικού οικοσυστήματος (Vuković et al., 2025).

Η πιο πρόσφατη φάση, μετά το 2020, χαρακτηρίζεται από την ωρίμανση και τη θεσμοθέτηση των τεχνολογιών TN στις χρηματοοικονομικές υπηρεσίες. Σύμφωνα με τους Vuković et al., η TN παύει πλέον να αποτελεί αποσπασματικό εργαλείο και ενσωματώνεται πλήρως σε ενοποιημένες, end-to-end ψηφιακές αρχιτεκτονικές που καλύπτουν ολόκληρο τον κύκλο ζωής του πελάτη (Vuković et al., 2025). Η έρευνα και η πρακτική εφαρμογή επικεντρώνονται στη σύνδεση της TN με το digital finance, το cloud computing, τη διαχείριση κινδύνου, τη χρηματοοικονομική συμπερίληψη και τα κριτήρια ESG, ενώ οι εξελίξεις στη βαθιά και την ενισχυτική μάθηση διευρύνουν τις δυνατότητες αυτοματοποιημένης διαχείρισης χαρτοφυλακίων και αλγοριθμικών στρατηγικών επένδυσης. Παράλληλα, ο συνδυασμός TN με robotic process automation επιτρέπει την περαιτέρω αυτοματοποίηση λειτουργιών back-office.

Ιδιαίτερη σημασία αποκτά πλέον η εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable AI – XAI), ιδίως σε εφαρμογές υψηλού ρίσκου, όπως η πιστοδότηση και η διαχείριση κινδύνου, καθώς οι

αυξημένες ρυθμιστικές απαιτήσεις και η ανάγκη αντιμετώπισης αλγοριθμικών μεροληψιών έχουν μετατοπίσει την έμφαση από απλώς «ακριβή» μοντέλα προς διαφανή και ελέγξιμα συστήματα. Παράλληλα, οι πρόσφατες εξελίξεις στη γενετική Τεχνητή Νοημοσύνη και στα μεγάλα γλωσσικά μοντέλα (LLMs) αρχίζουν να βρίσκουν πρακτική εφαρμογή σε τομείς όπως η ανάλυση οικονομικών αναφορών, η κανονιστική συμμόρφωση και τα εξελιγμένα συστήματα εξυπηρέτησης πελατών. Την ίδια στιγμή, η ευρεία υιοθέτηση δεδομενοκεντρικών και αυτοματοποιημένων μοντέλων ΤΝ εγείρει ζητήματα συστημικού κινδύνου και διακυβέρνησης, καθώς η αυξανόμενη ομοιομορφία αλγοριθμικών πρακτικών ενδέχεται να ενισχύσει τη χρηματοπιστωτική ευπάθεια σε περιόδους κρίσης (Vuković et al., 2025; Challoumis, 2024).

Συνολικά, η ιστορική εξέλιξη της Τεχνητής Νοημοσύνης στον χρηματοπιστωτικό τομέα δεν συνιστά απλώς μια γραμμική τεχνολογική πρόοδο, αλλά μια διαδοχή βαθύτερων μετασχηματισμών στον τρόπο με τον οποίο οργανώνεται, λειτουργεί και ρυθμίζεται το σύγχρονο χρηματοπιστωτικό σύστημα, με την ΤΝ να αποτελεί πλέον δομικό στοιχείο της σύγχρονης χρηματοοικονομικής αρχιτεκτονικής (Tarenzi, 2024; Vuković et al., 2025).

2.3 Κύριες Εφαρμογές της ΤΝ στα Χρηματοοικονομικά

Η τεχνητή νοημοσύνη (ΤΝ) αποτελεί πλέον κρίσιμο μοχλό μετασχηματισμού του χρηματοοικονομικού τομέα, καθώς επιτρέπει την αυτοματοποίηση διαδικασιών και την αξιοποίηση μεγάλων όγκων δεδομένων για ταχύτερη και ακριβέστερη λήψη αποφάσεων (Pattnaik et al., 2024; Maple et al., 2023). Εφαρμογές όπως η πιστωτική αξιολόγηση και πρόβλεψη αθέτησης, η ανίχνευση απάτης, η υποστήριξη επενδυτικών αποφάσεων και η διαχείριση κινδύνου αξιοποιούν τεχνικές μηχανικής μάθησης και αλγοριθμικής ανάλυσης, βελτιώνοντας την αποδοτικότητα και ενισχύοντας τη σταθερότητα των οργανισμών (Γκουνελάς, 2019; Κυδώνα, 2021).

Στην παρούσα ενότητα εξετάζονται αναλυτικότερα τρεις από τις πλέον καθοριστικές εφαρμογές της ΤΝ στα χρηματοοικονομικά: προγνωστικά συστήματα πιστωτικού κινδύνου, εφαρμογές στο trading και στις επενδυτικές αποφάσεις και αλγοριθμικές προσεγγίσεις διαχείρισης χρηματοοικονομικού κινδύνου.

2.3.1 Πιστωτικός Κίνδυνος και Πιθανότητα Αθέτησης

Η πρόβλεψη του πιστωτικού κινδύνου και της πιθανότητας αθέτησης υποχρεώσεων αποτελεί μία από τις πλέον καθιερωμένες και κρίσιμες εφαρμογές της τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα. Τα προγνωστικά συστήματα πιστωτικού κινδύνου χρησιμοποιούνται από τράπεζες και χρηματοπιστωτικούς οργανισμούς με σκοπό την έγκαιρη αξιολόγηση της φερεγγυότητας δανειοληπτών και τη βελτίωση της ποιότητας των χορηγήσεων, συμβάλλοντας στη διατήρηση της χρηματοπιστωτικής σταθερότητας (Παναγιωτοπούλου, 2021).

Τα σύγχρονα μοντέλα βασίζονται σε τεχνικές μηχανικής μάθησης και αξιοποιούν μεγάλα σύνολα δεδομένων που περιλαμβάνουν δημογραφικά, οικονομικά και συμπεριφορικά χαρακτηριστικά πελατών, καθώς και ιστορικά στοιχεία συναλλαγών και εξυπηρέτησης δανείων. Σε αντίθεση με τα παραδοσιακά στατιστικά μοντέλα, οι αλγόριθμοι αυτοί έχουν τη δυνατότητα να εντοπίζουν πολύπλοκες, μη γραμμικές σχέσεις μεταξύ των μεταβλητών και να

προσαρμόζονται δυναμικά σε μεταβαλλόμενες συνθήκες της αγοράς (Παναγιωτοπούλου, 2021; Pattnaik et al., 2024).

Η χρήση προγνωστικών συστημάτων επιτρέπει την ανάπτυξη μηχανισμών έγκαιρης προειδοποίησης (early warning systems), μέσω των οποίων μπορούν να εντοπιστούν δανειολήπτες που εμφανίζουν αυξημένο κίνδυνο αθέτησης, ακόμη και όταν οι υποχρεώσεις τους εξυπηρετούνται κανονικά. Με τον τρόπο αυτό, οι χρηματοπιστωτικοί οργανισμοί μπορούν να εφαρμόσουν προληπτικές στρατηγικές διαχείρισης κινδύνου, όπως αναδιάρθρωση δανείων ή προσαρμογή των πιστωτικών όρων (Παναγιωτοπούλου, 2021; Γκουνελάς, 2019).

Ιδιαίτερη σημασία αποδίδεται τα τελευταία χρόνια και στην ερμηνευσιμότητα των προγνωστικών μοντέλων, καθώς οι αποφάσεις πιστοδότησης υπόκεινται σε αυξημένες κανονιστικές απαιτήσεις διαφάνειας. Η δυνατότητα κατανόησης των παραγόντων που επηρεάζουν την εκτίμηση του πιστωτικού κινδύνου καθίσταται κρίσιμη τόσο για τη συμμόρφωση με το ρυθμιστικό πλαίσιο όσο και για την ενίσχυση της εμπιστοσύνης των πελατών προς τα αυτοματοποιημένα συστήματα λήψης αποφάσεων (Παναγιωτοπούλου, 2021; Maple et al., 2023).

2.3.2 Τεχνητή Νοημοσύνη στο Trading και στις Επενδυτικές Αποφάσεις

Η τεχνητή νοημοσύνη (TN) έχει μετασχηματίσει ριζικά τον τρόπο με τον οποίο λαμβάνονται επενδυτικές αποφάσεις και υλοποιούνται στρατηγικές trading. Οι αλγόριθμοι μηχανικής μάθησης επιτρέπουν την ανάλυση τεράστιου όγκου δεδομένων αγοράς σε πραγματικό χρόνο, εντοπίζοντας πρότυπα και ανωμαλίες που δύσκολα θα μπορούσαν να εντοπιστούν από τον άνθρωπο. Μέσω της αυτοματοποιημένης λήψης αποφάσεων, οι επενδυτικοί οργανισμοί επιτυγχάνουν αυξημένη ταχύτητα συναλλαγών, μείωση σφαλμάτων και προσαρμογή σε μεταβαλλόμενες συνθήκες αγοράς (Pattnaik et al., 2024; Maple et al., 2023). Οι εφαρμογές αυτές εκτείνονται από στρατηγικές εκτέλεσης και market making έως συστήματα υποστήριξης επενδυτικών αποφάσεων σε επίπεδο χαρτοφυλακίου.

Τα αλγοριθμικά συστήματα trading (algorithmic ή high-frequency trading systems) αξιοποιούν νευρωνικά δίκτυα και βαθιά μάθηση (deep learning) για να προβλέψουν τη βραχυπρόθεσμη κατεύθυνση των τιμών μετοχών, ομολόγων ή παραγώγων. Οι προβλέψεις αυτές βασίζονται όχι μόνο σε ιστορικά δεδομένα τιμών, αλλά και σε μη δομημένες πηγές πληροφόρησης, όπως οικονομικές ειδήσεις, ανακοινώσεις εταιρειών και ανάλυση συναισθήματος από τα μέσα κοινωνικής δικτύωσης (Maple et al., 2023).

Η χρήση ρομποτικών συμβούλων (robo-advisors) αποτελεί επίσης μια σημαντική εξέλιξη στη λιανική επενδυτική δραστηριότητα. Μέσω τεχνικών μηχανικής μάθησης, οι robo-advisors προτείνουν εξατομικευμένες επενδυτικές στρατηγικές, λαμβάνοντας υπόψη το προφίλ κινδύνου, τον χρονικό ορίζοντα και τις προτιμήσεις κάθε επενδυτή. Με αυτόν τον τρόπο, η TN συμβάλλει στη δημοκρατικοποίηση της επενδυτικής διαδικασίας, καθιστώντας την πρόσβαση στις αγορές πιο προσιτή και αποτελεσματική (Wallon, 2019).

Παράλληλα, οι αλγόριθμοι ενισχυτικής μάθησης (reinforcement learning) χρησιμοποιούνται όλο και συχνότερα για τη βελτιστοποίηση χαρτοφυλακίων και τη διαχείριση ρίσκου επενδύσεων. Μέσω συνεχούς προσαρμογής στις μεταβολές των τιμών και των μακροοικονομικών δεικτών, τα μοντέλα αυτά «μαθαίνουν» να λαμβάνουν δυναμικές αποφάσεις μεγιστοποίησης απόδοσης και ελαχιστοποίησης κινδύνου (Pattnaik et al., 2024).

Ωστόσο, η ευρεία υιοθέτηση της TN στο trading εγείρει ζητήματα διαφάνειας, ερμηνευσιμότητας και ηθικής ευθύνης, καθώς οι αυτοματοποιημένες αποφάσεις ενδέχεται να οδηγήσουν σε ανεξέλεγκτες διακυμάνσεις αγοράς ή σε μη προβλέψιμες στρατηγικές υψηλού κινδύνου. Η σωστή εποπτεία και η ανάπτυξη κανονιστικών πλαισίων παραμένουν κρίσιμα ζητήματα για τη βιώσιμη ενσωμάτωση των συστημάτων TN στο χρηματοοικονομικό οικοσύστημα (Maple et al., 2023).

2.3.3 Αλγοριθμική Διαχείριση Χρηματοοικονομικού Κινδύνου

Η διαχείριση χρηματοοικονομικού κινδύνου αποτελεί έναν από τους πλέον καίριους τομείς όπου η τεχνητή νοημοσύνη έχει επιφέρει ουσιαστικές βελτιώσεις, καθώς επιτρέπει την παρακολούθηση και αξιολόγηση πολλαπλών πηγών κινδύνου — όπως πιστωτικού, επιτοκιακού, συναλλαγματικού και λειτουργικού — σε πραγματικό χρόνο, συμβάλλοντας στη μείωση της αβεβαιότητας και στη βελτίωση της λήψης αποφάσεων. Με τη χρήση τεχνικών όπως τα μοντέλα πιθανοτικής πρόβλεψης και τα αλγοριθμικά μοντέλα stress testing, οι οργανισμοί μπορούν να προσομοιώνουν ακραία σενάρια και να εκτιμούν τη δυνητική επίδρασή τους στα χαρτοφυλάκια, ενσωματώνοντας εσωτερικά και εξωτερικά δεδομένα και επιτρέποντας τη δυναμική αναθεώρηση των εκτιμήσεων κινδύνου (Γκουνελάς, 2021; Maple et al., 2023).

Παράλληλα, μοντέλα πρόβλεψης μεταβλητότητας, έγκαιρης προειδοποίησης και αξιολόγησης αντισυμβαλλομένου, καθώς και τεχνικές ανίχνευσης ανωμαλιών (anomaly detection), αξιοποιούνται για τον εντοπισμό ασυνήθιστων μοτίβων στις χρηματοοικονομικές συναλλαγές, τα οποία ενδέχεται να υποδηλώνουν αυξημένο ή συστημικό κίνδυνο και παραβιάσεις κανονισμών. Η ενσωμάτωση τέτοιων αλγορίθμων ενισχύει την ικανότητα έγκαιρης παρέμβασης τόσο σε επίπεδο οργανισμών όσο και εποπτικών αρχών (Pattnaik et al., 2023).

Τέλος, στο πλαίσιο της διαχείρισης ρευστότητας και κεφαλαιακών αποθεμάτων, τα συστήματα TN χρησιμοποιούνται για την πρόβλεψη ταμειακών ροών και τη βελτιστοποίηση της χρήσης κεφαλαίων, ενισχύοντας τη σταθερότητα και την ανθεκτικότητα των οργανισμών σε περιόδους αβεβαιότητας (Κυδώνα, 2021). Συνολικά, η συνδυαστική εφαρμογή των τεχνολογιών αυτών διαμορφώνει ένα πιο ολιστικό και προληπτικό πλαίσιο διαχείρισης κινδύνου, στο οποίο η TN λειτουργεί όχι μόνο ως τεχνικό εργαλείο, αλλά και ως βασικός μηχανισμός στρατηγικού σχεδιασμού και οργανωσιακής ωρίμανσης (Maple et al., 2023).

2.4 Ηθικά και Ρυθμιστικά Ζητήματα στην TN και τη FinTech

Η αυξανόμενη ενσωμάτωση της τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα δημιουργεί σημαντικές ευκαιρίες, αλλά ταυτόχρονα αναδεικνύει κρίσιμα ηθικά και ρυθμιστικά ζητήματα. Η χρήση αλγοριθμικών συστημάτων σε αποφάσεις που αφορούν χρηματοδότηση, συμμόρφωση και διαχείριση κινδύνου εγείρει ερωτήματα διαφάνειας, λογοδοσίας, ισότητας και προστασίας προσωπικών δεδομένων (Aldboush & Ferdous, 2023; Agu et al., 2024; Lee, 2020; Dudu et al., 2024).

Στο πλαίσιο αυτό, η παρούσα ενότητα εξετάζει συνοπτικά τρεις κεντρικές διαστάσεις: ζητήματα εμπιστοσύνης, διαφάνειας και αλγοριθμικής μεροληψίας, τις προκλήσεις του ρυθμιστικού πλαισίου και της εποπτείας της Τεχνητής Νοημοσύνης στο χρηματοοικονομικό σύστημα και ειδικά ηθικά ζητήματα που ανακύπτουν σε τομείς όπως η φορολογία και η κανονιστική συμμόρφωση.

2.4.1 Ζητήματα Ισότητας, Αλγοριθμικής Δικαιοσύνης και Μεροληψίας

Η χρήση δεδομένων πραγματικού κόσμου στην εκπαίδευση μοντέλων ΑΙ μπορεί να οδηγήσει σε ενσωμάτωση συστηματικών μεροληψιών, οι οποίες με τη σειρά τους παράγουν άδικες ή διακριτικές εκβάσεις. Στον χώρο της χρηματοοικονομικής τεχνολογίας, η ενσωμάτωση μεροληπτικών δεδομένων στα μοντέλα τεχνητής νοημοσύνης μπορεί να οδηγήσει σε άνιση πρόσβαση στη χρηματοδότηση, σε ακατάλληλη ή στρεβλή αξιολόγηση της πιστοληπτικής ικανότητας και σε διακριτικές πρακτικές εις βάρος μειονοτικών ή ευάλωτων κοινωνικών ομάδων (Agu et al., 2024).

Η διασφάλιση της fairness απαιτεί την υιοθέτηση τόσο τεχνικών προσεγγίσεων (όπως bias mitigation, αναπλαισίωση χαρακτηριστικών και μοντέλα explainable AI), όσο και κατάλληλων θεσμικών παρεμβάσεων, καθώς η ηθική διάσταση της αλγοριθμικής δικαιοσύνης είναι κρίσιμη: μοντέλα που λειτουργούν χωρίς έλεγχο μπορούν να αναπαράγουν ή να ενισχύσουν υπάρχουσες κοινωνικές ανισότητες (Agu et al., 2024).

2.4.2 Ρυθμιστικό Πλαίσιο και Εποπτεία

Οι ρυθμιστικές αρχές αναγνωρίζουν ότι η αυξανόμενη χρήση τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα δημιουργεί νέες μορφές κινδύνου, οι οποίες δεν καλύπτονται επαρκώς από τα παραδοσιακά κανονιστικά πλαίσια. Όπως επισημαίνει ο Lee (2020), η ταχύτητα της τεχνολογικής εξέλιξης συχνά υπερβαίνει τη δυνατότητα των εποπτικών μηχανισμών να προσαρμοστούν εγκαίρως, δημιουργώντας φαινόμενα κανονιστικής υστέρησης (regulatory lag). Το πρόβλημα εντείνεται από την πολυπλοκότητα και τη σχετική αδιαφάνεια των αλγοριθμικών μοντέλων, καθώς και από την ασυμμετρία πληροφόρησης μεταξύ fintech επιχειρήσεων και εποπτικών αρχών (Lee, 2020).

Παράλληλα, η έλλειψη διεθνούς εναρμόνισης των ρυθμιστικών προσεγγίσεων και η ανάγκη προσαρμογής κανόνων που σχεδιάστηκαν για ανθρώπινους φορείς λήψης αποφάσεων σε αλγοριθμικά συστήματα δημιουργούν πρόσθετες προκλήσεις. Η αποτελεσματική εποπτεία προϋποθέτει νέες μεθοδολογίες ελέγχου και επαρκή θεσμική και τεχνική ενίσχυση των ρυθμιστικών αρχών, ώστε να διασφαλίζεται ταυτόχρονα η χρηματοπιστωτική σταθερότητα και η βιώσιμη καινοτομία στον κλάδο (Lee, 2020).

2.4.3 Ηθικά Ζητήματα ΑΙ στη Φορολογία & Συμμόρφωση

Η εφαρμογή συστημάτων Τεχνητής Νοημοσύνης στη φορολογική συμμόρφωση, έναν ταχέως αναπτυσσόμενο τομέα της χρηματοοικονομικής τεχνολογίας, δημιουργεί νέα και σύνθετα ηθικά ζητήματα. Η αξιοποίηση αλγορίθμων TN από φορολογικές αρχές και ψηφιακές πλατφόρμες συμμόρφωσης μπορεί να ενισχύσει τη διαφάνεια, την αποδοτικότητα και τη μείωση της φοροδιαφυγής. Παράλληλα, όμως, εγείρει σοβαρές ανησυχίες σχετικά με την επεξεργασία μεγάλου όγκου ευαίσθητων οικονομικών και προσωπικών δεδομένων, γεγονός που καθιστά κρίσιμη τη διασφάλιση της ιδιωτικότητας και της προστασίας των φορολογουμένων (Dudu et al., 2024; Aldboush & Ferdous, 2023).

Ένα επιπλέον ηθικό ζήτημα αφορά τον κίνδυνο αυτοματοποιημένων λανθασμένων κατηγοριοποιήσεων και αποφάσεων συμμόρφωσης, οι οποίες ενδέχεται να προκύψουν από σφάλματα στα δεδομένα ή από αλγοριθμική μεροληψία. Σε τέτοιες περιπτώσεις, η αδυναμία κατανόησης ή αμφισβήτησης των αποφάσεων αυτών μπορεί να επηρεάσει αρνητικά τη σχέση εμπιστοσύνης μεταξύ κράτους και πολίτη. Η αυξημένη αυτοματοποίηση στη φορολογική

εποπτεία μετατοπίζει τη λήψη αποφάσεων από ανθρώπινους ελεγκτές σε αλγοριθμικά συστήματα, δημιουργώντας ηθικά διλήμματα σχετικά με το αποδεκτό όριο αντικατάστασης της ανθρώπινης κρίσης (Dudu et al., 2024).

Η σύγχρονη βιβλιογραφία τονίζει την ανάγκη ανάπτυξης ενός ηθικού και διαφανούς πλαισίου χρήσης της Τεχνητής Νοημοσύνης στη φορολογική συμμόρφωση. Οι Dudu et al. (2024) υπογραμμίζουν τη σημασία μηχανισμών εποπτείας, της διατήρησης ανθρώπινης παρέμβασης στη διαδικασία λήψης αποφάσεων (human-in-the-loop) και της αυστηρής συμμόρφωσης με κανονιστικά και δεοντολογικά πρότυπα. Υπό αυτή την οπτική, η ηθική διακυβέρνηση της ΤΝ στη φορολογία δεν αποτελεί μόνο ζήτημα τεχνικής συμμόρφωσης, αλλά βασική προϋπόθεση για τη θεσμική νομιμοποίηση και τη βιώσιμη εφαρμογή των ψηφιακών τεχνολογιών στο σύγχρονο χρηματοοικονομικό κράτος (Dudu et al., 2024).

2.5 Συνολική Αξιολόγηση και Ερευνητικά Κενά

Η παρούσα ενότητα συνθέτει τα βασικά συμπεράσματα της προηγούμενης βιβλιογραφικής ανάλυσης και εστιάζει στα κύρια ερευνητικά κενά που αναδεικνύονται στη μελέτη της τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα. Συγκεκριμένα, αξιολογείται κριτικά η υφιστάμενη βιβλιογραφία, επισημαίνονται οι περιορισμοί και ο κατακερματισμός της, και αναδεικνύεται η ανάγκη για μια πιο ολιστική, διαδικασιοκεντρική και θεσμικά ενσωματωμένη προσέγγιση. Στο πλαίσιο αυτό, η ενότητα καταλήγει στη σαφή τοποθέτηση και τεκμηρίωση της συμβολής της παρούσας διπλωματικής εργασίας.

2.5.1 Αξιολόγηση της Βιβλιογραφίας και Ανάγκη Ολιστικής Προσέγγισης

Συνολικά, τα ηθικά και ρυθμιστικά ζητήματα στην AI/fintech απορρέουν από την ταυτόχρονη ανάγκη προώθησης της καινοτομίας, προστασίας των καταναλωτών, διασφάλισης δικαιοσύνης και ισότητας, ενίσχυσης της διαφάνειας και της λογοδοσίας, καθώς και ανάπτυξης ισχυρού θεσμικού πλαισίου. Η διεθνής βιβλιογραφία αναδεικνύει ότι οι προκλήσεις αυτές είναι πολυπαραγοντικές και απαιτούν προσεκτική ισορροπία μεταξύ τεχνικής προόδου και κοινωνικής υπευθυνότητας. Η ανάγκη για προστασία δεδομένων και ενίσχυση της εμπιστοσύνης συνδέεται άμεσα με ζητήματα διαφάνειας και υπεύθυνης χρήσης αλγορίθμων (Aldboush & Ferdous, 2023), ενώ οι ανησυχίες σχετικά με την αλγοριθμική μεροληψία και την ισότητα πρόσβασης επιβάλλουν συστηματικές παρεμβάσεις για την εξασφάλιση fairness στα χρηματοοικονομικά συστήματα (Agu et al., 2024). Παράλληλα, η διαμόρφωση αποτελεσματικού ρυθμιστικού πλαισίου αποτελεί προϋπόθεση για την ασφαλή και υπεύθυνη ανάπτυξη της AI, δεδομένων των αυξανόμενων τεχνικών, εποπτικών και χρηματοδοτικών απαιτήσεων (Lee, 2020). Στον χώρο της φορολογικής συμμόρφωσης και της ρυθμιστικής τεχνολογίας, αναδύονται επιπλέον ηθικά ζητήματα που σχετίζονται με την αυτοματοποίηση των αποφάσεων και τη διαχείριση ευαίσθητων δεδομένων, υπογραμμίζοντας την ανάγκη για ηθικά θεμελιωμένα μοντέλα και σταθερούς μηχανισμούς διακυβέρνησης (Dudu et al., 2024).

Η βιβλιογραφία συνολικά υποδεικνύει ότι η αποτελεσματική αντιμετώπιση των ηθικών και ρυθμιστικών ζητημάτων απαιτεί διεπιστημονικές προσεγγίσεις, συνεργασία μεταξύ τεχνικών ειδικών, ρυθμιστικών αρχών και κοινωνικών επιστημόνων, καθώς και συνεχή επικαιροποίηση των κανόνων, ώστε να συμβαδίζουν με τη γρήγορη εξέλιξη της τεχνολογίας και των χρηματοοικονομικών πρακτικών (Aldboush & Ferdous, 2023; Agu et al., 2024; Dudu et al., 2024; Lee, 2020).

2.5.2 Περιορισμοί και Κατακερματισμός της Υπάρχουσας Βιβλιογραφίας

Η βιβλιογραφία για την τεχνητή νοημοσύνη στα χρηματοοικονομικά είναι εκτεταμένη, αλλά έντονα κατακερματισμένη. Οι περισσότερες μελέτες εστιάζουν είτε σε επιμέρους τεχνολογικές προσεγγίσεις (π.χ. μηχανική μάθηση, βαθιά ή ενισχυτική μάθηση), είτε σε μεμονωμένα πεδία εφαρμογής, όπως η πιστοληπτική αξιολόγηση, η ανίχνευση απάτης ή το algorithmic trading (Pattnaik et al., 2024· Maple et al., 2023· Wallon, 2019). Παράλληλα, ένα σημαντικό μέρος της βιβλιογραφίας προσεγγίζει το φαινόμενο κυρίως μέσα από το πρίσμα του FinTech και του ψηφιακού μετασχηματισμού του τραπεζικού συστήματος (Μπαλοδήμος, 2024· Σκαρλατάκης, 2022).

Αν και υπάρχουν αναλυτικές ιστορικές και θεματικές επισκοπήσεις της εξέλιξης της ΤΝ στον χρηματοπιστωτικό τομέα (Golić, 2019· Tarenzi, 2024· Vuković et al., 2025), η ανάλυση συχνά παραμένει σε μακρο-επίπεδο και δεν συνδέει συστηματικά τις τεχνολογικές εξελίξεις με συγκεκριμένες χρηματοοικονομικές ροές εργασίας και επιχειρησιακές διαδικασίες. Αντίστοιχα, ενώ οι βασικές κατηγορίες εφαρμογών της ΤΝ έχουν χαρτογραφηθεί επαρκώς (πιστωτικός κίνδυνος, απάτη, επενδύσεις, χρηματοοικονομική διοίκηση), παραμένει περιορισμένη η ενοποιημένη θεώρηση που να αντιστοιχίζει συγκεκριμένες τεχνικές ΤΝ με επιμέρους στάδια των χρηματοοικονομικών διαδικασιών (Golić, 2019).

Τέλος, η βιβλιογραφία για τα ηθικά και ρυθμιστικά ζητήματα είναι αυξανόμενη αλλά επίσης αποσπασματική, με επιμέρους μελέτες να εξετάζουν ζητήματα διαφάνειας, μεροληψίας, προστασίας δεδομένων και εποπτείας χωρίς ένα ενιαίο, ολοκληρωμένο πλαίσιο ανάλυσης (Agu et al., 2024 · Aldboush & Ferdous, 2023 · Dudu et al., 2024· Lee, 2020). Συνολικά, αναδεικνύεται η ανάγκη για πιο συνθετικές και διαδικασιοκεντρικές προσεγγίσεις, οι οποίες να γεφυρώνουν το χάσμα μεταξύ τεχνολογικών δυνατοτήτων και πρακτικής εφαρμογής στον χρηματοοικονομικό τομέα.

2.5.3 Ερευνητικά Κενά στις Χρηματοοικονομικές Εργασίες

Ωστόσο, μεγάλο μέρος αυτής της συζήτησης παραμένει αποσυνδεδεμένο από μια ολοκληρωμένη θεώρηση των «χρηματοοικονομικών εργασιών» ως ενιαίου οικοσυστήματος δραστηριοτήτων — από τον πιστωτικό κύκλο και το trading μέχρι τη λογιστική, τα εσωτερικά reports και τις RegTech λειτουργίες. Συχνά λείπει μια σύνθεση που να δείχνει πώς τα ηθικά και ρυθμιστικά ζητήματα διαπερνούν εγκάρσια όλες τις εφαρμογές της ΤΝ στον χρηματοοικονομικό κύκλο και όχι μόνο μεμονωμένα πεδία, όπως το credit scoring ή το fraud detection (Maple et al., 2023).

Ένα επιπλέον κενό αφορά τη συστηματική σύνδεση της διεθνούς βιβλιογραφίας με την ελληνική και ευρύτερα ευρωπαϊκή πραγματικότητα. Παρότι ορισμένες ελληνικές μελέτες αναδεικνύουν την επίδραση της ΤΝ στις τράπεζες και στις επιχειρήσεις, παραμένοντας όμως κυρίως περιπτωσιολογικές ή εστιασμένες σε επιμέρους διαστάσεις όπως ο ψηφιακός μετασχηματισμός ή τα υποδείγματα αξιολόγησης κινδύνου (Γκουνελάς, 2019; Κυδώνα, 2021; Μάργαρης, 2025; Μπαλοδήμος, 2024), δεν έχει ακόμη αναπτυχθεί ένα ενιαίο εννοιολογικό πλαίσιο που να συνθέτει τις κύριες τεχνολογίες (AI, ML, FinTech), τα ιστορικά «κύματα» υιοθέτησης, τα πρακτικά πεδία εφαρμογής και τα συναφή ηθικά/ρυθμιστικά ζητήματα σε μοντέλο προσαρμοσμένο ειδικά στην ελληνική και ευρωπαϊκή χρηματοοικονομική αγορά.

Επιπλέον, παρότι οι νεότερες εργασίες (π.χ. Vuković et al., 2025· Maple et al., 2023· Pattnaik et al., 2024) αναγνωρίζουν την ανάδυση της γενετικής ΤΝ και των μεγάλων γλωσσικών

μοντέλων (LLMs) σε δραστηριότητες όπως η ανάλυση οικονομικών καταστάσεων, τα regulatory copilots και τα προηγμένα chatbots, η σχετική συζήτηση παραμένει σε μεγάλο βαθμό διερευνητική και λιγότερο συστηματική. Λείπει μία συγκροτημένη αποτύπωση του πώς αυτές οι νέες τεχνολογίες επηρεάζουν το σύνολο των χρηματοοικονομικών εργασιών (λογιστική, reporting, risk, compliance, συμβουλευτική), ποια επιπλέον ηθικά/ρυθμιστικά ζητήματα δημιουργούν και πώς μπορούν να ενταχθούν υπεύθυνα σε υφιστάμενα επιχειρησιακά μοντέλα.

2.5.4 Θέση και Συμβολή της Παρούσας Διπλωματικής Εργασίας

Με βάση τα βιβλιογραφικά κενά που αναδείχθηκαν, η παρούσα διπλωματική εργασία τοποθετείται στη διασταύρωση της τεχνολογικής, εφαρμοσμένης, ηθικής και ρυθμιστικής συζήτησης γύρω από τη χρήση της τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα. Στόχος της είναι η ανάπτυξη μιας ενιαίας και συστηματικής θεώρησης των χρηματοοικονομικών εργασιών υπό το πρίσμα της ΤΝ, υπερβαίνοντας τον κατακερματισμό που χαρακτηρίζει μεγάλο μέρος της υφιστάμενης βιβλιογραφίας. Στο πλαίσιο αυτό, η εργασία δεν περιορίζεται σε περιγραφική παρουσίαση εφαρμογών, αλλά επιχειρεί μια δομημένη σύνθεση που συνδέει τις βασικές τεχνολογίες (AI, ML, FinTech), την ιστορική τους εξέλιξη και τις σύγχρονες μορφές χρήσης τους στον χρηματοοικονομικό τομέα.

Κεντρική συμβολή της εργασίας είναι η διαμόρφωση ενός εννοιολογικού πλαισίου που αντιστοιχίζει κατηγορίες χρηματοοικονομικών δραστηριοτήτων (όπως πιστωτικός κίνδυνος, επενδύσεις, διαχείριση κινδύνου, λογιστική και συμμόρφωση) με κατάλληλες κατηγορίες αλγορίθμων, λαμβάνοντας υπόψη όχι μόνο τεχνικά κριτήρια, αλλά και απαιτήσεις διαφάνειας, ερμηνευσιμότητας και αλγοριθμικής δικαιοσύνης. Παράλληλα, τα ηθικά και ρυθμιστικά ζητήματα ενσωματώνονται στην ανάλυση ως δομικά στοιχεία κάθε πεδίου εφαρμογής και όχι ως εξωτερικές ή δευτερεύουσες παράμετροι (Maple et al., 2023· Vuković et al., 2025).

Ιδιαίτερη έμφαση δίνεται στην ελληνική και ευρωπαϊκή πραγματικότητα, καθώς και στη συστηματική αποτύπωση αναδυόμενων εφαρμογών γενετικής ΤΝ και μεγάλων γλωσσικών μοντέλων σε λειτουργίες όπως η αναφορά, η συμμόρφωση και η ανάλυση εγγράφων. Η εργασία οργανώνει κριτικά μια έως σήμερα διάσπαρτη συζήτηση και διατυπώνει κατευθυντήριες αρχές για την υπεύθυνη ενσωμάτωση των τεχνολογιών αυτών στο περιβάλλον του Banking 4.0. Συνολικά, φιλοδοξεί να αποτελέσει ένα συνθετικό θεωρητικό υπόβαθρο τόσο για μελλοντική εμπειρική έρευνα όσο και για την πρακτική κατανόηση της υπεύθυνης αξιοποίησης της ΤΝ στον χρηματοοικονομικό τομέα (Maple et al., 2023).

ΚΕΦΑΛΑΙΟ 3: ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΚΑΙ ΨΗΦΙΟΠΟΙΗΣΗ

Το προηγούμενο κεφάλαιο παρουσίασε το θεωρητικό υπόβαθρο της τεχνητής νοημοσύνης και τις βασικές εφαρμογές της στα χρηματοοικονομικά, καθώς και τα κύρια ηθικά και ρυθμιστικά ζητήματα. Στο παρόν κεφάλαιο, η ανάλυση μετατοπίζεται στο χρηματοοικονομικό σύστημα μέσα στο οποίο οι τεχνολογίες αυτές εφαρμόζονται και αποκτούν οικονομική και θεσμική σημασία.

Στόχος του Κεφαλαίου 3 είναι να περιγράψει συνοπτικά τις βασικές δομές του σύγχρονου χρηματοοικονομικού πλαισίου (τράπεζες, χρηματοοικονομικές αγορές και οικοσύστημα FinTech) και να αναδείξει τον ψηφιακό μετασχηματισμό ως καταλύτη αλλαγών στη

λειτουργία τους. Με αυτόν τον τρόπο, τίθεται το απαραίτητο πλαίσιο για την κατανόηση του τρόπου με τον οποίο η τεχνητή νοημοσύνη ενσωματώνεται στις καθημερινές λειτουργίες τραπεζών, αγορών και χρηματοοικονομικών οργανισμών.

3.1 Τράπεζες στο Σύγχρονο Χρηματοοικονομικό Σύστημα

Οι τράπεζες αποτελούν διαχρονικά τον κεντρικό πυλώνα του χρηματοοικονομικού συστήματος, επιτελώντας κρίσιμες λειτουργίες που σχετίζονται με τη χρηματοπιστωτική διαμεσολάβηση, τη διαχείριση ρευστότητας, τη δημιουργία πίστωσης και τη μετάδοση της νομισματικής πολιτικής. Μέσω της συγκέντρωσης καταθέσεων και της παροχής δανείων, οι τράπεζες διασυνδέουν τα πλεονάζοντα κεφάλαια με τις επενδυτικές ανάγκες της οικονομίας, συμβάλλοντας στη χρηματοδότηση της ανάπτυξης και στη σταθερότητα του οικονομικού συστήματος (Mishkin, 2019).

Στο σύγχρονο χρηματοοικονομικό περιβάλλον, ο ρόλος των τραπεζών έχει διευρυνθεί και πολυπλοκοποιηθεί. Πέρα από την παραδοσιακή τραπεζική δραστηριότητα, οι τράπεζες παρέχουν ένα ευρύ φάσμα υπηρεσιών, όπως επενδυτική τραπεζική, διαχείριση περιουσίας, πληρωμές, υπηρεσίες θεματοφυλακής και υποστήριξη επιχειρηματικών συναλλαγών. Παράλληλα, λειτουργούν ως κομβικοί φορείς διαχείρισης κινδύνου, αξιολογώντας πιστωτικούς, επιτοκιακούς, συναλλαγματικούς και λειτουργικούς κινδύνους, τόσο σε επίπεδο μεμονωμένων πελατών όσο και σε επίπεδο χαρτοφυλακίων (Mishkin, 2019; Vives, 2019).

Η σημασία των τραπεζών ενισχύεται περαιτέρω από τον συστημικό τους ρόλο. Η σταθερότητα του τραπεζικού τομέα συνδέεται άμεσα με τη συνολική χρηματοπιστωτική σταθερότητα, καθώς δυσλειτουργίες ή κρίσεις στο τραπεζικό σύστημα μπορούν να μεταδοθούν ταχύτατα στην πραγματική οικονομία. Για τον λόγο αυτό, οι τράπεζες υπόκεινται σε αυστηρά ρυθμιστικά και εποπτικά πλαίσια, τα οποία αποσκοπούν στη διασφάλιση της κεφαλαιακής τους επάρκειας, της ρευστότητας και της ορθής διαχείρισης κινδύνων (Mishkin, 2019; Lehmann, 2020).

Ωστόσο, τις τελευταίες δεκαετίες το τραπεζικό σύστημα αντιμετωπίζει σημαντικές προκλήσεις που μεταβάλλουν τον παραδοσιακό του ρόλο. Η παγκοσμιοποίηση των αγορών, η τεχνολογική πρόοδος, οι αυξημένες κανονιστικές απαιτήσεις μετά τη χρηματοπιστωτική κρίση, καθώς και η εμφάνιση νέων παικτών στο χρηματοοικονομικό οικοσύστημα, έχουν εντείνει τον ανταγωνισμό και έχουν περιορίσει τα περιθώρια κερδοφορίας των τραπεζών. Παράλληλα, οι πελάτες απαιτούν ταχύτερες, φθηνότερες και περισσότερο εξατομικευμένες υπηρεσίες, γεγονός που ασκεί πρόσθετη πίεση στα παραδοσιακά τραπεζικά επιχειρησιακά μοντέλα (Vives, 2019; Philippon, 2016).

Στο πλαίσιο αυτό, οι τράπεζες καλούνται να επαναπροσδιορίσουν τον ρόλο τους, υιοθετώντας νέες τεχνολογίες και αναδιαμορφώνοντας τις εσωτερικές τους διαδικασίες. Η μετάβαση από την παραδοσιακή τραπεζική σε πιο ψηφιακά και δεδομενοκεντρικά μοντέλα λειτουργίας δεν αποτελεί απλώς τεχνολογική αναβάθμιση, αλλά στρατηγική επιλογή επιβίωσης και ανταγωνιστικότητας. Οι τράπεζες μετατρέπονται σταδιακά σε οργανισμούς έντασης δεδομένων, όπου η ανάλυση πληροφορίας, η αυτοματοποίηση διαδικασιών και η αξιοποίηση ψηφιακών πλατφορμών διαδραματίζουν καθοριστικό ρόλο (Vives, 2019).

Η εξέλιξη αυτή δημιουργεί το υπόβαθρο για την ενσωμάτωση τεχνολογιών τεχνητής νοημοσύνης σε κρίσιμες τραπεζικές λειτουργίες, όπως η πιστοδοτική αξιολόγηση, η ανίχνευση απάτης, η διαχείριση κινδύνου, η εξυπηρέτηση πελατών και η κανονιστική συμμόρφωση.

Παράλληλα, αναδιαμορφώνει τη σχέση των τραπεζών με τις χρηματοοικονομικές αγορές και το οικοσύστημα FinTech, ενισχύοντας τόσο τις συνεργασίες όσο και τις ανταγωνιστικές δυναμικές (Vives, 2019; Gomber et al., 2017).

Συνολικά, οι τράπεζες εξακολουθούν να αποτελούν θεμέλιο του σύγχρονου χρηματοοικονομικού συστήματος, αλλά ο ρόλος τους μετασχηματίζεται ουσιαστικά. Από παραδοσιακοί διαμεσολαβητές κεφαλαίων, εξελίσσονται σε ψηφιακούς οργανισμούς παροχής χρηματοοικονομικών υπηρεσιών, όπου η τεχνολογία και τα δεδομένα διαμορφώνουν νέες μορφές αξίας, αποτελεσματικότητας και ανταγωνισμού (Vives, 2019).

3.2 Ο Ρόλος των Χρηματοοικονομικών Αγορών στη Σύγχρονη Οικονομία

Οι χρηματοοικονομικές αγορές αποτελούν βασικό μηχανισμό λειτουργίας της σύγχρονης οικονομίας, καθώς διευκολύνουν τη μεταφορά κεφαλαίων μεταξύ αποταμιευτών και επενδυτών, τη διαμόρφωση τιμών χρηματοοικονομικών μέσων και τη διαχείριση κινδύνου. Μέσω των αγορών χρήματος και κεφαλαίου, των αγορών μετοχών, ομολόγων και παραγώγων, δημιουργούνται οι συνθήκες για την αποτελεσματική κατανομή των πόρων και τη χρηματοδότηση της οικονομικής δραστηριότητας (Bodie et al., 2018).

Σε αντίθεση με τον τραπεζικό τομέα, ο οποίος βασίζεται κυρίως στη διαμεσολάβηση μέσω ισολογισμών, οι χρηματοοικονομικές αγορές λειτουργούν ως αποκεντρωμένα συστήματα συναλλαγών, όπου οι τιμές διαμορφώνονται δυναμικά μέσω της προσφοράς και της ζήτησης. Η διαδικασία αυτή ενσωματώνει διαθέσιμες πληροφορίες σχετικά με τις προσδοκίες των επενδυτών, τις μακροοικονομικές συνθήκες και τον κίνδυνο, συμβάλλοντας στη διαμόρφωση ενός μηχανισμού τιμολόγησης που θεωρητικά προάγει την αποδοτικότητα της αγοράς (Bodie et al., 2018).

Ο ρόλος των χρηματοοικονομικών αγορών εκτείνεται πέρα από τη χρηματοδότηση επιχειρήσεων και κρατών. Μέσω των αγορών παραγώγων, παρέχονται εργαλεία αντιστάθμισης κινδύνου (hedging) έναντι μεταβολών επιτοκίων, συναλλαγματικών ισοτιμιών και τιμών εμπορευμάτων, επιτρέποντας στους οικονομικούς φορείς να διαχειρίζονται την αβεβαιότητα. Παράλληλα, οι αγορές συμβάλλουν στη διαφοροποίηση χαρτοφυλακίων και στη βελτίωση της κατανομής κινδύνου σε διεθνές επίπεδο (Bodie et al., 2018).

Στη σύγχρονη οικονομία, οι χρηματοοικονομικές αγορές χαρακτηρίζονται από αυξημένο βαθμό παγκοσμιοποίησης, ταχύτητας και διασύνδεσης. Η ψηφιοποίηση των συναλλαγών και η ανάπτυξη ηλεκτρονικών πλατφορμών έχουν ενισχύσει τη ρευστότητα και έχουν μειώσει το κόστος συναλλαγών, αλλά ταυτόχρονα έχουν αυξήσει την πολυπλοκότητα των αγορών. Η ταχύτατη διάδοση της πληροφορίας και η συμμετοχή μεγάλου αριθμού θεσμικών και ιδιωτών επενδυτών δημιουργούν ένα περιβάλλον στο οποίο οι αγορές αντιδρούν άμεσα σε οικονομικά, πολιτικά και τεχνολογικά γεγονότα (FSB, 2019).

Ιδιαίτερη σημασία αποκτά ο ρόλος των χρηματοοικονομικών αγορών στη μετάδοση κινδύνου και στη διαμόρφωση συστημικών φαινομένων. Η στενή διασύνδεση αγορών και θεσμών σημαίνει ότι διαταραχές σε ένα τμήμα του χρηματοοικονομικού συστήματος μπορούν να μεταδοθούν ταχύτατα σε άλλους τομείς και γεωγραφικές περιοχές. Για τον λόγο αυτό, οι αγορές αποτελούν βασικό αντικείμενο εποπτείας και ρύθμισης, με στόχο τη διασφάλιση της

διαφάνειας, της ακεραιότητας και της χρηματοπιστωτικής σταθερότητας (FSB, 2019; Lehmann, 2020).

Ταυτόχρονα, οι χρηματοοικονομικές αγορές λειτουργούν ως πεδίο έντονης τεχνολογικής καινοτομίας. Η μετάβαση από τις παραδοσιακές φυσικές αγορές σε πλήρως ηλεκτρονικά περιβάλλοντα συναλλαγών έχει μεταβάλει ριζικά τον τρόπο λειτουργίας τους. Νέα μοντέλα συναλλαγών, αυτοματοποιημένα συστήματα εκτέλεσης εντολών και δεδομενοκεντρικές πρακτικές ανάλυσης έχουν αναδιαμορφώσει τη δομή και τη δυναμική των αγορών, ενισχύοντας τόσο τις ευκαιρίες όσο και τους κινδύνους (FSB, 2019; Gomber et al., 2017).

Συνολικά, οι χρηματοοικονομικές αγορές διαδραματίζουν κρίσιμο ρόλο στη σύγχρονη οικονομία, λειτουργώντας ως μηχανισμοί χρηματοδότησης, τιμολόγησης και διαχείρισης κινδύνου. Η εξέλιξή τους σε ψηφιακά, ταχύτατα και υψηλής πολυπλοκότητας συστήματα δημιουργεί το πλαίσιο εντός του οποίου αναπτύσσονται νέες τεχνολογίες και επιχειρηματικά μοντέλα, προετοιμάζοντας το έδαφος για την ανάδυση του οικοσυστήματος FinTech και τον ευρύτερο ψηφιακό μετασχηματισμό του χρηματοοικονομικού τομέα (FSB, 2019; Gomber et al., 2017).

3.3 Το Οικοσύστημα FinTech στο Σύγχρονο Χρηματοοικονομικό Σύστημα

Το οικοσύστημα FinTech (Financial Technology) έχει αναδειχθεί τα τελευταία χρόνια σε έναν από τους βασικότερους παράγοντες αναδιάρθρωσης του χρηματοοικονομικού τομέα, επηρεάζοντας ουσιαστικά τον τρόπο παροχής, διάθεσης και κατανάλωσης χρηματοοικονομικών υπηρεσιών. Το FinTech δεν συνιστά έναν ενιαίο κλάδο, αλλά ένα πολυδιάστατο οικοσύστημα που περιλαμβάνει νεοφυείς επιχειρήσεις, τεχνολογικούς παρόχους, παραδοσιακούς χρηματοπιστωτικούς οργανισμούς, ρυθμιστικές αρχές και τελικούς χρήστες, οι οποίοι αλληλεπιδρούν σε ένα διαρκώς εξελισσόμενο ψηφιακό περιβάλλον (Gomber et al., 2017).

Η ανάπτυξη του οικοσυστήματος FinTech συνδέεται άμεσα με τη ραγδαία πρόοδο των ψηφιακών τεχνολογιών, τη διάδοση του διαδικτύου και των κινητών συσκευών, καθώς και με την αυξανόμενη διαθεσιμότητα και αξιοποίηση δεδομένων. Σε αντίθεση με τα παραδοσιακά χρηματοοικονομικά ιδρύματα, οι FinTech επιχειρήσεις χαρακτηρίζονται από ευέλικτες οργανωτικές δομές, έντονη καινοτομία και έμφαση στην εμπειρία του χρήστη, γεγονός που τους επιτρέπει να αναπτύσσουν ταχύτερα νέα προϊόντα και υπηρεσίες προσαρμοσμένες στις ανάγκες της ψηφιακής οικονομίας (Gomber et al., 2017; Vives, 2019).

Το οικοσύστημα FinTech καλύπτει ένα ευρύ φάσμα χρηματοοικονομικών δραστηριοτήτων, όπως οι ψηφιακές πληρωμές, οι πλατφόρμες δανεισμού και χρηματοδότησης (peer-to-peer lending, crowdfunding), οι αυτοματοποιημένες επενδυτικές υπηρεσίες (robo-advisors), η κανονιστική τεχνολογία (RegTech), η ασφαλιστική τεχνολογία (InsurTech) και οι εφαρμογές ανοικτής τραπεζικής (Open Banking). Μέσω αυτών των εφαρμογών, το FinTech συμβάλλει στη μείωση του κόστους συναλλαγών, στη βελτίωση της προσβασιμότητας σε χρηματοοικονομικές υπηρεσίες και στην ενίσχυση της χρηματοοικονομικής συμπερίληψης (Gomber et al., 2017; Philippon, 2016).

Η επίδραση του FinTech στις παραδοσιακές χρηματοοικονομικές υπηρεσίες είναι πολυεπίπεδη. Αφενός, δημιουργεί έντονο ανταγωνισμό προς τις τράπεζες και τους

καθιερωμένους χρηματοπιστωτικούς οργανισμούς, πιέζοντας τα επιχειρησιακά τους μοντέλα και τα περιθώρια κερδοφορίας. Αφετέρου, λειτουργεί ως καταλύτης συνεργασιών, καθώς πολλές τράπεζες επιλέγουν να συνεργαστούν με FinTech επιχειρήσεις ή να ενσωματώσουν τις τεχνολογικές τους λύσεις στις δικές τους πλατφόρμες, προκειμένου να επιταχύνουν τον ψηφιακό μετασχηματισμό τους (Vives, 2019; Philippon, 2016).

Ιδιαίτερη σημασία αποκτά η μετάβαση από το παραδοσιακό, κλειστό τραπεζικό μοντέλο σε πιο ανοικτά και διασυνδεδεμένα οικοσυστήματα χρηματοοικονομικών υπηρεσιών. Η ανάπτυξη του Open Banking και η χρήση application programming interfaces (APIs) επιτρέπουν τη διασύνδεση διαφορετικών παρόχων και την ανταλλαγή δεδομένων με τη συγκατάθεση των χρηστών, δημιουργώντας νέες δυνατότητες για καινοτόμα προϊόντα και εξατομικευμένες υπηρεσίες. Στο πλαίσιο αυτό, το FinTech συμβάλλει στη μετάβαση από την τραπεζοκεντρική παροχή υπηρεσιών σε πελατοκεντρικά ψηφιακά οικοσυστήματα (Vives, 2019; Gomber et al., 2017).

Παρά τα οφέλη που προσφέρει, η ανάπτυξη του οικοσυστήματος FinTech συνοδεύεται και από σημαντικές προκλήσεις. Ζητήματα που σχετίζονται με την ασφάλεια των δεδομένων, την προστασία της ιδιωτικότητας, τη διαχείριση κινδύνου, καθώς και τη ρυθμιστική συμμόρφωση αποκτούν αυξανόμενη σημασία. Η ταχεία καινοτομία συχνά προηγείται της προσαρμογής των κανονιστικών πλαισίων, γεγονός που δημιουργεί προκλήσεις τόσο για τις επιχειρήσεις όσο και για τις εποπτικές αρχές (Lehmann, 2020; Gomber et al., 2017).

Συνολικά, το οικοσύστημα FinTech αναδιαμορφώνει ριζικά τον χρηματοοικονομικό τομέα, μεταβάλλοντας τον τρόπο με τον οποίο σχεδιάζονται, προσφέρονται και χρησιμοποιούνται οι χρηματοοικονομικές υπηρεσίες. Η δυναμική αυτή δεν αντικαθιστά πλήρως τους παραδοσιακούς χρηματοπιστωτικούς θεσμούς, αλλά οδηγεί σε ένα υβριδικό μοντέλο, όπου συνεργασία και ανταγωνισμός συνυπάρχουν. Το FinTech λειτουργεί έτσι ως βασικός μοχλός του ευρύτερου ψηφιακού μετασχηματισμού του χρηματοοικονομικού συστήματος, προετοιμάζοντας το έδαφος για την ενσωμάτωση προηγμένων τεχνολογιών και νέων μορφών αυτοματοποίησης στις χρηματοοικονομικές εργασίες (Vives, 2019; Gomber et al., 2017).

3.4 Ο Ψηφιακός Μετασχηματισμός του Χρηματοοικονομικού Τομέα

Ο ψηφιακός μετασχηματισμός του χρηματοοικονομικού τομέα αποτελεί μια πολυδιάστατη και διαρκή διαδικασία, η οποία υπερβαίνει την απλή υιοθέτηση νέων τεχνολογιών και επηρεάζει σε βάθος τα επιχειρησιακά μοντέλα, τις οργανωτικές δομές και τον τρόπο παροχής χρηματοοικονομικών υπηρεσιών. Στον πυρήνα του ψηφιακού μετασχηματισμού βρίσκεται η συστηματική αξιοποίηση ψηφιακών τεχνολογιών, δεδομένων και αυτοματοποιημένων συστημάτων, με στόχο τη βελτίωση της αποδοτικότητας, της ευελιξίας και της εμπειρίας των χρηστών (Vives, 2019; Gomber et al., 2017).

Στον τραπεζικό και χρηματοοικονομικό κλάδο, ο ψηφιακός μετασχηματισμός εκδηλώνεται μέσα από την εκτεταμένη ψηφιοποίηση διαδικασιών, όπως το άνοιγμα λογαριασμών, η παροχή δανείων, η εκτέλεση συναλλαγών, η διαχείριση κινδύνου και η κανονιστική συμμόρφωση. Η μετάβαση από έντυπες και χειροκίνητες διαδικασίες σε πλήρως ψηφιακές ροές εργασίας μειώνει το λειτουργικό κόστος, περιορίζει τα σφάλματα και επιταχύνει τον χρόνο εξυπηρέτησης πελατών, ενισχύοντας παράλληλα τη διαφάνεια και τον έλεγχο (Vives, 2019).

Ιδιαίτερη σημασία στον ψηφιακό μετασχηματισμό έχει η μετάβαση σε δεδομένοκεντρικά (data-driven) επιχειρησιακά μοντέλα. Οι χρηματοοικονομικοί οργανισμοί συλλέγουν και

επεξεργάζονται πλέον μεγάλους όγκους δομημένων και μη δομημένων δεδομένων, προερχόμενων από συναλλαγές, αγορές, ψηφιακές πλατφόρμες και συμπεριφορά πελατών. Η αξιοποίηση των δεδομένων αυτών επιτρέπει την ανάπτυξη πιο ακριβών μοντέλων πρόβλεψης, τη βελτίωση της διαχείρισης κινδύνου και τη δημιουργία προσωποποιημένων χρηματοοικονομικών προϊόντων (Vives, 2019; Philippon, 2016).

Ο ψηφιακός μετασχηματισμός συνδέεται άμεσα και με την αλλαγή της σχέσης μεταξύ χρηματοπιστωτικών οργανισμών και πελατών. Η υιοθέτηση ψηφιακών καναλιών, όπως mobile banking, ψηφιακές πλατφόρμες επενδύσεων και αυτοματοποιημένα συστήματα εξυπηρέτησης, έχει μετατοπίσει το επίκεντρο από τη φυσική παρουσία στο κατάστημα προς μια συνεχώς διαθέσιμη, εξατομικευμένη και πολυκαναλική εμπειρία. Στο πλαίσιο αυτό, η εμπειρία χρήστη (user experience) και η ευκολία πρόσβασης καθίστανται κρίσιμα στοιχεία ανταγωνιστικότητας (Vives, 2019).

Παράλληλα, ο ψηφιακός μετασχηματισμός επηρεάζει βαθιά τις εσωτερικές λειτουργίες των χρηματοοικονομικών οργανισμών. Η αυτοματοποίηση διαδικασιών μέσω ψηφιακών εργαλείων και τεχνολογιών, όπως robotic process automation (RPA), επιτρέπει την αποδέσμευση ανθρώπινων πόρων από επαναλαμβανόμενες εργασίες και τη μετατόπισή τους σε πιο σύνθετες και στρατηγικές δραστηριότητες. Οι αλλαγές αυτές συνοδεύονται από νέες απαιτήσεις σε δεξιότητες, οργανωτική κουλτούρα και διαχείριση ανθρώπινου δυναμικού (Vives, 2019; Gomber et al., 2017).

Ωστόσο, ο ψηφιακός μετασχηματισμός του χρηματοοικονομικού τομέα συνοδεύεται και από σημαντικές προκλήσεις. Ζητήματα κυβερνοασφάλειας, προστασίας προσωπικών δεδομένων, επιχειρησιακής ανθεκτικότητας και κανονιστικής συμμόρφωσης αποκτούν αυξανόμενη σημασία σε ένα περιβάλλον υψηλής ψηφιοποίησης. Επιπλέον, η ταχύτητα των τεχνολογικών εξελίξεων δημιουργεί πιέσεις για συνεχή προσαρμογή τόσο των οργανισμών όσο και των εποπτικών αρχών (Lehmann, 2020; FSB, 2019).

Συνολικά, ο ψηφιακός μετασχηματισμός δεν αποτελεί ένα μεμονωμένο έργο ή μια προσωρινή στρατηγική επιλογή, αλλά έναν διαρκή μετασχηματισμό του χρηματοοικονομικού τομέα. Η ενοποίηση τραπεζών, χρηματοοικονομικών αγορών και οικοσυστημάτων FinTech σε ένα ενιαίο ψηφιακό περιβάλλον δημιουργεί τις προϋποθέσεις για νέες μορφές χρηματοοικονομικών υπηρεσιών και επιχειρησιακών μοντέλων. Το πλαίσιο αυτό λειτουργεί ως βάση για την περαιτέρω ενσωμάτωση προηγμένων τεχνολογιών, όπως η τεχνητή νοημοσύνη, οι οποίες αναλύονται στα επόμενα κεφάλαια της παρούσας εργασίας (Vives, 2019; FSB, 2019; Gomber et al., 2017).

ΚΕΦΑΛΑΙΟ 4: ΜΕΘΟΔΟΛΟΓΙΑ

4.1 Σχεδιασμός της Έρευνας

Η παρούσα μελέτη υιοθετεί ποσοτική και πειραματική ερευνητική προσέγγιση, με στόχο την ανάπτυξη και αξιολόγηση μοντέλων μηχανικής μάθησης για την πρόβλεψη της αθέτησης πληρωμής πιστωτικής κάρτας. Το πρόβλημα διατυπώνεται ως πρόβλημα επιτηρούμενης μάθησης (supervised learning) δυαδικής ταξινόμησης, όπου κάθε παρατήρηση αντιστοιχεί σε έναν πελάτη και η μεταβλητή-στόχος δηλώνει αν θα υπάρξει ή όχι αθέτηση πληρωμής κατά τον επόμενο μήνα (Lessmann et al., 2015).

Ο σχεδιασμός της έρευνας βασίστηκε σε μια σαφώς καθορισμένη μεθοδολογική ροή, η οποία περιλαμβάνει: τη συλλογή και κατανόηση των δεδομένων, την εξερευνητική ανάλυση δεδομένων (Exploratory Data Analysis – EDA), την προεπεξεργασία και τον μετασχηματισμό των χαρακτηριστικών, την εκπαίδευση επιλεγμένων μοντέλων ταξινόμησης και, τέλος, την αξιολόγηση και σύγκρισή τους με τη χρήση κατάλληλων μετρικών απόδοσης.

Για την εξαγωγή αξιόπιστων συμπερασμάτων, το σύνολο δεδομένων διαχωρίστηκε σε σύνολο εκπαίδευσης και σύνολο ελέγχου, ενώ ιδιαίτερη έμφαση δόθηκε στη διατήρηση της κατανομής των κλάσεων και στην αποφυγή διαρροής πληροφορίας κατά τη διαδικασία εκπαίδευσης των μοντέλων. Η επιλογή των μοντέλων και των μετρικών αξιολόγησης πραγματοποιήθηκε λαμβάνοντας υπόψη τη φύση του προβλήματος και την ανισορροπία των κλάσεων, ώστε η ανάλυση να αντικατοπτρίζει ρεαλιστικά σενάρια εφαρμογής στον χώρο της διαχείρισης πιστωτικού κινδύνου (Weiss, 2004).

4.2 Δεδομένα και Πηγή Δεδομένων

Η παρούσα μελέτη αξιοποιεί ένα ευρέως χρησιμοποιούμενο και αξιόπιστο σύνολο δεδομένων (dataset) που προέρχεται από την πλατφόρμα Kaggle, με τίτλο «Default of Credit Card Clients Dataset». Το συγκεκριμένο dataset δημιουργήθηκε αρχικά από το UCI Machine Learning Repository και έχει χρησιμοποιηθεί εκτενώς στη διεθνή βιβλιογραφία για την ανάπτυξη και αξιολόγηση μοντέλων πρόβλεψης πιστωτικού κινδύνου και εφαρμογών τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα (Yeh & Lien, 2009).

Το dataset περιέχει συνολικά 30.000 εγγραφές πελατών πιστωτικών καρτών και περιλαμβάνει 25 μεταβλητές, οι οποίες καλύπτουν τόσο δημογραφικά χαρακτηριστικά (όπως φύλο, ηλικία, εκπαίδευση και οικογενειακή κατάσταση), όσο και χρηματοοικονομικά στοιχεία (όπως πιστωτικό όριο, ιστορικό πληρωμών, ποσά λογαριασμών και προηγούμενες εξοφλήσεις). Η μεταβλητή-στόχος αφορά την ένδειξη αθέτησης πληρωμής (default) τον επόμενο μήνα, και λαμβάνει δυαδικές τιμές (0 = μη αθέτηση, 1 = αθέτηση) (UCI Machine Learning Repository, Kaggle Dataset).

Κεντρικός σκοπός της αξιοποίησης του συγκεκριμένου dataset είναι η ανάπτυξη και αξιολόγηση μοντέλων τεχνητής νοημοσύνης και μηχανικής μάθησης που μπορούν να προβλέψουν με υψηλή ακρίβεια την πιθανότητα αθέτησης υποχρεώσεων από πελάτες πιστωτικών καρτών. Η πρόβλεψη αυτή αποτελεί κρίσιμο εργαλείο για τα χρηματοπιστωτικά ιδρύματα, καθώς επιτρέπει την αποτελεσματικότερη διαχείριση πιστωτικού κινδύνου, τη βελτιστοποίηση της πιστοδοτικής πολιτικής και τη μείωση των επισφαλειών (Yeh & Lien, 2009; Lessmann et al., 2015).

Η επιλογή του συγκεκριμένου dataset βασίστηκε στην καταλληλότητά του για εφαρμογές τεχνητής νοημοσύνης στον χρηματοοικονομικό τομέα. Πιο συγκεκριμένα, το σχετικά μεγάλο μέγεθος του δείγματος, η ποικιλία των μεταβλητών και η ύπαρξη τόσο συμπεριφορικών όσο και δημογραφικών χαρακτηριστικών δημιουργούν τις κατάλληλες προϋποθέσεις για την εκπαίδευση, τη δοκιμή και τη συγκριτική αξιολόγηση διαφορετικών αλγορίθμων μηχανικής μάθησης. Επιπλέον, η δομή του προβλήματος (δυαδική ταξινόμηση) το καθιστά ιδανικό για την εφαρμογή τεχνικών πρόβλεψης που χρησιμοποιούνται ευρέως στη σύγχρονη χρηματοοικονομική τεχνολογία (Lessmann et al., 2015).

Το συγκεκριμένο σύνολο δεδομένων έχει χρησιμοποιηθεί σε πλήθος επιστημονικών μελετών που αφορούν την πρόβλεψη πιστωτικού κινδύνου, γεγονός που επιβεβαιώνει την αξιοπιστία

και τη χρησιμότητά του. Παράλληλα, η χρήση ενός πραγματικού και ευρέως αναγνωρισμένου συνόλου δεδομένων ενισχύει την εγκυρότητα των αποτελεσμάτων και επιτρέπει τη σύνδεση της παρούσας μελέτης με πραγματικές συνθήκες λειτουργίας των τραπεζικών και χρηματοπιστωτικών οργανισμών.

Το dataset είναι δημόσια διαθέσιμο στον ακόλουθο σύνδεσμο της πλατφόρμας Kaggle: <https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset>

4.3 Προετοιμασία Δεδομένων

Η προετοιμασία των δεδομένων αποτελεί ένα από τα πιο κρίσιμα και απαιτητικά στάδια της μεθοδολογίας, καθώς επηρεάζει άμεσα την ποιότητα, την αξιοπιστία και την προβλεπτική ικανότητα των μοντέλων τεχνητής νοημοσύνης που αναπτύσσονται. Η διαδικασία αυτή περιλαμβάνει μια σειρά από βήματα καθαρισμού, μετασχηματισμού και οργάνωσης των δεδομένων, με στόχο τη δημιουργία ενός συνεκτικού και κατάλληλου συνόλου εισόδου για την εκπαίδευση και αξιολόγηση των αλγορίθμων μηχανικής μάθησης.

Η υλοποίηση πραγματοποιήθηκε σε περιβάλλον Python, αξιοποιώντας τις βιβλιοθήκες pandas και NumPy για τη διαχείριση και επεξεργασία των δεδομένων, scikit-learn για την προεπεξεργασία, την εκπαίδευση και την αξιολόγηση των μοντέλων μηχανικής μάθησης, καθώς και τις βιβλιοθήκες matplotlib και seaborn για την οπτικοποίηση και την εξερευνητική ανάλυση των δεδομένων. Ο συνδυασμός των παραπάνω εργαλείων επιτρέπει την αποτελεσματική διαχείριση μεγάλων συνόλων δεδομένων και την αξιόπιστη εφαρμογή σύγχρονων τεχνικών ανάλυσης δεδομένων και τεχνητής νοημοσύνης.

Στα επόμενα υποκεφάλαια παρουσιάζονται αναλυτικά τα επιμέρους στάδια της διαδικασίας προετοιμασίας, τα οποία περιλαμβάνουν την εξερευνητική ανάλυση δεδομένων (Exploratory Data Analysis – EDA), τον καθαρισμό και τον μετασχηματισμό των μεταβλητών, καθώς και τη διαμόρφωση του τελικού συνόλου δεδομένων που χρησιμοποιήθηκε για την εκπαίδευση και την αξιολόγηση των μοντέλων.

4.3.1 Εξερευνητική Ανάλυση Δεδομένων

Η Εξερευνητική Ανάλυση Δεδομένων (Exploratory Data Analysis – EDA) αποτελεί ένα κρίσιμο αρχικό στάδιο στη διαδικασία ανάλυσης δεδομένων και ανάπτυξης μοντέλων μηχανικής μάθησης. Σκοπός της είναι η κατανόηση της δομής του συνόλου δεδομένων, η διερεύνηση της κατανομής των μεταβλητών, ο εντοπισμός πιθανών ανωμαλιών ή ακραίων τιμών, καθώς και η αξιολόγηση της συνολικής ποιότητας των δεδομένων πριν από το στάδιο της μοντελοποίησης (Komorowski et al., 2016). Μέσω της EDA καθίσταται δυνατή η έγκαιρη ανίχνευση προβλημάτων που θα μπορούσαν να επηρεάσουν αρνητικά την απόδοση και την αξιοπιστία των αλγορίθμων τεχνητής νοημοσύνης.

Στο πλαίσιο της παρούσας μελέτης, η EDA πραγματοποιήθηκε με χρήση της γλώσσας Python και των βιβλιοθηκών pandas, numpy, matplotlib και seaborn. Ο πλήρης κώδικας της διαδικασίας παρατίθεται στο Παράρτημα Α.

1. Επισκόπηση Δομής Δεδομένων και Τύπων Μεταβλητών

Αρχικά πραγματοποιήθηκε η φόρτωση του συνόλου δεδομένων και ο έλεγχος των διαστάσεων του, καθώς και η προβολή των πρώτων εγγραφών με τη χρήση της εντολής `df.head()`, προκειμένου να επιβεβαιωθεί η ορθή εισαγωγή των δεδομένων. Στη συνέχεια, χρησιμοποιήθηκε η εντολή `df.info()` για την επισκόπηση των τύπων δεδομένων κάθε μεταβλητής και τον έλεγχο της πληρότητας των στηλών.

Από τον έλεγχο αυτό διαπιστώθηκε ότι το dataset αποτελείται εξολοκλήρου από αριθμητικές μεταβλητές, εκ των οποίων ορισμένες αντιστοιχούν σε κατηγορικά χαρακτηριστικά κωδικοποιημένα αριθμητικά (όπως το φύλο, το επίπεδο εκπαίδευσης και η οικογενειακή κατάσταση). Επιπλέον, η μεταβλητή ID αναγνωρίστηκε ως απλό αναγνωριστικό εγγραφής χωρίς προβλεπτική αξία και, για τον λόγο αυτό, αφαιρέθηκε πριν από τη διαδικασία μοντελοποίησης.

2. Στατιστική Περιγραφή των Μεταβλητών

Στη συνέχεια πραγματοποιήθηκε περιγραφική στατιστική ανάλυση των αριθμητικών μεταβλητών με τη χρήση της εντολής `df.describe()`. Για κάθε μεταβλητή υπολογίστηκαν βασικοί στατιστικοί δείκτες, όπως ο μέσος όρος, η τυπική απόκλιση, οι ελάχιστες και μέγιστες τιμές, καθώς και τα τεταρτημόρια.

Η ανάλυση αυτή παρείχε μια συνοπτική εικόνα της κεντρικής τάσης και της διασποράς των δεδομένων, ενώ επέτρεψε τον αρχικό εντοπισμό μεταβλητών με μεγάλη διακύμανση ή πιθανές ακραίες τιμές, οι οποίες θα μπορούσαν να επηρεάσουν αρνητικά τη συμπεριφορά και την απόδοση των μοντέλων μηχανικής μάθησης (Komorowski et al., 2016).

3. Έλεγχος Ελλειπουσών Τιμών

Ο έλεγχος για ελλείπουσες τιμές πραγματοποιήθηκε με τη χρήση της εντολής `df.isnull().sum()`. Από τα αποτελέσματα προέκυψε ότι το σύνολο δεδομένων δεν περιέχει ελλείπουσες τιμές σε καμία από τις μεταβλητές, γεγονός που απλοποίησε σημαντικά το στάδιο της προεπεξεργασίας και ενισχύει την αξιοπιστία της ανάλυσης (Kwak & Kim, 2017)..

Η απουσία ελλειπόν τιμών μειώνει τον κίνδυνο εισαγωγής μεροληψίας και απώλειας πληροφορίας κατά τη διαδικασία εκπαίδευσης των μοντέλων, φαινόμενα που συχνά παρατηρούνται σε πραγματικά χρηματοοικονομικά δεδομένα (Weiss, 2004).

4. Ανάλυση της Μεταβλητής-στόχου και Ανισορροπία Κλάσεων

Η μεταβλητή-στόχος της παρούσας μελέτης είναι η `default.payment.next.month`, η οποία λαμβάνει δυαδικές τιμές και δηλώνει αν ένας πελάτης αθέτησε ή όχι την υποχρέωση πληρωμής του κατά τον επόμενο μήνα. Αρχικά υπολογίστηκε η κατανομή των κλάσεων τόσο σε απόλυτους αριθμούς όσο και σε ποσοστιαία μορφή.

Η ανάλυση έδειξε ότι το σύνολο δεδομένων παρουσιάζει σαφή ανισορροπία κλάσεων (class imbalance), με την πλειονότητα των παρατηρήσεων να ανήκει στην κατηγορία των μη αθετήσεων. Το φαινόμενο αυτό είναι ιδιαίτερα συνηθισμένο σε προβλήματα πιστωτικού κινδύνου και δύναται να επηρεάσει σημαντικά τόσο τη διαδικασία εκπαίδευσης όσο και την αξιολόγηση των μοντέλων, καθιστώντας αναγκαία τη χρήση κατάλληλων μετρικών απόδοσης

πέραν της απλής ακρίβειας (accuracy). Για τον λόγο αυτό, η κατανομή της μεταβλητής-στόχου απεικονίστηκε γραφικά μέσω ραβδογράμματος (bar plot), ώστε να καταστεί εμφανής η έκταση της ανισορροπίας μεταξύ των δύο κλάσεων (Weiss, 2004; Fawcett, 2006).

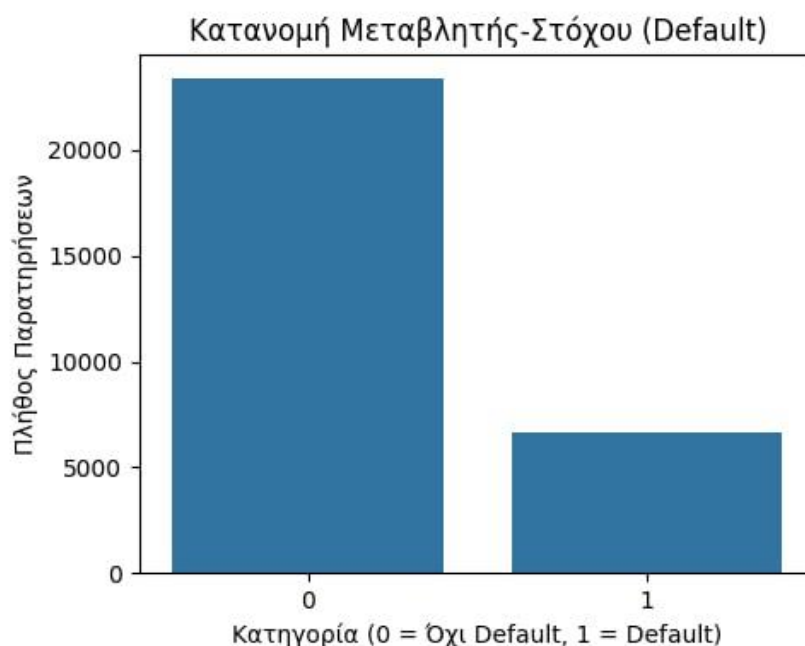
Όπως φαίνεται και στον Πίνακα 2 και στο αντίστοιχο ραβδόγραμμα, η κλάση της αθέτησης αντιπροσωπεύει περίπου το 22% του συνόλου των παρατηρήσεων.

Πίνακας 1 Κατανομή κατηγορικών μεταβλητών

Μεταβλητή	Κατηγορία	Περιγραφή	Πλήθος
SEX	1	Άνδρας	11.888
	2	Γυναίκα	18.112
EDUCATION	0	Άγνωστο	14
	1	Μεταπτυχιακό / Διδακτορικό	10.585
	2	Πανεπιστήμιο	14.030
	3	Λύκειο	4.917
	4	Άλλο	123
	5	Άγνωστο	280
	6	Άγνωστο	51
MARRIAGE	0	Άγνωστο	54
	1	Έγγαμος/η	13.659
	2	Άγαμος/η	15.964
	3	Άλλο	323

Πίνακας 2 Μεταβλητή-στόχος Default

Default	Πλήθος	Ποσοστό
0 Όχι Default	23364	77.88
1 Default	6636	22.12



Εικόνα 1 Κατανομή μεταβλητής-στόχου (Default)

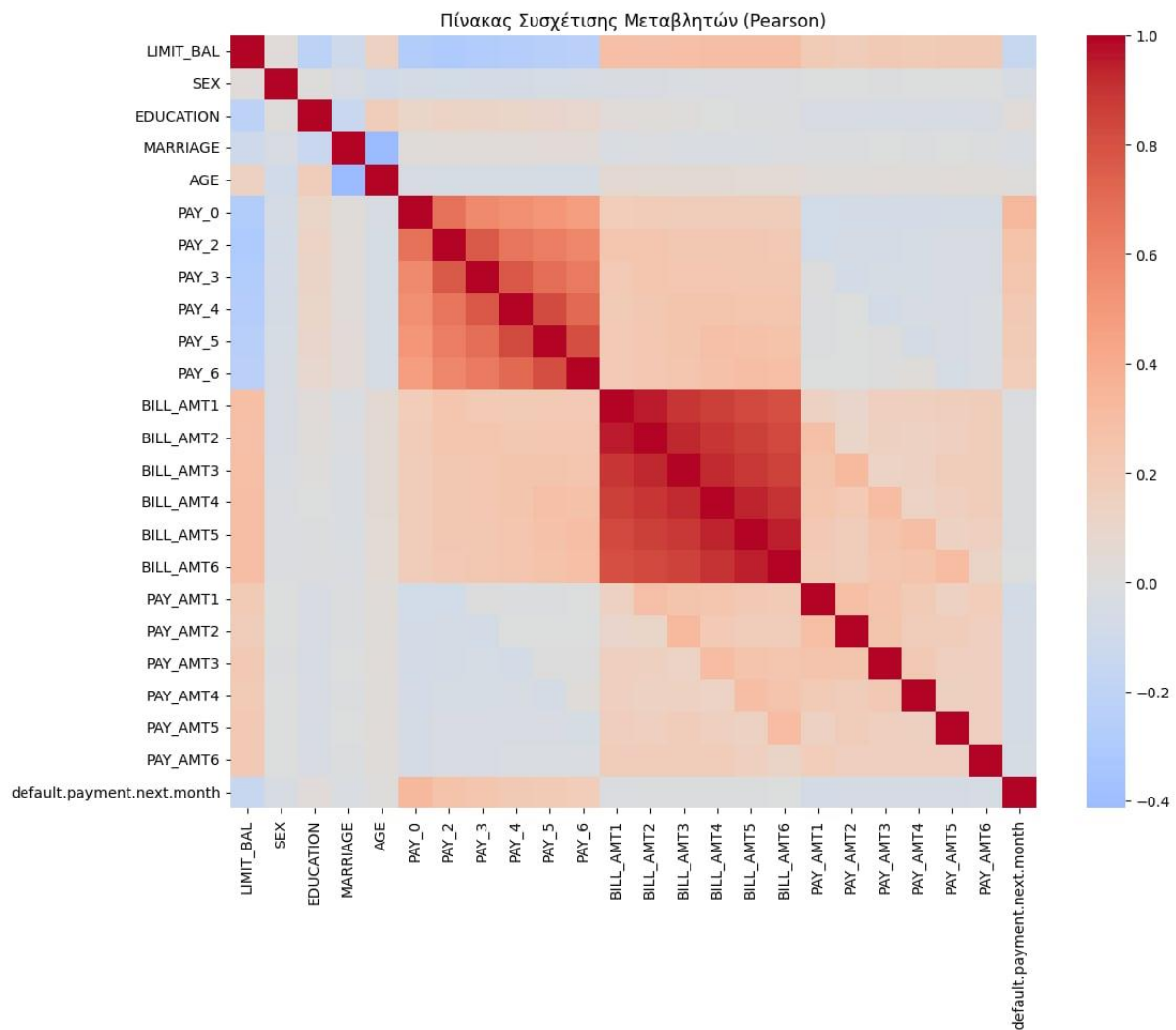
5. Ανάλυση Συσχετίσεων Μεταξύ Μεταβλητών

Για τη διερεύνηση των γραμμικών σχέσεων μεταξύ των αριθμητικών μεταβλητών του συνόλου δεδομένων υπολογίστηκε ο πίνακας συσχέτισης Pearson και απεικονίστηκε γραφικά μέσω θερμικού χάρτη (heatmap), όπως φαίνεται στην Εικόνα 2. Ο συντελεστής συσχέτισης Pearson ποσοτικοποιεί τον βαθμό γραμμικής συσχέτισης μεταξύ δύο μεταβλητών και λαμβάνει τιμές από -1 έως +1.

Από την ανάλυση της Εικόνας 3 προκύπτει ότι ομάδες μεταβλητών που αναφέρονται σε διαδοχικές χρονικές περιόδους, όπως τα υπόλοιπα λογαριασμών (BILL_AMT1–BILL_AMT6), τα ποσά πληρωμών (PAY_AMT1–PAY_AMT6), καθώς και οι δείκτες καθυστέρησης πληρωμών (PAY_0–PAY_6), εμφανίζουν ισχυρές μεταξύ τους συσχετίσεις. Το εύρημα αυτό είναι αναμενόμενο, καθώς οι μεταβλητές αυτές αποτυπώνουν την ίδια οικονομική συμπεριφορά σε διαδοχικούς μήνες.

Αντιθέτως, η μεταβλητή-στόχος `default.payment.next.month` δεν παρουσιάζει ισχυρή γραμμική συσχέτιση με καμία μεμονωμένη μεταβλητή, γεγονός που υποδηλώνει ότι η αθέτηση πληρωμής αποτελεί αποτέλεσμα συνδυασμού πολλών παραγόντων και όχι ενός μόνο χαρακτηριστικού.

Η παρατήρηση αυτή, η οποία αποτυπώνεται καθαρά στον θερμικό χάρτη συσχέτισης, υποστηρίζει τη χρήση μη γραμμικών μοντέλων μηχανικής μάθησης, όπως τα Δέντρα Απόφασης, τα οποία είναι σε θέση να συλλάβουν σύνθετες και μη γραμμικές σχέσεις μεταξύ των μεταβλητών (Lessmann et al., 2015).



Εικόνα 2 Πίνακας συσχέτισης Pearson σε μορφή θερμικού χάρτη (heatmap)

6. Συνοπτική Αποτίμηση της Εξερευνητικής Ανάλυσης Δεδομένων (EDA)

Η εξερευνητική ανάλυση δεδομένων παρείχε μια ολοκληρωμένη εικόνα της δομής, της ποιότητας και των βασικών στατιστικών χαρακτηριστικών του συνόλου δεδομένων. Διαπιστώθηκε ότι το dataset είναι πλήρες, χωρίς ελλείπουσες τιμές, και αποτελείται κυρίως από αριθμητικές μεταβλητές, καθώς και από ορισμένες κατηγορικές μεταβλητές κωδικοποιημένες αριθμητικά.

Η ανάλυση της μεταβλητής-στόχου ανέδειξε σαφή ανισορροπία κλάσεων, χαρακτηριστικό σύνηθες σε προβλήματα πιστωτικού κινδύνου, γεγονός που καθιστά αναγκαία τη χρήση κατάλληλων μετρικών αξιολόγησης πέραν της απλής ακρίβειας. Παράλληλα, η ανάλυση συσχετίσεων έδειξε ότι δεν υπάρχει ισχυρή γραμμική σχέση μεταξύ της μεταβλητής-στόχου και κάποιας μεμονωμένης ανεξάρτητης μεταβλητής, ενώ παρατηρούνται ισχυρές συσχετίσεις μεταξύ ομάδων μεταβλητών που αφορούν διαδοχικές χρονικές περιόδους.

Τα ευρήματα αυτά υποδεικνύουν ότι το πρόβλημα της πρόβλεψης αθέτησης πληρωμής είναι πολυπαραγοντικό και δεν μπορεί να περιγραφεί επαρκώς μέσω απλών γραμμικών σχέσεων.

Με βάση τα παραπάνω ευρήματα, στην επόμενη ενότητα ακολουθεί η ανάπτυξη και σύγκριση (Weiss, 2004).

4.3.2 Προεπεξεργασία Δεδομένων και Μετασχηματισμοί

Πριν από την εκπαίδευση των μοντέλων μηχανικής μάθησης, πραγματοποιήθηκε προεπεξεργασία των δεδομένων, με στόχο την κατάλληλη διαμόρφωσή τους και τη διασφάλιση της ορθής λειτουργίας των αλγορίθμων.

Αρχικά, η μεταβλητή ID, η οποία λειτουργεί αποκλειστικά ως αναγνωριστικό εγγραφής και δεν φέρει προβλεπτική πληροφορία, αφαιρέθηκε από το σύνολο δεδομένων. Στη συνέχεια, τα δεδομένα διαχωρίστηκαν σε ανεξάρτητες μεταβλητές (X) και στη μεταβλητή-στόχο (y).

Το σύνολο δεδομένων χωρίστηκε σε σύνολο εκπαίδευσης (training set) και σύνολο ελέγχου (test set) με αναλογία 80% – 20%, χρησιμοποιώντας στρατηγική διαστρωμάτωσης (stratified split), ώστε να διατηρηθεί η ίδια αναλογία κλάσεων και στα δύο σύνολα.

Για την κατάλληλη προετοιμασία των μεταβλητών, εφαρμόστηκε διαφορετική προεπεξεργασία σε αριθμητικά και κατηγορικά χαρακτηριστικά. Συγκεκριμένα, οι αριθμητικές μεταβλητές κανονικοποιήθηκαν με τη μέθοδο της τυποποίησης (Standardization) μέσω του μετασχηματισμού StandardScaler, ώστε να έχουν μηδενικό μέσο όρο και μοναδιαία τυπική απόκλιση. Οι κατηγορικές μεταβλητές (SEX, EDUCATION, MARRIAGE), οι οποίες είναι κωδικοποιημένες αριθμητικά, μετατράπηκαν σε δυαδικές μεταβλητές με τη χρήση της τεχνικής One-Hot Encoding.

Η όλη διαδικασία προεπεξεργασίας και εκπαίδευσης των μοντέλων υλοποιήθηκε μέσω της χρήσης Pipelines και ColumnTransformer της βιβλιοθήκης scikit-learn, ώστε να διασφαλιστεί η ορθή και αναπαραγώγιμη εφαρμογή των ίδιων μετασχηματισμών τόσο στα δεδομένα εκπαίδευσης όσο και στα δεδομένα ελέγχου, αποφεύγοντας φαινόμενα διαρροής πληροφορίας (data leakage).

4.4 Μοντέλα Ταξινόμησης

Για την πρόβλεψη της αθέτησης πληρωμής πιστωτικής κάρτας χρησιμοποιήθηκαν δύο διαφορετικά μοντέλα μηχανικής μάθησης: η Λογιστική Παλινδρόμηση (Logistic Regression) και το Δέντρο Απόφασης (Decision Tree). Η επιλογή αυτή επιτρέπει τη σύγκριση ενός γραμμικού και ενός μη γραμμικού ταξινομητή, προσφέροντας μια πιο ολοκληρωμένη εικόνα της φύσης του προβλήματος και της ικανότητας διαφορετικών μεθόδων να συλλάβουν τις υποκείμενες σχέσεις στα δεδομένα (Lessmann et al., 2015).

Πριν την εκπαίδευση των μοντέλων, το σύνολο δεδομένων διαχωρίστηκε σε σύνολο εκπαίδευσης (training set) και σύνολο ελέγχου (test set) με αναλογία 80/20, χρησιμοποιώντας στρωματοποιημένο διαχωρισμό (stratified split), ώστε να διατηρηθεί η ίδια κατανομή των κλάσεων και στα δύο σύνολα. Η προεπεξεργασία των δεδομένων (κανονικοποίηση των αριθμητικών μεταβλητών και κωδικοποίηση των κατηγορικών) ενσωματώθηκε σε ενιαίο pipeline μαζί με κάθε μοντέλο, ώστε να αποφεύγεται η διαρροή πληροφορίας (data leakage) από το σύνολο εκπαίδευσης στο σύνολο ελέγχου (Weiss, 2004).

4.4.1 Λογιστική Παλινδρόμηση (Logistic Regression)

Η Λογιστική Παλινδρόμηση αποτελεί ένα από τα πιο διαδεδομένα και θεμελιώδη μοντέλα για προβλήματα δυαδικής ταξινόμησης. Το μοντέλο υπολογίζει την πιθανότητα μια παρατήρηση να ανήκει στην κλάση της αθέτησης πληρωμής μέσω της λογιστικής (sigmoid) συνάρτησης, η οποία εφαρμόζεται σε έναν γραμμικό συνδυασμό των εισόδων. Με τον τρόπο αυτό, οι προβλέψεις λαμβάνουν τιμές στο διάστημα $[0, 1]$ και μπορούν να ερμηνευθούν ως πιθανότητες.

Η τελική απόφαση προκύπτει με την εφαρμογή ενός κατωφλίου (threshold), το οποίο στην παρούσα εργασία ορίστηκε ίσο με 0.5. Παρότι πρόκειται για γραμμικό μοντέλο, το οποίο υποθέτει γραμμικό διαχωρισμό μεταξύ των κλάσεων στον χώρο των χαρακτηριστικών, η Λογιστική Παλινδρόμηση χρησιμοποιείται ευρέως ως μοντέλο αναφοράς (baseline), λόγω της σταθερότητας, της ταχύτητας εκπαίδευσης και της υψηλής ερμηνευσιμότητάς της.

Στην υλοποίηση χρησιμοποιήθηκε η LogisticRegression της βιβλιοθήκης scikit-learn με κανονικοποίηση L2 (L2 regularization), η οποία προσθέτει όρο ποινής στο μέγεθος των συντελεστών του μοντέλου με στόχο τον περιορισμό της υπερεκμάθησης (overfitting). Ο μέγιστος αριθμός επαναλήψεων (max_iter) ορίστηκε σε 2000, ώστε να διασφαλιστεί η σύγκλιση της διαδικασίας βελτιστοποίησης, δεδομένου του μεγάλου αριθμού χαρακτηριστικών που προκύπτουν μετά την κωδικοποίηση των κατηγορικών μεταβλητών.

Η Λογιστική Παλινδρόμηση χρησιμοποιείται στο πλαίσιο της παρούσας μελέτης ως μοντέλο αναφοράς, πάνω στο οποίο συγκρίνεται η απόδοση πιο σύνθετων μη γραμμικών μεθόδων.

4.4.2 Δέντρο Απόφασης (Decision Tree)

Το Δέντρο Απόφασης αποτελεί ένα μη παραμετρικό και μη γραμμικό μοντέλο ταξινόμησης, το οποίο αναπαριστά τη διαδικασία λήψης απόφασης ως μια ιεραρχική δομή διαδοχικών διαχωρισμών (splits) πάνω στα χαρακτηριστικά των δεδομένων. Κάθε εσωτερικός κόμβος του δέντρου αντιστοιχεί σε έναν έλεγχο πάνω σε μία μεταβλητή, ενώ τα φύλλα του δέντρου αντιστοιχούν στις τελικές προβλέψεις κλάσης.

Ένα από τα βασικά πλεονεκτήματα των Δέντρων Απόφασης είναι η ικανότητά τους να μοντελοποιούν πολύπλοκες και μη γραμμικές σχέσεις μεταξύ των μεταβλητών, χωρίς να απαιτούν υποθέσεις για τη μορφή των κατανομών των δεδομένων. Επιπλέον, τα δέντρα χαρακτηρίζονται από υψηλή ερμηνευσιμότητα, καθώς η λογική των αποφάσεων μπορεί να αναλυθεί βήμα προς βήμα.

Ωστόσο, τα Δέντρα Απόφασης είναι ιδιαίτερα επιρρεπή στην υπερεκμάθηση. Για τον λόγο αυτό, στην παρούσα εργασία χρησιμοποιήθηκε περιορισμένο δέντρο με μέγιστο βάθος (max_depth = 6) και ελάχιστο αριθμό δειγμάτων ανά φύλλο (min_samples_leaf = 50). Επιπλέον, εφαρμόστηκε στάθμιση κλάσεων (class_weight = "balanced"), ώστε να αντιμετωπιστεί η ανισορροπία των κλάσεων στο σύνολο δεδομένων και να δοθεί μεγαλύτερη σημασία στην ορθή πρόβλεψη της μειοψηφικής κλάσης της αθέτησης.

Το Δέντρο Απόφασης χρησιμοποιείται στην παρούσα μελέτη ως αντιπροσωπευτικό μη γραμμικό μοντέλο, ώστε να συγκριθεί η απόδοσή του με εκείνη της Λογιστικής Παλινδρόμησης.

4.4.3 Συγκριτική Θεώρηση των δύο Μοντέλων

Η ταυτόχρονη χρήση της Λογιστικής Παλινδρόμησης και του Δέντρου Απόφασης επιτρέπει τη σύγκριση ενός απλού, γραμμικού και ερμηνεύσιμου μοντέλου με ένα πιο ευέλικτο, μη γραμμικό μοντέλο μεγαλύτερης εκφραστικής ικανότητας. Με τον τρόπο αυτό, καθίσταται δυνατή η διερεύνηση του κατά πόσο το πρόβλημα της πρόβλεψης αθέτησης πληρωμής μπορεί να περιγραφεί επαρκώς μέσω γραμμικών σχέσεων ή αν απαιτείται η χρήση πιο σύνθετων μη γραμμικών προσεγγίσεων.

Η συμπληρωματική αυτή προσέγγιση επιτρέπει μια πιο σφαιρική αξιολόγηση των μοντέλων, η οποία παρουσιάζεται στην επόμενη ενότητα.

4.5 Αξιολόγηση Μοντέλων

Η απόδοση των δύο μοντέλων, της Λογιστικής Παλινδρόμησης και του Δέντρου Απόφασης, αξιολογήθηκε με τη χρήση πολλαπλών μετρικών ταξινόμησης, ώστε να προκύψει μια ολοκληρωμένη και αξιόπιστη εικόνα της συμπεριφοράς τους στο πρόβλημα της πρόβλεψης αθέτησης πληρωμής πιστωτικής κάρτας. Δεδομένης της ανισορροπίας των κλάσεων στο σύνολο δεδομένων, η αξιολόγηση δεν βασίστηκε αποκλειστικά στην ακρίβεια (accuracy), αλλά δόθηκε έμφαση σε μετρικές που εστιάζουν στη σωστή ανίχνευση της μειοψηφικής κλάσης της αθέτησης (Weiss, 2004).

Οι βασικές μετρικές που χρησιμοποιήθηκαν είναι οι εξής: Accuracy, Precision, Recall, F1-score και το εμβαδόν κάτω από την καμπύλη ROC (ROC-AUC).

Η ακρίβεια (Accuracy) εκφράζει το ποσοστό των συνολικά σωστών προβλέψεων. Ωστόσο, σε προβλήματα με έντονη ανισορροπία κλάσεων, όπως το παρόν, η ακρίβεια μπορεί να είναι παραπλανητική, καθώς ένα μοντέλο μπορεί να επιτύχει υψηλή ακρίβεια προβλέποντας σχεδόν πάντα την πλειοψηφική κλάση των μη αθετήσεων, χωρίς να ανιχνεύει επαρκώς τις πραγματικές περιπτώσεις αθέτησης.

Για τον λόγο αυτό, χρησιμοποιούνται οι μετρικές Precision και Recall για την κλάση της αθέτησης (θετική κλάση). Το Precision εκφράζει το ποσοστό των προβλέψεων αθέτησης που είναι πράγματι σωστές, ενώ το Recall (ευαισθησία) εκφράζει το ποσοστό των πραγματικών αθετήσεων που το μοντέλο καταφέρνει να εντοπίσει. Μεταξύ των δύο αυτών μετρικών υπάρχει συνήθως ένας συμβιβασμός (trade-off): η αύξηση της ευαισθησίας οδηγεί συχνά σε περισσότερους ψευδείς συναγερμούς, ενώ η αύξηση της ακρίβειας των θετικών προβλέψεων μπορεί να μειώσει την κάλυψη των πραγματικών αθετήσεων.

Το F1-score αποτελεί τον αρμονικό μέσο των Precision και Recall και παρέχει μια ισορροπημένη συνολική εκτίμηση της απόδοσης ως προς την ανίχνευση της κλάσης της αθέτησης. Στο παρόν πρόβλημα, το F1-score θεωρείται ιδιαίτερα κατάλληλο μέτρο, καθώς τιμωρεί τόσο την χαμηλή ευαισθησία όσο και την χαμηλή ακρίβεια των θετικών προβλέψεων.

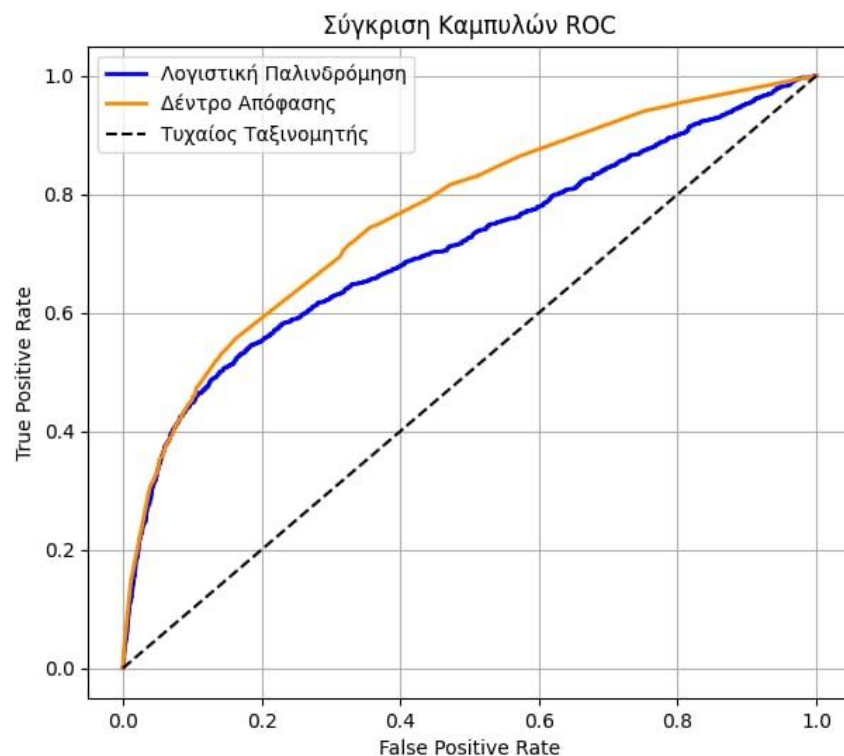
Για μια πιο συνολική και ανεξάρτητη από το κατώφλι ταξινόμησης αξιολόγηση, χρησιμοποιήθηκε η καμπύλη ROC (Receiver Operating Characteristic) και το αντίστοιχο εμβαδόν κάτω από αυτήν (Area Under the Curve – AUC). Η καμπύλη ROC απεικονίζει τον ρυθμό αληθώς θετικών προβλέψεων (True Positive Rate) ως προς τον ρυθμό ψευδώς θετικών προβλέψεων (False Positive Rate) για όλα τα δυνατά κατώφλια απόφασης. Ένα ιδανικό

μοντέλο θα είχε καμπύλη που προσεγγίζει την επάνω αριστερή γωνία του διαγράμματος, ενώ η διαγώνιος αντιστοιχεί σε τυχαίο ταξινομητή (Fawcett, 2006).

Στην Εικόνα 5 παρουσιάζεται η σύγκριση των καμπυλών ROC για τα δύο μοντέλα. Παρατηρείται ότι το Δέντρο Απόφασης παρουσιάζει ελαφρώς καλύτερη διακριτική ικανότητα σε σχέση με τη Λογιστική Παλινδρόμηση, γεγονός που αποτυπώνεται και στις τιμές του ROC-AUC. Το αποτέλεσμα αυτό υποδηλώνει ότι το μη γραμμικό μοντέλο καταφέρνει να συλλάβει πιο αποτελεσματικά τις σύνθετες σχέσεις μεταξύ των μεταβλητών που σχετίζονται με τον πιστωτικό κίνδυνο.

Συγκεκριμένα, από τα πειραματικά αποτελέσματα προκύπτει ότι η Λογιστική Παλινδρόμηση πέτυχε τιμή ROC-AUC ίση με 0.71, ενώ το Δέντρο Απόφασης παρουσίασε υψηλότερη τιμή, ίση με 0.765, γεγονός που υποδηλώνει καλύτερη συνολική διακριτική ικανότητα του δεύτερου μοντέλου στον διαχωρισμό μεταξύ πελατών που θα αθετήσουν και αυτών που δεν θα αθετήσουν την πληρωμή τους. Παράλληλα, το Δέντρο Απόφασης εμφάνισε σημαντικά υψηλότερη ευαισθησία (recall) για την κλάση της αθέτησης (περίπου 54%) σε σύγκριση με τη Λογιστική Παλινδρόμηση (περίπου 24%), γεγονός ιδιαίτερα κρίσιμο για εφαρμογές διαχείρισης πιστωτικού κινδύνου, όπου το κόστος μη ανίχνευσης ενός επισφαλούς πελάτη είναι ιδιαίτερα υψηλό. Αντίθετα, η Λογιστική Παλινδρόμηση παρουσίασε ελαφρώς υψηλότερη συνολική ακρίβεια (accuracy), ωστόσο το αποτέλεσμα αυτό οφείλεται κυρίως στην καλή πρόβλεψη της πλειοψηφικής κλάσης των μη αθετήσεων και δεν αντικατοπτρίζει την ικανότητά της να εντοπίζει αποτελεσματικά τις πραγματικές περιπτώσεις αθέτησης (Fawcett, 2006; Weiss, 2004).

Τα αποτελέσματα αυτά συζητούνται περαιτέρω στο επόμενο κεφάλαιο, στο πλαίσιο των δυνατοτήτων και των περιορισμών κάθε μοντέλου.



Εικόνα 3 Σύγκριση καμπυλών ROC

ΚΕΦΑΛΑΙΟ 5: ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

5.1 Λογιστική Παλινδρόμηση

Στην ενότητα 5.1 εξετάζεται η Λογιστική Παλινδρόμηση ως βασικό γραμμικό μοντέλο αναφοράς (baseline) για την πρόβλεψη της αθέτησης πληρωμής πιστωτικής κάρτας. Το μοντέλο αυτό επιλέγεται λόγω της απλότητας και της ερμηνευσιμότητάς του, παρέχοντας ένα σημείο σύγκρισης για την αξιολόγηση πιο σύνθετων, μη γραμμικών προσεγγίσεων που ακολουθούν (Thomas, 2000; Wahab et al., 2024). Αρχικά παρουσιάζεται η διαδικασία εκπαίδευσης και προεπεξεργασίας των δεδομένων, στη συνέχεια αναλύονται τα ποσοτικά αποτελέσματα και τέλος σχολιάζεται η απόδοση του μοντέλου με έμφαση στους περιορισμούς του ως προς την ανίχνευση της κλάσης της αθέτησης.

5.1.1 Περιγραφή και Εκπαίδευση Λογιστικής Παλινδρόμησης

Η Λογιστική Παλινδρόμηση χρησιμοποιήθηκε ως μοντέλο αναφοράς (baseline) για την πρόβλεψη της αθέτησης πληρωμής πιστωτικής κάρτας. Η επιλογή του συγκεκριμένου μοντέλου επιτρέπει την αξιολόγηση της απόδοσης ενός γραμμικού ταξινομητή σε ένα πρόβλημα που, όπως προέκυψε από την εξερευνητική ανάλυση δεδομένων, χαρακτηρίζεται από σύνθετες και μη γραμμικές σχέσεις μεταξύ των μεταβλητών (Thomas, 2000).

Η εκπαίδευση του μοντέλου πραγματοποιήθηκε στο σύνολο εκπαίδευσης, το οποίο προέκυψε από στρωματοποιημένο διαχωρισμό του αρχικού dataset σε ποσοστό 80% για εκπαίδευση και 20% για έλεγχο. Η προεπεξεργασία των δεδομένων ενσωματώθηκε σε ενιαίο pipeline μαζί με το μοντέλο, ώστε να διασφαλιστεί η ορθή εφαρμογή των ίδιων μετασχηματισμών τόσο στα δεδομένα εκπαίδευσης όσο και στα δεδομένα ελέγχου και να αποφευχθεί η διαρροή πληροφορίας (data leakage).

Οι αριθμητικές μεταβλητές κανονικοποιήθηκαν μέσω τυποποίησης (StandardScaler), ενώ οι κατηγορικές μεταβλητές κωδικοποιήθηκαν με τη μέθοδο One-Hot Encoding. Για την υλοποίηση του μοντέλου χρησιμοποιήθηκε η κλάση LogisticRegression της βιβλιοθήκης scikit-learn με κανονικοποίηση L2, η οποία συμβάλλει στον περιορισμό της υπερεκμάθησης. Ο μέγιστος αριθμός επαναλήψεων ορίστηκε σε 2000, ώστε να διασφαλιστεί η σύγκλιση της διαδικασίας βελτιστοποίησης λόγω του αυξημένου αριθμού χαρακτηριστικών μετά την κωδικοποίηση των κατηγορικών μεταβλητών. Η έξοδος του μοντέλου ερμηνεύεται ως πιθανότητα αθέτησης πληρωμής, ενώ η τελική ταξινόμηση προκύπτει με την εφαρμογή κατωφλίου απόφασης ίσου με 0.5.

5.1.2 Ποσοτικά Αποτελέσματα Λογιστικής Παλινδρόμησης

Η απόδοση της Λογιστικής Παλινδρόμησης αξιολογήθηκε στο σύνολο ελέγχου με τη χρήση πολλαπλών μετρικών ταξινόμησης, προκειμένου να αποτυπωθεί με πληρότητα η ικανότητά της να διαχωρίζει τις περιπτώσεις αθέτησης από τις περιπτώσεις μη αθέτησης πληρωμής. Στον Πίνακα 3 παρουσιάζονται τα βασικά ποσοτικά αποτελέσματα του μοντέλου, ενώ η αντίστοιχη μήτρα σύγχυσης απεικονίζεται στην Εικόνα 6 (Wahab et al., 2024).

Συγκεκριμένα, η Λογιστική Παλινδρόμηση πέτυχε συνολική ακρίβεια (accuracy) ίση με 0.8088. Ωστόσο, η ανάλυση των επιμέρους μετρικών ανά κλάση δείχνει σημαντικές διαφοροποιήσεις μεταξύ της πλειοψηφικής κλάσης των μη αθετήσεων και της μειοψηφικής

κλάσης της αθέτησης. Για την κλάση των μη αθετήσεων (κλάση 0), το μοντέλο παρουσίασε υψηλή ευαισθησία ($\text{recall} \approx 0.97$), γεγονός που υποδηλώνει ότι αναγνωρίζει με μεγάλη επιτυχία τους πελάτες που δεν θα αθετήσουν την πληρωμή τους.

Αντιθέτως, για την κλάση της αθέτησης (κλάση 1), η απόδοση του μοντέλου ήταν σαφώς χαμηλότερη. Συγκεκριμένα, το recall για την κλάση της αθέτησης ανήλθε σε περίπου 0.24, γεγονός που σημαίνει ότι το μοντέλο εντοπίζει λιγότερο από το ένα τέταρτο των πραγματικών περιπτώσεων αθέτησης. Το precision για την ίδια κλάση διαμορφώθηκε σε περίπου 0.69, υποδηλώνοντας ότι όταν το μοντέλο προβλέπει αθέτηση, η πρόβλεψη είναι σχετικά αξιόπιστη, αλλά πραγματοποιείται σε περιορισμένο αριθμό περιπτώσεων (Thomas, 2000)..

Το F1-score για την κλάση της αθέτησης ανήλθε σε περίπου 0.36, αντικατοπτρίζοντας τον συμβιβασμό μεταξύ χαμηλής ευαισθησίας και μέτριας ακρίβειας των θετικών προβλέψεων. Τέλος, η συνολική διακριτική ικανότητα του μοντέλου, όπως αποτυπώνεται από τον δείκτη ROC-AUC, ήταν ίση με 0.71, τιμή που υποδηλώνει μέτρια ικανότητα διάκρισης μεταξύ των δύο κλάσεων.

Τα παραπάνω αποτελέσματα αναδεικνύουν τους περιορισμούς της Λογιστικής Παλινδρόμησης ως γραμμικού μοντέλου στην αποτελεσματική ανίχνευση της μειοψηφικής κλάσης της αθέτησης, γεγονός που καθιστά αναγκαία τη διερεύνηση πιο ευέλικτων μη γραμμικών προσεγγίσεων (Thomas, 2000).



Εικόνα 4 Πίνακας σύγχυσης Λογιστικής Παλινδρόμησης

Πίνακας 3 Μετρικές απόδοσης Λογιστικής Παλινδρόμησης ανά κλάση

Μετρική	Κλάση 0	Κλάση 1
Precision	0.81	0.69
Recall	0.97	0.24
F1-score	0.89	0.36

Πίνακας 4 Συνολικές μετρικές απόδοσης Λογιστικής Παλινδρόμησης

Μετρική	Τιμή
Accuracy	0.8088
ROC-AUC	0.71

5.1.3 Ερμηνεία και Σχολιασμός Αποτελεσμάτων

Τα αποτελέσματα της Λογιστικής Παλινδρόμησης δείχνουν ότι το μοντέλο επιτυγχάνει ικανοποιητική συνολική ακρίβεια, ωστόσο παρουσιάζει σημαντικούς περιορισμούς ως προς την ανίχνευση της κλάσης της αθέτησης πληρωμής. Συγκεκριμένα, η τιμή της ακρίβειας (accuracy) είναι σχετικά υψηλή, γεγονός που οφείλεται κυρίως στη σωστή ταξινόμηση της πλειοψηφικής κλάσης των μη αθετήσεων. Το εύρημα αυτό επιβεβαιώνει ότι, σε συνθήκες έντονης ανισορροπίας κλάσεων, η ακρίβεια από μόνη της δεν αποτελεί επαρκές κριτήριο αξιολόγησης της απόδοσης ενός μοντέλου (Wahab et al., 2024).

Η ανάλυση των μετρικών για την κλάση της αθέτησης αποκαλύπτει ότι το μοντέλο παρουσιάζει χαμηλή ευαισθησία (recall $\approx 24\%$), γεγονός που σημαίνει ότι εντοπίζει μόνο ένα μικρό ποσοστό των πραγματικών περιπτώσεων αθέτησης. Αντίθετα, η τιμή του precision για την ίδια κλάση είναι σχετικά υψηλή, υποδηλώνοντας ότι όταν το μοντέλο προβλέπει αθέτηση, η πρόβλεψη αυτή είναι συχνά σωστή. Το μοτίβο αυτό δείχνει ότι η Λογιστική Παλινδρόμηση τείνει να είναι συντηρητική στις προβλέψεις της για τη θετική κλάση, αποφεύγοντας τις ψευδώς θετικές προβλέψεις, αλλά με κόστος τη μη ανίχνευση σημαντικού αριθμού επισφαλών πελατών.

Το χαμηλό F1-score για την κλάση της αθέτησης αντανακλά τον ανωτέρω συμβιβασμό μεταξύ precision και recall και επιβεβαιώνει ότι το μοντέλο δεν επιτυγχάνει ισορροπημένη απόδοση ως προς την ανίχνευση του πιστωτικού κινδύνου. Από επιχειρησιακή σκοπιά, η αδυναμία αυτή είναι ιδιαίτερα κρίσιμη, καθώς το κόστος ενός ψευδώς αρνητικού σφάλματος (false negative), δηλαδή η μη αναγνώριση ενός πελάτη που τελικά θα αθετήσει την πληρωμή του, είναι υψηλό για ένα χρηματοπιστωτικό ίδρυμα (Thomas, 2000).

Η συνολική διακριτική ικανότητα του μοντέλου, όπως αποτυπώνεται από την τιμή του ROC-AUC (≈ 0.71), δείχνει ότι η Λογιστική Παλινδρόμηση αποδίδει καλύτερα από έναν τυχαίο ταξινομητή, αλλά παρουσιάζει περιορισμένη ικανότητα διαχωρισμού των δύο κλάσεων. Το αποτέλεσμα αυτό συνάδει με τη γραμμική φύση του μοντέλου και με τα ευρήματα της

εξερευνητικής ανάλυσης δεδομένων, σύμφωνα με τα οποία η μεταβλητή-στόχος δεν εμφανίζει ισχυρές γραμμικές συσχετίσεις με μεμονωμένα χαρακτηριστικά (Wahab et al., 2024)..

Συνολικά, η Λογιστική Παλινδρόμηση μπορεί να θεωρηθεί ως ένα σταθερό και ερμηνεύσιμο μοντέλο αναφοράς (baseline), το οποίο παρέχει μια αρχική εικόνα της δυσκολίας του προβλήματος. Ωστόσο, τα αποτελέσματά της υποδεικνύουν ότι η πρόβλεψη της αθέτησης πληρωμής δεν μπορεί να περιγραφεί επαρκώς μέσω απλών γραμμικών σχέσεων, γεγονός που καθιστά αναγκαία τη διερεύνηση πιο ευέλικτων και μη γραμμικών μοντέλων, όπως τα Δέντρα Απόφασης, τα οποία εξετάζονται στην επόμενη ενότητα (Thomas, 2000).

5.2 Εκπαίδευση και Αξιολόγηση Δέντρου Απόφασης

Στην ενότητα 5.2 εξετάζεται το Δέντρο Απόφασης ως δεύτερο μοντέλο ταξινόμησης για την πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας. Σε αντίθεση με τη Λογιστική Παλινδρόμηση, το Δέντρο Απόφασης μπορεί να αποτυπώσει μη γραμμικές σχέσεις και αλληλεπιδράσεις μεταξύ μεταβλητών, γεγονός ιδιαίτερα χρήσιμο σε προβλήματα πιστωτικού κινδύνου. Αρχικά παρουσιάζεται συνοπτικά η διαδικασία εκπαίδευσης και οι βασικές παράμετροι του μοντέλου, στη συνέχεια παρατίθενται τα ποσοτικά αποτελέσματα αξιολόγησης, και τέλος γίνεται ερμηνεία των ευρημάτων με έμφαση στη συμπεριφορά του μοντέλου ως προς την κλάση της αθέτησης (Wahab et al., 2024).

5.2.1 Περιγραφή και Εκπαίδευση Δέντρου Απόφασης

Στο πλαίσιο της παρούσας μελέτης, ως δεύτερο μοντέλο ταξινόμησης επιλέχθηκε το Δέντρο Απόφασης (Decision Tree), με στόχο τη διερεύνηση της απόδοσης ενός μη γραμμικού μοντέλου στην πρόβλεψη της αθέτησης πληρωμής πιστωτικής κάρτας. Το Δέντρο Απόφασης αποτελεί ένα ευέλικτο μοντέλο μηχανικής μάθησης, το οποίο διαχωρίζει τον χώρο των δεδομένων μέσω μιας ακολουθίας ιεραρχικών κανόνων απόφασης, βασισμένων στις τιμές των χαρακτηριστικών (Wahab et al., 2024).

Σε αντίθεση με τη Λογιστική Παλινδρόμηση, το Δέντρο Απόφασης δεν προϋποθέτει γραμμικές σχέσεις μεταξύ των μεταβλητών και είναι σε θέση να συλλάβει σύνθετες και μη γραμμικές αλληλεπιδράσεις, γεγονός ιδιαίτερα σημαντικό για προβλήματα πιστωτικού κινδύνου, όπου η συμπεριφορά των πελατών εξαρτάται από συνδυασμούς πολλαπλών παραγόντων.

Η εκπαίδευση του μοντέλου πραγματοποιήθηκε στο ίδιο σύνολο εκπαίδευσης που χρησιμοποιήθηκε και για τη Λογιστική Παλινδρόμηση, ώστε να διασφαλιστεί η συγκρισιμότητα των αποτελεσμάτων. Για την αποφυγή υπερπροσαρμογής (overfitting), εφαρμόστηκαν περιορισμοί στη δομή του δέντρου, και συγκεκριμένα ορίστηκε μέγιστο βάθος δέντρου ίσο με 6 (`max_depth = 6`) και ελάχιστος αριθμός παρατηρήσεων ανά φύλλο ίσος με 50 (`min_samples_leaf = 50`). Οι παράμετροι αυτές επιλέχθηκαν με στόχο την εξισορρόπηση μεταξύ εκφραστικής ικανότητας του μοντέλου και γενίκευσης σε νέα δεδομένα.

Επιπλέον, δεδομένης της ανισορροπίας των κλάσεων στο σύνολο δεδομένων, εφαρμόστηκε στάθμιση κλάσεων (`class_weight = "balanced"`), ώστε να δοθεί μεγαλύτερη σημασία στη σωστή ταξινόμηση της μειοψηφικής κλάσης της αθέτησης πληρωμής. Με τον τρόπο αυτό, το μοντέλο ενθαρρύνθηκε να λαμβάνει υπόψη το κόστος των σφαλμάτων ταξινόμησης και να βελτιώσει την ικανότητά του να εντοπίζει επισφαλείς πελάτες (Wahab et al., 2024).

Η διαδικασία προεπεξεργασίας των δεδομένων και εκπαίδευσης του μοντέλου ενσωματώθηκε σε ενιαίο pipeline, εξασφαλίζοντας ότι οι ίδιοι μετασχηματισμοί εφαρμόστηκαν με συνεπή τρόπο τόσο στα δεδομένα εκπαίδευσης όσο και στα δεδομένα ελέγχου, χωρίς διαρροή πληροφορίας μεταξύ των συνόλων.

5.2.2 Ποσοτικά Αποτελέσματα Δέντρου Απόφασης

Η απόδοση του μοντέλου Δέντρου Απόφασης αξιολογήθηκε στο σύνολο ελέγχου (test set) με τη χρήση των μετρικών Accuracy, Precision, Recall, F1-score και ROC-AUC, ώστε να αποτυπωθεί με πληρότητα η συμπεριφορά του τόσο στη συνολική ταξινόμηση όσο και ειδικότερα στην κλάση της αθέτησης πληρωμής (Wahab et al., 2024).

Τα ποσοτικά αποτελέσματα δείχνουν ότι το Δέντρο Απόφασης πέτυχε συνολική ακρίβεια (accuracy) ίση με 0.7812. Παρότι η τιμή αυτή είναι ελαφρώς χαμηλότερη σε σύγκριση με τη Λογιστική Παλινδρόμηση, παρέχει μια πιο ρεαλιστική εικόνα της απόδοσης του μοντέλου σε συνθήκες ανισορροπίας κλάσεων.

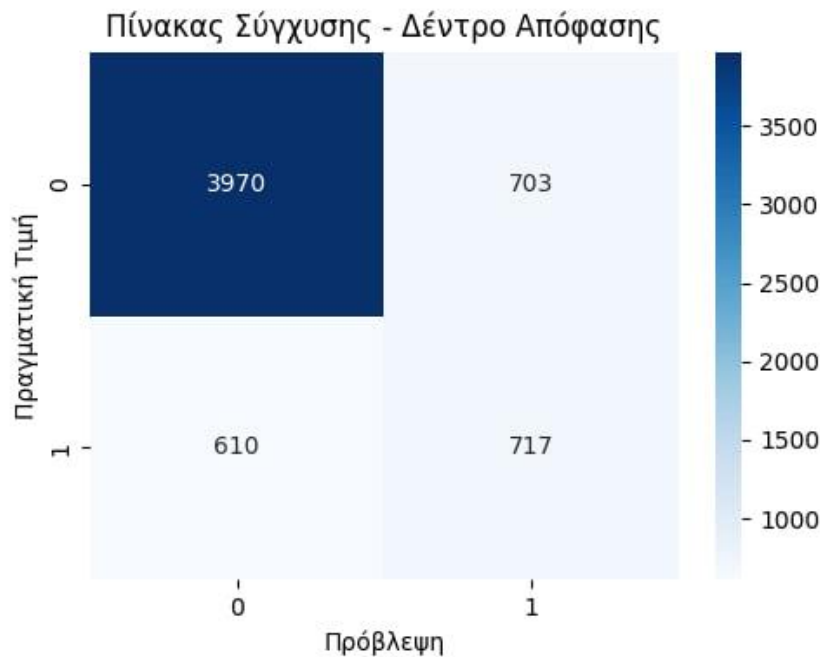
Όσον αφορά την κλάση της αθέτησης πληρωμής, το μοντέλο παρουσίασε precision περίπου ίσο με 0.505 και recall περίπου ίσο με 0.540. Τα αποτελέσματα αυτά υποδηλώνουν ότι το Δέντρο Απόφασης καταφέρνει να εντοπίζει πάνω από τις μισές πραγματικές περιπτώσεις αθέτησης, ενώ ταυτόχρονα διατηρεί ικανοποιητικό επίπεδο ακρίβειας στις προβλέψεις της θετικής κλάσης. Το αντίστοιχο F1-score για την κλάση της αθέτησης ανέρχεται περίπου σε 0.522, αντανakλώντας μια πιο ισορροπημένη απόδοση σε σχέση με το γραμμικό μοντέλο.

Η συνολική διακριτική ικανότητα του μοντέλου, όπως αποτυπώνεται από την τιμή του δείκτη ROC-AUC, είναι ίση με 0.7651. Η τιμή αυτή υποδηλώνει καλή ικανότητα διαχωρισμού μεταξύ των δύο κλάσεων και επιβεβαιώνει ότι το Δέντρο Απόφασης υπερέχει σαφώς του τυχαίου ταξινομητή.

Τα αποτελέσματα της ταξινόμησης αποτυπώνονται αναλυτικά στον πίνακα αποτελεσμάτων και στη μήτρα σύγχυσης (confusion matrix), όπου παρουσιάζεται η κατανομή των σωστών και λανθασμένων προβλέψεων για κάθε κλάση. Επιπλέον, η καμπύλη ROC του μοντέλου απεικονίζει γραφικά τη σχέση μεταξύ του ρυθμού αληθώς θετικών και ψευδώς θετικών προβλέψεων για διαφορετικά κατώφλια απόφασης.

Πίνακας 5 Ποσοτικά Αποτελέσματα Δέντρου Απόφασης

Μετρική	Τιμή
Accuracy	0.7812
Precision (κλάση Default = 1)	0.5049
Recall (κλάση Default = 1)	0.5403
F1-score (κλάση Default = 1)	0.5220
ROC-AUC	0.7651



Εικόνα 5 Πίνακας σύγκρισης Δέντρου Απόφασης

5.2.3 Ερμηνεία και Σχολιασμός Αποτελεσμάτων

Τα αποτελέσματα του Δέντρου Απόφασης δείχνουν ότι το μοντέλο παρουσιάζει πιο ισορροπημένη συμπεριφορά ως προς την ανίχνευση της κλάσης της αθέτησης πληρωμής σε σύγκριση με τη Λογιστική Παλινδρόμηση. Παρότι η συνολική ακρίβεια του μοντέλου είναι ελαφρώς χαμηλότερη, η απόδοσή του ως προς τις μετρικές που σχετίζονται άμεσα με τον πιστωτικό κίνδυνο είναι σαφώς βελτιωμένη.

Ειδικότερα, το Δέντρο Απόφασης εμφανίζει σημαντικά υψηλότερη ευαισθησία ($\text{recall} \approx 54\%$) για την κλάση της αθέτησης, γεγονός που υποδηλώνει ότι το μοντέλο καταφέρνει να εντοπίσει περισσότερες από τις μισές πραγματικές περιπτώσεις πελατών που τελικά αθετούν την πληρωμή τους. Το εύρημα αυτό είναι ιδιαίτερα κρίσιμο σε εφαρμογές διαχείρισης πιστωτικού κινδύνου, όπου το κόστος μη ανίχνευσης ενός επισφαλούς πελάτη (false negative) είναι σημαντικά υψηλότερο από το κόστος ενός ψευδώς θετικού σφάλματος (Thomas, 2000).

Παράλληλα, η τιμή του precision για την κλάση της αθέτησης ($\approx 50\%$) δείχνει ότι περίπου οι μισές από τις προβλέψεις αθέτησης του μοντέλου είναι ορθές. Αν και το ποσοστό αυτό υποδηλώνει την ύπαρξη ψευδών θετικών προβλέψεων, ο συγκεκριμένος συμβιβασμός θεωρείται αποδεκτός στο πλαίσιο της παρούσας μελέτης, δεδομένου ότι η έγκαιρη αναγνώριση επισφαλών πελατών έχει μεγαλύτερη επιχειρησιακή σημασία από την αποφυγή πρόσθετων ελέγχων σε πελάτες χαμηλότερου κινδύνου (Thomas, 2000).

Το F1-score του μοντέλου, το οποίο ανέρχεται περίπου σε 0.522, επιβεβαιώνει τη βελτιωμένη ισορροπία μεταξύ precision και recall σε σχέση με τη Λογιστική Παλινδρόμηση. Το αποτέλεσμα αυτό υποδηλώνει ότι το Δέντρο Απόφασης επιτυγχάνει πιο αποτελεσματική

συνολική απόδοση ως προς την ανίχνευση της αθέτησης πληρωμής, χωρίς να θυσιάζει υπερβολικά την ακρίβεια των θετικών προβλέψεων.

Η συνολική διακριτική ικανότητα του μοντέλου, όπως αποτυπώνεται από την τιμή του δείκτη ROC-AUC (≈ 0.765), καταδεικνύει ότι το Δέντρο Απόφασης έχει καλύτερη ικανότητα διαχωρισμού των δύο κλάσεων σε σύγκριση με το γραμμικό μοντέλο. Το εύρημα αυτό συνάδει με τη μη γραμμική φύση του μοντέλου και επιβεβαιώνει ότι η πρόβλεψη της αθέτησης πληρωμής επηρεάζεται από σύνθετες αλληλεπιδράσεις μεταξύ των μεταβλητών, οι οποίες δεν μπορούν να αποτυπωθούν επαρκώς μέσω γραμμικών υποθέσεων (Wahab et al., 2024)..

Συνολικά, το Δέντρο Απόφασης αναδεικνύεται ως πιο κατάλληλο μοντέλο για το συγκεκριμένο πρόβλημα, ιδίως όταν το ζητούμενο είναι η αποτελεσματική ανίχνευση επισφαλών πελατών. Τα αποτελέσματά του δικαιολογούν τη χρήση του ως βασικού μοντέλου πρόβλεψης πιστωτικού κινδύνου στο πλαίσιο της παρούσας μελέτης και προετοιμάζουν το έδαφος για τη συγκριτική αξιολόγηση των δύο μοντέλων που ακολουθεί στην επόμενη ενότητα.

5.3 Σύγκριση Μοντέλων

Στην ενότητα 5.3 πραγματοποιείται η συνολική σύγκριση των δύο μοντέλων που εξετάστηκαν στην παρούσα μελέτη, της Λογιστικής Παλινδρόμησης και του Δέντρου Απόφασης, με στόχο την αξιολόγηση της καταλληλότητάς τους για την πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας. Η σύγκριση δεν περιορίζεται μόνο στις ποσοτικές μετρικές απόδοσης, αλλά επεκτείνεται και σε ποιοτικά χαρακτηριστικά των μοντέλων, καθώς και στην ανάλυση των σφαλμάτων ταξινόμησης και των περιορισμών τους. Με τον τρόπο αυτό επιδιώκεται μια πιο σφαιρική και ουσιαστική αποτίμηση των πλεονεκτημάτων και των αδυναμιών κάθε προσέγγισης στο πλαίσιο της διαχείρισης πιστωτικού κινδύνου (Thomas, 2000).

5.3.1 Ποσοτική Σύγκριση

Για τη συγκριτική αξιολόγηση των δύο μοντέλων, της Λογιστικής Παλινδρόμησης και του Δέντρου Απόφασης, χρησιμοποιήθηκαν οι ίδιες μετρικές απόδοσης, ώστε τα αποτελέσματα να είναι άμεσα συγκρίσιμα. Η σύγκριση πραγματοποιήθηκε με βάση τη συνολική ακρίβεια (accuracy), καθώς και μετρικές που εστιάζουν στην κλάση της αθέτησης πληρωμής, όπως το precision, το recall, το F1-score και ο δείκτης ROC-AUC.

Από τα ποσοτικά αποτελέσματα προκύπτει ότι η Λογιστική Παλινδρόμηση παρουσιάζει ελαφρώς υψηλότερη συνολική ακρίβεια σε σχέση με το Δέντρο Απόφασης. Ωστόσο, η διαφορά αυτή οφείλεται κυρίως στην καλύτερη ταξινόμηση της πλειοψηφικής κλάσης των μη αθετήσεων και δεν αντικατοπτρίζει την ικανότητα των μοντέλων να εντοπίζουν αποτελεσματικά τις πραγματικές περιπτώσεις αθέτησης.

Αντίθετα, το Δέντρο Απόφασης υπερέχει σαφώς ως προς την ευαισθησία (recall) για την κλάση της αθέτησης, επιτυγχάνοντας τιμή περίπου 54%, έναντι περίπου 24% της Λογιστικής Παλινδρόμησης. Το εύρημα αυτό υποδηλώνει ότι το μη γραμμικό μοντέλο είναι σημαντικά πιο αποτελεσματικό στον εντοπισμό επισφαλών πελατών. Παράλληλα, το F1-score του Δέντρου Απόφασης είναι αισθητά υψηλότερο, γεγονός που επιβεβαιώνει την πιο ισορροπημένη απόδοσή του ως προς την ανίχνευση της μειοψηφικής κλάσης.

Όσον αφορά τη συνολική διακριτική ικανότητα των μοντέλων, ο δείκτης ROC-AUC ενισχύει περαιτέρω τα παραπάνω συμπεράσματα. Συγκεκριμένα, το Δέντρο Απόφασης πέτυχε τιμή

ROC-AUC ίση με 0.765, ενώ η Λογιστική Παλινδρόμηση περιορίστηκε σε τιμή 0.71. Η διαφορά αυτή υποδηλώνει ότι το Δέντρο Απόφασης διαθέτει καλύτερη ικανότητα διαχωρισμού μεταξύ πελατών που θα αθετήσουν και αυτών που δεν θα αθετήσουν την πληρωμή τους, ανεξάρτητα από το επιλεγμένο κατώφλι ταξινόμησης.

Συνοψίζοντας, η ποσοτική σύγκριση δείχνει ότι, παρότι η Λογιστική Παλινδρόμηση εμφανίζει υψηλότερη συνολική ακρίβεια, το Δέντρο Απόφασης υπερέχει στις μετρικές που είναι κρίσιμες για την πρόβλεψη πιστωτικού κινδύνου. Το αποτέλεσμα αυτό καθιστά το Δέντρο Απόφασης πιο κατάλληλο μοντέλο για εφαρμογές όπου η έγκαιρη ανίχνευση αθέτησης πληρωμής αποτελεί πρωταρχικό στόχο (Wahab et al., 2024).

Πίνακας 6 Σύγκριση βασικών μετρικών μεταξύ Λογιστικής Παλινδρόμησης και Δέντρο Αποφάσεων

Μοντέλο	Accuracy	Precision (Default=1)	Recall (Default=1)	F1-score (Default=1)	ROC-AUC
Λογιστική Παλινδρόμηση	0.8088	0.6923	0.2442	0.3610	0.7100
Δέντρο Απόφασης	0.7812	0.5049	0.5403	0.5220	0.7651

5.3.2 Ποιοτική Σύγκριση

Πέραν της ποσοτικής αξιολόγησης μέσω μετρικών απόδοσης, κρίνεται σκόπιμη και η ποιοτική σύγκριση των δύο μοντέλων, λαμβάνοντας υπόψη χαρακτηριστικά όπως η ερμηνευσιμότητα, η ευελιξία, η ικανότητα σύλληψης μη γραμμικών σχέσεων και η καταλληλότητά τους για εφαρμογές πιστωτικού κινδύνου (Thomas, 2000).

Η Λογιστική Παλινδρόμηση αποτελεί ένα γραμμικό μοντέλο με υψηλό βαθμό ερμηνευσιμότητας, καθώς οι συντελεστές του μοντέλου επιτρέπουν την άμεση κατανόηση της κατεύθυνσης και της σχετικής επίδρασης κάθε μεταβλητής στην πιθανότητα αθέτησης. Το χαρακτηριστικό αυτό καθιστά τη Λογιστική Παλινδρόμηση ιδιαίτερα χρήσιμη σε περιβάλλοντα όπου απαιτείται διαφάνεια και αιτιολόγηση των αποφάσεων, όπως σε ρυθμιστικά πλαίσια ή σε διαδικασίες συμμόρφωσης. Ωστόσο, η γραμμική της φύση περιορίζει την ικανότητά της να αποτυπώσει πολύπλοκες και μη γραμμικές σχέσεις μεταξύ των μεταβλητών, γεγονός που αντανακλάται και στη χαμηλή ευαισθησία της ως προς την ανίχνευση των πραγματικών περιπτώσεων αθέτησης.

Αντίθετα, το Δέντρο Απόφασης αποτελεί ένα μη γραμμικό μοντέλο, το οποίο μπορεί να συλλάβει σύνθετες αλληλεπιδράσεις μεταξύ των χαρακτηριστικών χωρίς να απαιτεί γραμμικές υποθέσεις. Μέσω της ιεραρχικής δομής διαδοχικών κανόνων απόφασης, το μοντέλο αυτό προσαρμόζεται καλύτερα στην πολυπαραγοντική φύση του πιστωτικού κινδύνου. Επιπλέον, παρά την αυξημένη ευελιξία του, το Δέντρο Απόφασης διατηρεί σχετικά υψηλό επίπεδο ερμηνευσιμότητας, καθώς η λογική των αποφάσεων μπορεί να αναλυθεί μέσω των κόμβων και των διαχωρισμών του δέντρου. Η αυξημένη ικανότητά του να εντοπίζει περιπτώσεις αθέτησης το καθιστά ιδιαίτερα κατάλληλο για εφαρμογές διαχείρισης κινδύνου, όπου η έγκαιρη αναγνώριση επισφαλών πελατών έχει κρίσιμη σημασία.

Συνολικά, η ποιοτική σύγκριση δείχνει ότι τα δύο μοντέλα εξυπηρετούν διαφορετικές ανάγκες: η Λογιστική Παλινδρόμηση προσφέρει απλότητα, σταθερότητα και διαφάνεια, ενώ το Δέντρο Απόφασης παρέχει μεγαλύτερη εκφραστική ισχύ και αυξημένη αποτελεσματικότητα στην ανίχνευση της μειοψηφικής κλάσης της αθέτησης. Η συμπληρωματική χρήση τους επιτρέπει μια πιο σφαιρική κατανόηση του προβλήματος και ενισχύει την αξιοπιστία των συμπερασμάτων της παρούσας μελέτης.

5.3.3 Ανάλυση Σφαλμάτων: Κόστος False Positives και False Negatives

Στα προβλήματα πρόβλεψης πιστωτικού κινδύνου, η αξιολόγηση των μοντέλων δεν περιορίζεται αποκλειστικά στις συνολικές μετρικές απόδοσης, αλλά επεκτείνεται στην ανάλυση των σφαλμάτων ταξινόμησης και στο σχετικό κόστος που αυτά συνεπάγονται. Ιδιαίτερη σημασία έχουν τα σφάλματα τύπου false positive και false negative, καθώς επηρεάζουν άμεσα τις αποφάσεις διαχείρισης κινδύνου ενός χρηματοπιστωτικού ιδρύματος.

Ένα false positive αντιστοιχεί στην περίπτωση όπου το μοντέλο προβλέπει ότι ένας πελάτης θα αθετήσει την πληρωμή του (default), ενώ στην πραγματικότητα ο πελάτης είναι φερέγγυος. Το κόστος ενός τέτοιου σφάλματος αφορά κυρίως την απώλεια δυνητικών εσόδων, καθώς ένας αξιόπιστος πελάτης μπορεί να απορριφθεί ή να λάβει δυσμενέστερους όρους χρηματοδότησης. Αν και το κόστος αυτό δεν είναι αμελητέο, θεωρείται συνήθως χαμηλότερο σε σχέση με το κόστος των σφαλμάτων false negative (Thomas, 2000).

Αντίθετα, ένα false negative προκύπτει όταν το μοντέλο αποτυγχάνει να εντοπίσει έναν πραγματικά επισφαλής πελάτη, προβλέποντας λανθασμένα ότι δεν θα υπάρξει αθέτηση πληρωμής. Το σφάλμα αυτό έχει ιδιαίτερα υψηλό κόστος, καθώς μπορεί να οδηγήσει σε πραγματικές οικονομικές απώλειες για τον οργανισμό, αυξημένες προβλέψεις επισφαλών απαιτήσεων και επιβάρυνση της συνολικής πιστωτικής έκθεσης. Για τον λόγο αυτό, στα συστήματα πιστωτικής αξιολόγησης δίνεται συχνά μεγαλύτερη έμφαση στη μείωση των false negatives, ακόμη και εις βάρος μιας αύξησης των false positives.

Η ανάλυση των πινάκων σύγχυσης έδειξε ότι η Λογιστική Παλινδρόμηση παρουσίασε σημαντικά υψηλότερο αριθμό false negatives σε σύγκριση με το Δέντρο Απόφασης, γεγονός που αντανακλάται και στη χαμηλή τιμή της ευαισθησίας (recall) για την κλάση της αθέτησης. Αντιθέτως, το Δέντρο Απόφασης κατάφερε να εντοπίσει μεγαλύτερο ποσοστό των πραγματικών περιπτώσεων αθέτησης, μειώνοντας τον αριθμό των false negatives, αν και με το τίμημα ενός αυξημένου αριθμού false positives.

Η συμπεριφορά αυτή υποδηλώνει ότι το Δέντρο Απόφασης είναι περισσότερο κατάλληλο για εφαρμογές όπου η έγκαιρη ανίχνευση επισφαλών πελατών αποτελεί προτεραιότητα και το κόστος μη ανίχνευσης μιας αθέτησης είναι υψηλό. Αντίθετα, η Λογιστική Παλινδρόμηση εμφανίζεται πιο συντηρητική στις προβλέψεις αθέτησης, γεγονός που οδηγεί σε λιγότερες ψευδείς απορρίψεις φερέγγυων πελατών, αλλά αυξάνει τον κίνδυνο αποδοχής επισφαλών περιπτώσεων.

Συνεπώς, η επιλογή του κατάλληλου μοντέλου εξαρτάται σε μεγάλο βαθμό από τη στρατηγική διαχείρισης κινδύνου και τη σχετική στάθμιση του κόστους των διαφορετικών τύπων σφαλμάτων. Στο πλαίσιο της παρούσας μελέτης, όπου η πρόβλεψη της αθέτησης πληρωμής

αποτελεί κρίσιμο στόχο, το Δέντρο Απόφασης εμφανίζεται ως πιο κατάλληλη επιλογή λόγω της μειωμένης πιθανότητας εμφάνισης σφαλμάτων false negative.

5.3.4 Περιορισμοί και Πηγές Μεροληψίας

Παρά τα ικανοποιητικά αποτελέσματα που προέκυψαν από την εκπαίδευση και αξιολόγηση των μοντέλων, η παρούσα μελέτη υπόκειται σε ορισμένους περιορισμούς και πιθανές πηγές μεροληψίας, οι οποίοι πρέπει να ληφθούν υπόψη κατά την ερμηνεία των αποτελεσμάτων.

Αρχικά, το σύνολο δεδομένων που χρησιμοποιήθηκε προέρχεται από συγκεκριμένο χρονικό και γεωγραφικό πλαίσιο και αφορά πελάτες πιστωτικών καρτών συγκεκριμένου χρηματοπιστωτικού ιδρύματος. Ως εκ τούτου, τα μοτίβα συμπεριφοράς που εντοπίστηκαν ενδέχεται να μην γενικεύονται πλήρως σε διαφορετικά οικονομικά περιβάλλοντα, χρονικές περιόδους ή πληθυσμούς με διαφορετικά δημογραφικά και κοινωνικοοικονομικά χαρακτηριστικά.

Ένας ακόμη σημαντικός περιορισμός αφορά την έντονη ανισορροπία των κλάσεων στη μεταβλητή-στόχο. Παρότι εφαρμόστηκαν κατάλληλες μετρικές αξιολόγησης και τεχνικές στάθμισης κλάσεων, η ανισορροπία αυτή ενδέχεται να επηρεάζει τη σταθερότητα και τη γενικευσιμότητα των μοντέλων, ιδίως σε περιπτώσεις όπου η κατανομή των κλάσεων μεταβάλλεται στο χρόνο (concept drift).

Επιπλέον, ορισμένες μεταβλητές του dataset είναι κωδικοποιημένες αριθμητικά, παρότι αντιπροσωπεύουν κατηγορικά χαρακτηριστικά (π.χ. επίπεδο εκπαίδευσης, οικογενειακή κατάσταση). Η αριθμητική αυτή αναπαράσταση μπορεί να εισάγει έμμεσες υποθέσεις τάξης ή απόστασης μεταξύ των κατηγοριών, οι οποίες δεν αντικατοπτρίζουν απαραίτητα την πραγματική σημασιολογική τους σχέση, επηρεάζοντας ενδεχομένως τη συμπεριφορά των μοντέλων.

Όσον αφορά τα ίδια τα μοντέλα, η Λογιστική Παλινδρόμηση βασίζεται στην υπόθεση γραμμικού διαχωρισμού των κλάσεων και ενδέχεται να μην αποτυπώνει πλήρως σύνθετες και μη γραμμικές σχέσεις μεταξύ των μεταβλητών. Από την άλλη πλευρά, το Δέντρο Απόφασης, παρότι πιο ευέλικτο, είναι ευαίσθητο στις επιλογές υπερπαραμέτρων και μπορεί να παρουσιάσει φαινόμενα υπερεκμάθησης, ακόμη και όταν εφαρμόζονται τεχνικές περιορισμού της πολυπλοκότητάς του.

Τέλος, η αξιολόγηση των μοντέλων βασίστηκε σε στατικό διαχωρισμό του dataset σε σύνολα εκπαίδευσης και ελέγχου. Αν και η προσέγγιση αυτή είναι ευρέως αποδεκτή, δεν λαμβάνει υπόψη δυναμικές αλλαγές στη συμπεριφορά των πελατών με την πάροδο του χρόνου, οι οποίες είναι ιδιαίτερα σημαντικές σε πραγματικά συστήματα πιστωτικής αξιολόγησης.

Συνοψίζοντας, οι παραπάνω περιορισμοί δεν αναιρούν την εγκυρότητα των αποτελεσμάτων, αλλά υπογραμμίζουν την ανάγκη προσεκτικής ερμηνείας τους και την περαιτέρω διερεύνηση του προβλήματος με πιο εκτεταμένες και δυναμικές προσεγγίσεις.

5.4 Καταλληλότητα των Μοντέλων για την Πρόβλεψη Αθέτησης

Η πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας αποτελεί ένα σύνθετο πρόβλημα δυαδικής ταξινόμησης, το οποίο χαρακτηρίζεται από ανισορροπία κλάσεων, πολυδιάστατα δεδομένα και αλληλεπιδράσεις μεταξύ οικονομικών και δημογραφικών μεταβλητών. Στο πλαίσιο αυτό, η επιλογή κατάλληλων τεχνικών μηχανικής μάθησης είναι κρίσιμη για την αποτελεσματική μοντελοποίηση του πιστωτικού κινδύνου.

Η Λογιστική Παλινδρόμηση αποδείχθηκε μια αξιόπιστη και σταθερή τεχνική, προσφέροντας ικανοποιητική συνολική ακρίβεια και υψηλή ερμηνευσιμότητα. Η γραμμική φύση του μοντέλου επιτρέπει την κατανόηση της επίδρασης κάθε μεταβλητής στην πιθανότητα αθέτησης, γεγονός που το καθιστά ιδιαίτερα κατάλληλο για εφαρμογές όπου η διαφάνεια και η επεξηγησιμότητα των αποφάσεων είναι απαραίτητες, όπως σε κανονιστικά πλαίσια χρηματοπιστωτικών οργανισμών. Ωστόσο, η περιορισμένη ικανότητά του να εντοπίζει μεγάλο ποσοστό πραγματικών περιπτώσεων αθέτησης περιορίζει την αποτελεσματικότητά του σε περιβάλλοντα όπου το κόστος των ψευδώς αρνητικών προβλέψεων είναι υψηλό.

Αντίθετα, το Δέντρο Απόφασης παρουσίασε βελτιωμένη απόδοση ως προς την ανίχνευση της μειοψηφικής κλάσης της αθέτησης, επιτυγχάνοντας υψηλότερες τιμές recall, F1-score και ROC-AUC. Η μη γραμμική του δομή του επιτρέπει να συλλαμβάνει πιο σύνθετες σχέσεις και αλληλεπιδράσεις μεταξύ των μεταβλητών, γεγονός που το καθιστά πιο κατάλληλο για προβλήματα πιστωτικού κινδύνου όπου η συμπεριφορά των πελατών δεν ακολουθεί απλές γραμμικές σχέσεις. Παράλληλα, η δυνατότητα οπτικοποίησης της λογικής των αποφάσεων διατηρεί ένα ικανοποιητικό επίπεδο ερμηνευσιμότητας.

Συνολικά, τα αποτελέσματα δείχνουν ότι καμία τεχνική δεν υπερέχει απόλυτα σε όλα τα κριτήρια. Η Λογιστική Παλινδρόμηση είναι περισσότερο κατάλληλη ως μοντέλο αναφοράς και για περιπτώσεις όπου προτεραιότητα αποτελεί η ερμηνευσιμότητα και η σταθερότητα, ενώ το Δέντρο Απόφασης εμφανίζεται καταλληλότερο για την αποτελεσματική ανίχνευση επισφαλών πελατών, όπου η έγκαιρη αναγνώριση του κινδύνου είναι καθοριστικής σημασίας.

Η σύγκριση αυτή αναδεικνύει ότι η επιλογή της κατάλληλης τεχνικής εξαρτάται άμεσα από τους επιχειρησιακούς στόχους και το κόστος των λαθών, επιβεβαιώνοντας τη σημασία της πολυκριτηριακής αξιολόγησης και της συνδυαστικής προσέγγισης στη μοντελοποίηση πιστωτικού κινδύνου (Thomas, 2000; Wahab et al., 2024).

5.5 Προτάσεις για Βελτιώσεις και Μελλοντική Έρευνα

Παρότι τα αποτελέσματα της παρούσας μελέτης είναι ικανοποιητικά και επιβεβαιώνουν τη χρησιμότητα των τεχνικών μηχανικής μάθησης στην πρόβλεψη αθέτησης πληρωμής, υπάρχουν αρκετές δυνατότητες περαιτέρω βελτίωσης και επέκτασης της ανάλυσης.

Μία πρώτη κατεύθυνση βελτίωσης αφορά τη δοκιμή πιο σύνθετων και ισχυρών αλγορίθμων μηχανικής μάθησης, όπως σύνολα δέντρων (ensemble methods), συμπεριλαμβανομένων των Random Forests και των Gradient Boosting μοντέλων. Οι μέθοδοι αυτές είναι σε θέση να μειώσουν τη μεταβλητότητα των μεμονωμένων δέντρων και να βελτιώσουν τη γενίκευση του μοντέλου, διατηρώντας παράλληλα υψηλή ικανότητα ανίχνευσης της μειοψηφικής κλάσης.

Επιπλέον, η περαιτέρω βελτιστοποίηση των υπερπαραμέτρων των μοντέλων μέσω συστηματικών διαδικασιών, όπως η διασταυρούμενη επικύρωση (cross-validation) και η αναζήτηση πλέγματος (grid search), θα μπορούσε να οδηγήσει σε βελτιωμένη απόδοση και μεγαλύτερη σταθερότητα των αποτελεσμάτων. Στην παρούσα εργασία επιλέχθηκαν συγκεκριμένες τιμές υπερπαραμέτρων με στόχο την αποφυγή υπερεκμάθησης, ωστόσο μια πιο εκτεταμένη διαδικασία βελτιστοποίησης ενδέχεται να αποδώσει περαιτέρω βελτιώσεις.

Μια ακόμη σημαντική προοπτική αφορά την πιο ενεργή αντιμετώπιση της ανισορροπίας των κλάσεων μέσω τεχνικών επαναδειγματοληψίας, όπως το SMOTE ή συνδυαστικές μέθοδοι υποδειγματοληψίας και υποδειγματοληψίας (over- και under-sampling). Η εφαρμογή τέτοιων τεχνικών θα μπορούσε να ενισχύσει περαιτέρω την ικανότητα των μοντέλων να εντοπίζουν περιπτώσεις αθέτησης, μειώνοντας τον αριθμό των σφαλμάτων false negative.

Παράλληλα, η ενσωμάτωση επιπλέον χαρακτηριστικών, όπως μακροοικονομικοί δείκτες ή δεδομένα συμπεριφοράς πελατών σε μεγαλύτερο χρονικό ορίζοντα, θα μπορούσε να εμπλουτίσει το σύνολο δεδομένων και να βελτιώσει την προβλεπτική ικανότητα των μοντέλων. Η δυναμική φύση του πιστωτικού κινδύνου υποδηλώνει ότι μοντέλα τα οποία λαμβάνουν υπόψη χρονικές μεταβολές και μοτίβα συμπεριφοράς ενδέχεται να είναι πιο αποτελεσματικά σε πραγματικές εφαρμογές.

Τέλος, ιδιαίτερο ενδιαφέρον παρουσιάζει η διερεύνηση τεχνικών επεξηγήσιμης τεχνητής νοημοσύνης (Explainable AI – XAI), όπως οι μέθοδοι SHAP και LIME, οι οποίες μπορούν να προσφέρουν πιο λεπτομερή και διαφανή ερμηνεία των προβλέψεων των μοντέλων. Η ενσωμάτωση τέτοιων τεχνικών θα ενίσχυε την εμπιστοσύνη των χρηστών στα συστήματα πρόβλεψης και θα διευκόλυνε τη συμμόρφωση με ρυθμιστικά και ηθικά πλαίσια στον χρηματοπιστωτικό τομέα.

Συνολικά, οι προτάσεις αυτές αναδεικνύουν ότι η παρούσα εργασία αποτελεί ένα σταθερό θεμέλιο για περαιτέρω έρευνα και εφαρμογή πιο προηγμένων τεχνικών μηχανικής μάθησης στη διαχείριση πιστωτικού κινδύνου (Wahab et al., 2024).

ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Ανακεφαλαίωση

Η παρούσα διπλωματική εργασία είχε ως βασικό στόχο τη διερεύνηση της αξιοποίησης τεχνικών Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης σε χρηματοοικονομικές εργασίες, με έμφαση στην πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας. Μέσα από συνδυασμό θεωρητικής προσέγγισης και εμπειρικής ανάλυσης, επιχειρήθηκε η σύνδεση της σύγχρονης βιβλιογραφίας με πρακτικές μεθοδολογίες ανάλυσης δεδομένων που εφαρμόζονται ευρέως στον τραπεζικό και χρηματοπιστωτικό τομέα, αναδεικνύοντας τον ρόλο της ΤΝ ως εργαλείο υποστήριξης κρίσιμων οικονομικών αποφάσεων.

Στα αρχικά κεφάλαια της εργασίας παρουσιάστηκε το θεωρητικό και θεσμικό πλαίσιο της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης και της Χρηματοοικονομικής Τεχνολογίας (FinTech), καθώς και η συμβολή τους στον ψηφιακό μετασχηματισμό του χρηματοπιστωτικού συστήματος. Παράλληλα, αναλύθηκαν βασικές εφαρμογές της ΤΝ στις χρηματοοικονομικές εργασίες, όπως η διαχείριση πιστωτικού κινδύνου και η ανίχνευση απάτης, ενώ δόθηκε ιδιαίτερη έμφαση σε ζητήματα διαφάνειας, ερμηνευσιμότητας και κανονιστικής

συμμόρφωσης, τα οποία καθίστανται κρίσιμα σε περιβάλλοντα υψηλού ρίσκου όπως οι τραπεζικές πιστοδοτήσεις.

Ιδιαίτερη βαρύτητα στην παρούσα εργασία δόθηκε στο Κεφάλαιο 4 (Μεθοδολογία), όπου αναπτύχθηκε αναλυτικά η ερευνητική διαδικασία και η πειραματική προσέγγιση που ακολουθήθηκε. Το πρόβλημα της πρόβλεψης αθέτησης πληρωμής διατυπώθηκε ως πρόβλημα επιτηρούμενης μάθησης δυαδικής ταξινόμησης, αξιοποιώντας πραγματικά χρηματοοικονομικά δεδομένα. Παρουσιάστηκαν διεξοδικά τα στάδια προεπεξεργασίας των δεδομένων, η εξερευνητική ανάλυση, η αντιμετώπιση της ανισορροπίας κλάσεων, καθώς και η επιλογή κατάλληλων μετρικών αξιολόγησης, ώστε να διασφαλιστεί η αξιοπιστία και η εγκυρότητα των αποτελεσμάτων.

Στο ίδιο κεφάλαιο πραγματοποιήθηκε η εκπαίδευση και αξιολόγηση δύο διαφορετικών μοντέλων Μηχανικής Μάθησης: της Λογιστικής Παλινδρόμησης και του Δέντρου Απόφασης. Η επιλογή των συγκεκριμένων μοντέλων επέτρεψε τη σύγκριση ενός απλού και ερμηνεύσιμου γραμμικού μοντέλου με ένα πιο ευέλικτο, μη γραμμικό μοντέλο, αναδεικνύοντας τα πλεονεκτήματα και τους περιορισμούς κάθε προσέγγισης στο πλαίσιο της διαχείρισης πιστωτικού κινδύνου.

Το Κεφάλαιο 5 (Ανάλυση και Συζήτηση Αποτελεσμάτων) αποτέλεσε τον βασικό πυρήνα της εμπειρικής ανάλυσης της εργασίας. Τα αποτελέσματα έδειξαν ότι η Λογιστική Παλινδρόμηση παρουσίασε υψηλή συνολική ακρίβεια, ωστόσο υστέρησε σημαντικά στην ανίχνευση της κλάσης της αθέτησης πληρωμής, εμφανίζοντας χαμηλές τιμές recall και αυξημένο αριθμό false negatives. Το εύρημα αυτό ανέδειξε τη μη επάρκεια της ακρίβειας ως μοναδικού κριτηρίου αξιολόγησης σε προβλήματα με έντονη ανισορροπία κλάσεων.

Αντίθετα, το Δέντρο Απόφασης παρουσίασε σαφώς βελτιωμένη απόδοση στην ανίχνευση επισφαλών πελατών, επιτυγχάνοντας υψηλότερες τιμές recall, F1-score και ROC-AUC για την κλάση της αθέτησης. Παρότι η συνολική ακρίβεια ήταν ελαφρώς χαμηλότερη, το μοντέλο κρίθηκε καταλληλότερο από επιχειρησιακή σκοπιά, καθώς μείωσε σημαντικά τον κίνδυνο μη εντοπισμού πελατών με υψηλή πιθανότητα αθέτησης. Η ανάλυση των false positives και false negatives ανέδειξε τη σημασία της προσαρμογής των μοντέλων στις πραγματικές οικονομικές συνέπειες των σφαλμάτων ταξινόμησης.

Συνολικά, τα αποτελέσματα των Κεφαλαίων 4 και 5 καταδεικνύουν ότι η αποτελεσματική εφαρμογή τεχνικών Μηχανικής Μάθησης στον χρηματοοικονομικό τομέα δεν εξαρτάται μόνο από την επιλογή του αλγορίθμου, αλλά κυρίως από τη σωστή μεθοδολογία, την κατάλληλη επιλογή μετρικών αξιολόγησης και την ερμηνεία των αποτελεσμάτων με βάση το πραγματικό κόστος του κινδύνου. Η εργασία επιβεβαιώνει ότι τα μοντέλα TN μπορούν να λειτουργήσουν ως αξιόπιστα εργαλεία υποστήριξης αποφάσεων στη διαχείριση πιστωτικού κινδύνου, όταν χρησιμοποιούνται με κριτικό και υπεύθυνο τρόπο.

6.2 Μελλοντική Έρευνα

Παρότι η παρούσα εργασία κατέδειξε τη χρησιμότητα και την αποτελεσματικότητα βασικών τεχνικών Μηχανικής Μάθησης στην πρόβλεψη αθέτησης πληρωμής πιστωτικής κάρτας, υπάρχουν σημαντικά περιθώρια περαιτέρω έρευνας και βελτίωσης. Ένα πρώτο πεδίο μελλοντικής διερεύνησης αφορά την αξιοποίηση πιο προηγμένων αλγοριθμικών προσεγγίσεων, όπως τα ensemble μοντέλα (Random Forest, Gradient Boosting, XGBoost), τα οποία έχουν αποδειχθεί ιδιαίτερα αποδοτικά σε προβλήματα ταξινόμησης με σύνθετες μη

γραμμικές σχέσεις και ανισορροπία κλάσεων. Η σύγκρισή τους με τα απλούστερα μοντέλα που εξετάστηκαν στην παρούσα εργασία θα μπορούσε να προσφέρει περαιτέρω βελτίωση στην ακρίβεια πρόβλεψης και στη μείωση του πιστωτικού κινδύνου.

Επιπλέον, μελλοντικές μελέτες θα μπορούσαν να ενσωματώσουν τεχνικές βελτιστοποίησης υπερπαραμέτρων και μεθόδους αντιμετώπισης της ανισορροπίας κλάσεων, όπως oversampling, undersampling ή cost-sensitive learning. Η συστηματική διερεύνηση της επίδρασης αυτών των τεχνικών στα false positives και false negatives θα επέτρεπε την καλύτερη προσαρμογή των μοντέλων στις πραγματικές ανάγκες των χρηματοπιστωτικών οργανισμών, όπου το κόστος των σφαλμάτων δεν είναι συμμετρικό.

Ένα ακόμη σημαντικό πεδίο μελλοντικής έρευνας αφορά την ερμηνευσιμότητα και τη διαφάνεια των μοντέλων Τεχνητής Νοημοσύνης. Αν και η παρούσα εργασία εστίασε σε σχετικά ερμηνεύσιμα μοντέλα, η εφαρμογή πιο σύνθετων αλγορίθμων καθιστά αναγκαία τη χρήση τεχνικών Explainable AI (XAI), όπως SHAP και LIME, ώστε οι αποφάσεις των μοντέλων να μπορούν να αιτιολογηθούν επαρκώς. Η ενίσχυση της ερμηνευσιμότητας είναι κρίσιμη τόσο για λόγους κανονιστικής συμμόρφωσης όσο και για την ενίσχυση της εμπιστοσύνης των χρηστών στα αυτοματοποιημένα συστήματα λήψης αποφάσεων.

Παράλληλα, η μελλοντική έρευνα θα μπορούσε να επεκτείνει το πλαίσιο ανάλυσης με τη χρήση δυναμικών και χρονικών δεδομένων (time-series), λαμβάνοντας υπόψη τη μεταβολή της οικονομικής συμπεριφοράς των πελατών στο χρόνο. Η ενσωμάτωση μακροοικονομικών μεταβλητών ή δεδομένων πραγματικού χρόνου θα μπορούσε να βελτιώσει σημαντικά την προσαρμοστικότητα και τη ρεαλιστικότητα των μοντέλων σε πραγματικές συνθήκες αγοράς.

Τέλος, ιδιαίτερο ενδιαφέρον παρουσιάζει η εφαρμογή των προτεινόμενων μοντέλων σε πραγματικά επιχειρησιακά περιβάλλοντα, σε συνδυασμό με κανονιστικά πλαίσια και πρακτικές διαχείρισης κινδύνου των τραπεζών. Η μελλοντική έρευνα θα μπορούσε να εστιάσει στη μελέτη της ενσωμάτωσης συστημάτων Τεχνητής Νοημοσύνης σε διαδικασίες πιστοδότησης, λαμβάνοντας υπόψη ζητήματα ηθικής, διακυβέρνησης δεδομένων και προστασίας προσωπικών δεδομένων. Με τον τρόπο αυτό, θα καταστεί δυνατή η ανάπτυξη ολοκληρωμένων, υπεύθυνων και βιώσιμων συστημάτων ΤΝ στον χρηματοοικονομικό τομέα.

Βιβλιογραφία

1. Σκαρλατάκης, Φ. (2022). Η χρηματοοικονομική τεχνολογία (FinTech) και ο ρόλος των παραδοσιακών τραπεζών στο μέλλον. Ελληνικό Ανοικτό Πανεπιστήμιο. Available at: <https://apothesis.eap.gr/archive/download/bb13141f-309e-4c6c-b829-b3eb7ba94a6d.pdf>
2. Μελισσάρη, Α. (2025). Χρηματοοικονομική Τεχνολογία (FinTech) και ο ρόλος των παραδοσιακών τραπεζών στο μέλλον. Διπλωματική εργασία, Ελληνικό Ανοικτό Πανεπιστήμιο. Available at: <https://apothesis.eap.gr/archive/download/bb3a9cfa-0415-4e24-b4ba-4f8307691b36.pdf>
3. Μπαλοδήμος, Π. (2024). Ψηφιακός μετασχηματισμός στα Πιστωτικά Ιδρύματα. Η επίδραση της υιοθέτησης καινοτομικών εφαρμογών Τεχνητής Νοημοσύνης & Μηχανικής Μάθησης στη λειτουργία των Ελληνικών τραπεζών. Διπλωματική εργασία, Πανεπιστήμιο Πατρών. Available at: <https://apothesis.eap.gr/archive/download/1b2bb70e-9928-4864-91b2-8911e01c6b7a.pdf>
4. Μάργαρης, Ι. (2025). Νέα εποχή στον τραπεζικό τομέα με τη χρήση της τεχνητής νοημοσύνης. Ελληνικό Ανοικτό Πανεπιστήμιο. Available at: <https://apothesis.eap.gr/archive/download/3008c797-d0f5-4a27-a28b-78cf21a1318a.pdf>
5. Challoumis, C. (2024, October). The evolution of financial systems-AI'S role in reshaping money management. In XVI International Scientific Conference (pp. 128-151). Available at: <https://conference-w.com/wp-content/uploads/2024/10/EST.T-1718102024.pdf#page=129>
6. Golić, Z. (2019). Finance and artificial intelligence: The fifth industrial revolution and its impact on the financial sector. Zbornik radova Ekonomskog fakulteta u Istočnom Sarajevu, (19), 67-81. Available at: <https://zrefis.ekofis.ues.rs.ba/images/zrefis2019-19/67-81%20Golic.pdf>
7. Vuković, D. B., Dekpo-Adza, S., & Matović, S. (2025). AI integration in financial services: a systematic review of trends and regulatory challenges. Humanities and Social Sciences Communications, 12(1), 1-29. Available at: [AI integration in financial services: a systematic review of trends and regulatory challenges](#)
8. Tarenzi, G. (2024). Artificial Intelligence: history, evolution and application with a focus in the Financial Services sector (Doctoral dissertation, Politecnico di Torino).
9. Russell, S. J., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach, Global Edition 4e. Available at: https://api.pageplace.de/preview/DT0400.9781292401171_A41586057/preview-9781292401171_A41586057.pdf
10. McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1956). Dartmouth summer research conference on artificial intelligence. Dartmouth College. Available at: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1903/1801>
11. Aldboush, H. H., & Ferdous, M. (2023). Building trust in fintech: an analysis of ethical and privacy considerations in the intersection of big data, AI, and customer

- trust. International Journal of Financial Studies, 11(3), 90. Available at: <https://www.mdpi.com/2227-7072/11/3/90/pdf>
12. Agu, E. E., Abhulimen, A. O., Obiki-Osafiele, A. N., Osundare, O. S., Adeniran, I. A., & Efunniyi, C. P. (2024). Discussing ethical considerations and solutions for ensuring fairness in AI-driven financial services. International Journal of Frontier Research in Science, 3(2), 001-009. Available at: [Discussing ethical considerations and solutions for ensuring fairness in AI-driven financial services](#)
 13. Lee, J. (2020). Access to finance for artificial intelligence regulation in the financial services industry. European Business Organization Law Review, 21(4), 731-757. Available at: [Access to Finance for Artificial Intelligence Regulation in the Financial Services Industry](#)
 14. Dudu, O. F., Alao, O. B., & Alonge, E. O. (2024). Conceptual framework for AI-driven tax compliance in fintech ecosystems. Journal of Fintech and Taxation, 12(3), 45-60. Available at: [Conceptual framework for AI-driven tax compliance in fintech ecosystems](#)
 15. Γκουνελάς, Π. (2019). Η εφαρμογή της τεχνητής νοημοσύνης στον τραπεζικό τομέα. Ελληνικό Ανοικτό Πανεπιστήμιο. Available at: <https://apothesis.eap.gr>
 16. Κυδώνα, Ε. (2021). Εφαρμογές της τεχνητής νοημοσύνης στην χρηματοοικονομική διοίκηση των επιχειρήσεων στην Ελλάδα και οι επιδοτήσεις ως χρηματοδοτικό εργαλείο. Available at: <https://dspace.lib.uom.gr>
 17. Παναγιωτοπούλου, Μ. (2021). Κίνδυνος χρηματοπιστωτικών ιδρυμάτων και αναγκαιότητα ελέγχου
 18. Maple, C., Szpruch, L., Epiphaniou, G., Staykova, K., Singh, S., Penwarden, W., ... & Avramovic, P. (2023). The AI revolution: Opportunities and challenges for the finance sector. arXiv preprint arXiv:2308.16538 Available at: <https://arxiv.org>
 19. Pattnaik, D., Ray, S., & Raman, R. (2024). Applications of artificial intelligence and machine learning in the financial services industry: A bibliometric review. Heliyon, 10(1). Available at: <https://pdf.sciencedirectassets.com>
 20. Wallon, C. (2019). Artificial intelligence applications in corporate finance Available at: <https://matheo.uliege.be>
 21. Vives, X. (2019). Digital disruption in banking. Annual Review of Financial Economics, 11(1), 243-272. Available at: https://blog.iese.edu/xvives/files/2020/01/Digital-Disruption-in-Banking_Nov.2019.pdf
 22. Gomber, P., Koch, J. A., & Siering, M. (2017). Digital Finance and FinTech: current research and future research directions. Journal of business economics, 87(5), 537-580. Available at: <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=2928833>
 23. Philippon, T. (2016). The fintech opportunity (No. w22476). National Bureau of Economic Research. Available at: https://www.nber.org/system/files/working_papers/w22476/w22476.pdf

24. Lehmann, M. (2020). Global rules for a global market place?-Regulation and supervision of Fintech providers. BU Int'l LJ, 38, 118. Available at: <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3421963>
25. Yeh, I. C., & Lien, C. H. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Expert systems with applications, 36(2), 2473-2480. Available at: https://bradzzz.gitbooks.io/ga-seattle-dsi/dsi/dsi_05_classification_databases/2.1-lesson/assets/datasets/DefaultCreditCardClients_yeh_2009.pdf
26. Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European journal of operational research, 247(1), 124-136. Available at: https://eprints.soton.ac.uk/377196/1/Lessmann_Benchmarking.pdf
27. Fawcett, T. (2006). An introduction to ROC analysis. Pattern recognition letters, 27(8), 861-874. Available at: <https://www.math.ucdavis.edu/~saito/data/roc/fawcett-roc.pdf>
28. Weiss, G. M. (2004). Mining with rarity: a unifying framework. ACM Sigkdd Explorations Newsletter, 6(1), 7-19. Available at: <https://dl.acm.org/doi/pdf/10.1145/1007730.1007734>
29. Komorowski, M., Marshall, D.C., Saliccioli, J.D., Crutain, Y. (2016). Exploratory Data Analysis. Secondary Analysis of Electronic Health Records. Springer, Cham. Available at: https://link.springer.com/content/pdf/10.1007/978-3-319-43742-2_15.pdf
30. Thomas, L. C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. International journal of forecasting, 16(2), 149-172. Available at: <http://www.yaroslavvb.com/papers/thomas-survey.pdf>
31. Wahab, F., Khan, I., & Sabada, S. (2024). Credit card default prediction using ML and DL techniques. Internet of Things and Cyber-Physical Systems, 4, 293-306. Available at: <https://www.sciencedirect.com/science/article/pii/S2667345224000087>
32. Henley, W., & Hand, D. J. (1996). Ak-nearest-neighbour classifier for assessing consumer credit risk. Journal of the Royal Statistical Society Series D: The Statistician, 45(1), 77-95. Available at: <http://pages.ucsd.edu/~aronatas/project/academic/Nearest%20neighbor%20analysis%20in%20credit%20scoring.pdf>

Παράρτημα Α: Πλήρης κώδικας σε Python

```
# Εισαγωγή βιβλιοθηκών
import pandas as pd
import numpy as np

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, OneHotEncoder
from sklearn.pipeline import Pipeline
from sklearn.compose import ColumnTransformer

from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier

from sklearn.metrics import (
    classification_report,
    confusion_matrix,
    roc_auc_score,
    roc_curve,
    accuracy_score,
    precision_score,
    recall_score,
    f1_score
)

import matplotlib.pyplot as plt
import seaborn as sns

# Φόρτωση dataset
df = pd.read_csv("UCI_Credit_Card.csv")

print("Διαστάσεις dataset:", df.shape)
display(df.head())

# ΕΞΕΡΕΥΝΗΤΙΚΗ ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ (EDA)

print("\nΠληροφορίες για τις στήλες:")
df.info()

print("\nΣτατιστική σύνοψη αριθμητικών μεταβλητών:")
display(df.describe())

print("\nΕλεγχος ελλειπουσών τιμών ανά στήλη:")
print(df.isnull().sum())

# Ορισμός μεταβλητής-στόχου
target = "default.payment.next.month"

print("\nΚατανομή της μεταβλητής-στόχου (πλήθος):")
print(df[target].value_counts())

print("\nΚατανομή της μεταβλητής-στόχου (ποσοστά %):")
print((df[target].value_counts(normalize=True) * 100).round(2))
```

```

# ΓΡΑΦΗΜΑ 1: Κατανομή μεταβλητής-στόχου
plt.figure(figsize=(5,4))
sns.countplot(x=target, data=df)
plt.title("Κατανομή Μεταβλητής-Στόχου (Default)")
plt.xlabel("Κατηγορία (0 = Όχι Default, 1 = Default)")
plt.ylabel("Πλήθος Παρατηρήσεων")
plt.tight_layout()
plt.show()

# ΓΡΑΦΗΜΑ 2: Πίνακας συσχέτισης (Correlation Heatmap)
plt.figure(figsize=(12,10))
corr = df.drop(columns=["ID"], errors="ignore").corr()
sns.heatmap(corr, cmap="coolwarm", center=0)
plt.title("Πίνακας Συσχέτισης Μεταβλητών (Pearson)")
plt.tight_layout()
plt.show()

# Καθαρισμός δεδομένων & ορισμός X, y
# Αφαίρεση της στήλης ID (δεν είναι χρήσιμη για πρόβλεψη)
if "ID" in df.columns:
    df = df.drop(columns=["ID"])

X = df.drop(columns=[target])
y = df[target]

# Ορισμός κατηγορικών και αριθμητικών μεταβλητών
categorical_features = ["SEX", "EDUCATION", "MARRIAGE"]
categorical_features = [c for c in categorical_features if c in X.columns]

numerical_features = [c for c in X.columns if c not in categorical_features]

print("\nΚατηγορικές μεταβλητές:", categorical_features)
print("Αριθμητικές μεταβλητές (ενδεικτικά):", numerical_features[:10])

# Pipeline προεπεξεργασίας δεδομένων
numeric_transformer = Pipeline(steps=[
    ("Κλιμάκωση", StandardScaler())
])

categorical_transformer = Pipeline(steps=[
    ("OneHot", OneHotEncoder(handle_unknown="ignore"))
])

preprocessor = ColumnTransformer(
    transformers=[
        ("Αριθμητικά", numeric_transformer, numerical_features),
        ("Κατηγορικά", categorical_transformer, categorical_features),
    ]
)

# Διαχωρισμός σε σύνολο εκπαίδευσης & ελέγχου
X_train, X_test, y_train, y_test = train_test_split(
    X, y,
    test_size=0.2,
    random_state=42,
    stratify=y
)

print("\nΜέγεθος train set:", X_train.shape)
print("Μέγεθος test set:", X_test.shape)

```

```

# Ορισμός των 2 Μοντέλων

# Λογιστική Παλινδρόμηση (Baseline)
model_log = Pipeline(steps=[
    ("Προεπεξεργασία", preprocessor),
    ("Μοντέλο", LogisticRegression(max_iter=2000))
])

# Δέντρο Απόφασης
model_tree = Pipeline(steps=[
    ("Προεπεξεργασία", preprocessor),
    ("Μοντέλο", DecisionTreeClassifier(
        max_depth=6,
        min_samples_leaf=50,
        class_weight="balanced",
        random_state=42
    ))
])

models = {
    "Λογιστική Παλινδρόμηση": model_log,
    "Δέντρο Απόφασης": model_tree
}

# Εκπαίδευση και Αξιολόγηση Μοντέλων
results = []
predictions = {}
probabilities = {}

for name, model in models.items():
    print("\n")
    print(f"Εκπαίδευση και Αξιολόγηση: {name}")
    print("")

    model.fit(X_train, y_train)

    y_pred = model.predict(X_test)
    y_proba = model.predict_proba(X_test)[:, 1]

    predictions[name] = y_pred
    probabilities[name] = y_proba

    print(classification_report(y_test, y_pred, digits=4))
    auc = roc_auc_score(y_test, y_proba)
    print("ROC AUC:", round(auc, 4))

    results.append({
        "Μοντέλο": name,
        "Accuracy": accuracy_score(y_test, y_pred),
        "Precision (Default=1)": precision_score(y_test, y_pred, pos_label=1, zero_division=0),
        "Recall (Default=1)": recall_score(y_test, y_pred, pos_label=1, zero_division=0),
        "F1-score (Default=1)": f1_score(y_test, y_pred, pos_label=1, zero_division=0),
        "ROC AUC": auc
    })

# Πίνακας σύγκρισης αποτελεσμάτων
results_df = pd.DataFrame(results).round(4)

print("\n")
print("ΣΥΓΚΡΙΣΗ ΤΩΝ 2 ΜΟΝΤΕΛΩΝ")
print("")
display(results_df)

```

```

# Πίνακες Σύγχυσης (Confusion Matrices)
for name, y_pred in predictions.items():
    cm = confusion_matrix(y_test, y_pred)

    plt.figure(figsize=(5,4))
    sns.heatmap(cm, annot=True, fmt="d", cmap="Blues")
    plt.title(f"Πίνακας Σύγχυσης - {name}")
    plt.xlabel("Πρόβλεψη")
    plt.ylabel("Πραγματική Τιμή")
    plt.tight_layout()
    plt.show()

# Καμπύλες ROC
plt.figure(figsize=(7,6))

colors = {
    "Λογιστική Παλινδρόμηση": "blue",
    "Δέντρο Απόφασης": "darkorange"
}

for name, y_proba in probabilities.items():
    fpr, tpr, _ = roc_curve(y_test, y_proba)
    plt.plot(fpr, tpr, label=name, color=colors.get(name, None), linewidth=2)

# Τυχάιος ταξινομητής
plt.plot([0,1], [0,1], "k--", label="Τυχάιος Ταξινομητής")

# Μορφοποίηση και εμφάνιση καμπυλών ROC
plt.xlabel("False Positive Rate")
plt.ylabel("True Positive Rate")
plt.title("Σύγκριση Καμπυλών ROC")
plt.legend()
plt.grid(True)
plt.show()

```