



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

**Η Εφαρμογή της Τεχνητής Νοημοσύνης στον Τραπεζικό  
κλάδο: Πολυσταδιακή Βαθιά Μάθηση για τον Εντοπισμό  
Απάτης και την Ερμηνευτικότητα Μοντέλων**

Διατριβή η οποία υποβλήθηκε για τη μερική εκπλήρωση  
των υποχρεώσεων απόκτησης του Διδακτορικού Διπλώματος

**Γεώργιος Ζιοβίρης**

Απρίλιος 2024





ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

**Η Εφαρμογή της Τεχνητής Νοημοσύνης στον Τραπεζικό  
κλάδο: Πολυσταδιακή Βαθιά Μάθηση για τον Εντοπισμό  
Απάτης και την Ερμηνευτικότητα Μοντέλων**

Διατριβή η οποία υποβλήθηκε για τη μερική εκπλήρωση  
των υποχρεώσεων απόκτησης του Διδακτορικού Διπλώματος

**Γεώργιος Ζιοβίρης**

Απρίλιος 2024





UNIVERSITY OF THESSALY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

# **AI-Driven Application in Banking: Multistage Deep Learning for Fraud Detection and Model Interpretability**

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy

**Georgios Zioviris**

April 2024





ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

**Η Εφαρμογή της Τεχνητής Νοημοσύνης στον Τραπεζικό κλάδο:  
Πολυσταδιακή Βαθιά Μάθηση για τον Εντοπισμό Απάτης και την  
Ερμηνευτικότητα Μοντέλων**

Διδακτορική Διατριβή

**Γεώργιος Ζιοβίρης**

Συμβουλευτική επιτροπή

**Γεώργιος Σταμούλης**, Καθηγητής του τμήματος ΗΜΜΥ, Πανεπιστημίου  
Θεσσαλίας (Επιβλέπων)

**Ασπασία Δασκαλοπούλου**, Αναπληρώτρια Καθηγήτρια του τμήματος ΗΜΜΥ,  
Πανεπιστημίου Θεσσαλίας

**Κότιος Άγγελος**, Καθηγητής του τμήματος Διεθνών και Ευρωπαϊκών Οικονομικών και  
Αναπτυξιακών Σχέσεων, Πανεπιστημίου Πειραιώς

Επταμελής εξεταστική επιτροπή

**Γεώργιος Σταμούλης**, Καθηγητής του τμήματος ΗΜΜΥ, Πανεπιστημίου  
Θεσσαλίας (Επιβλέπων)

**Ασπασία Δασκαλοπούλου**, Αναπληρώτρια Καθηγήτρια του τμήματος ΗΜΜΥ,  
Πανεπιστημίου Θεσσαλίας

**Κότιος Άγγελος**, Καθηγητής του τμήματος Διεθνών και Ευρωπαϊκών Οικονομικών και  
Αναπτυξιακών Σχέσεων, Πανεπιστημίου Πειραιώς

**Κωνσταντίνος Κολομβάτσος**, Επίκουρος Καθηγητής του τμήματος Πληροφορικής και  
Τηλεπικοινωνιών, Πανεπιστημίου Θεσσαλίας

**Ελευθέριος Τσουκαλάς**, Καθηγητής του τμήματος ΗΜΜΥ, Πανεπιστημίου Θεσσαλίας

**Βασίλειος Βλάχος**, Αναπληρωτής Καθηγητής του τμήματος Οικονομικών,  
Πανεπιστημίου Θεσσαλίας

**Ιωάννης Μουντάνος**, Αναπληρωτής Καθηγητής του τμήματος ΗΜΜΥ, Πανεπιστημίου  
Θεσσαλίας

Απρίλιος 2024

## **ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ ΠΕΡΙ ΑΚΑΔΗΜΑΪΚΗΣ ΔΕΟΝΤΟΛΟΓΙΑΣ ΚΑΙ ΠΝΕΥΜΑΤΙΚΩΝ ΔΙΚΑΙΩΜΑΤΩΝ**

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ρητά ότι η παρούσα διδακτορική διατριβή, καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας, αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία 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. Δηλώνω επίσης ότι τα αποτελέσματα της εργασίας δεν έχουν χρησιμοποιηθεί για την απόκτηση άλλου πτυχίου. Αναλαμβάνω πλήρως, ατομικά και προσωπικά, όλες τις νομικές και διοικητικές συνέπειες που δύναται να προκύψουν στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής.

Ο/Η Δηλών/ούσα

Γεώργιος Ζιοβίρης



# Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμότερες ευχαριστίες μου προς όσους συνέβαλαν στην επίτευξη αυτού του επιστημονικού ορόσημου στη ζωή μου. Η πορεία προς την απόκτηση του Διδακτορικού Διπλώματος ήταν μια σημαντική και πολύτιμη εμπειρία, και θέλω να εκφράσω τις ευγνωμοσύνες μου προς εκείνους που έπαιξαν σημαντικό ρόλο σε αυτό το ταξίδι.

Καταρχάς, θέλω να ευχαριστήσω τη μικρή μου κόρη Μαρία, η οποία στα τέσσερα της χρόνια αποτελεί πηγή απεριόριστης έμπνευσης. Η αθώα παιδική χαρά και η αγάπη της ανέδειξαν τον μαγικό κόσμο της ανυποκρισίας, και είμαι ευγνώμων για τη χαρά που μου χαρίζει καθημερινά.

Στη συνέχεια, θέλω να εκφράσω την βαθιά μου ευγνωμοσύνη προς την αγαπημένη μου σύζυγο Τατιάνα, η οποία με υποστήριξε ανιδιοτέλως καθ' όλη τη διάρκεια αυτής της πορείας. Η ανυποχώρητη συνεισφορά της και ο ενθουσιασμός της για την επίτευξη των στόχων μου είναι το καταφύγιο που με βοήθαι να προχωρώ μπροστά.

Τέλος, θέλω να ευχαριστήσω δύο εξαιρετικούς καθηγητές, κ.κ. Γεώργιο Σταμούλη & Κωνσταντίνο Κολομβάτσο που μου παρείχαν την ευκαιρία να ολοκληρώσω αυτό το επιστημονικό ταξίδι. Η εμπιστοσύνη και η καθοδήγησή τους ήταν καθοριστικής σημασίας για την επίτευξη του στόχου μου, και τους εκφράζω την βαθύτατη ευγνωμοσύνη μου για την ευγενική τους υποστήριξη.

Ευχαριστώ όλους τους αγαπημένους μου, οι οποίοι με στήριξαν κατά τη διάρκεια αυτής της αναπτυξιακής πορείας. Η εμπειρία αυτή θα μείνει αξέχαστη, και οι συνεισφορές όλων είναι ανεκτίμητες.

Με εκτίμηση,

Γεώργιος Ζιοβίρης



# Διδακτορική Διατριβή

## Η Εφαρμογή της Τεχνητής Νοημοσύνης στον Τραπεζικό κλάδο: Πολυσταδιακή Βαθιά Μάθηση για τον Εντοπισμό Απάτης και την Ερμηνευτικότητα Μοντέλων

Γεώργιος Ζιοβίρης

### Περίληψη

Ο τραπεζικός τομέας βρίσκεται στα πρόθυρα ενός σημαντικού μετασχηματισμού, με την κινητήρια δύναμη πίσω από αυτόν να είναι η Τεχνητή Νοημοσύνη (TN). Νέες εφαρμογές της τεχνητής νοημοσύνης έχουν ήδη προταθεί για την αντιμετώπιση προκλήσεων σε τομείς όπως η πιστοληπτική αξιολόγηση, η αξιολόγηση κινδύνου, η βελτίωση της εμπειρίας των πελατών και η αποτελεσματική διαχείριση του χαρτοφυλακίου. Ένα από τα πιο κρίσιμα ζητήματα στον ανωτέρω τομέα είναι ο εντοπισμός απάτης σε συναλλαγές. Πρόσφατα, εισήχθησαν μοντέλα βαθιάς μάθησης για την αντιμετώπιση αυτού του προβλήματος, στοχεύοντας στον εντοπισμό και την πρόβλεψη πιθανών δόλιων ενεργειών. Ο στόχος είναι να εκτιμηθεί η άγνωστη κατανομή των κανονικών/δόλιων συναλλαγών και, στη συνέχεια, να ανιχνευθούν αποκλίσεις που μπορεί να υποδηλώνουν πιθανή απάτη.

Σε αυτήν τη διατριβή, παρουσιάζουμε ένα νέο πολυεπίπεδο μοντέλο βαθιάς μάθησης που στοχεύει στην αποτελεσματική διαχείριση των εισερχομένων ροών συναλλαγών και στον εντοπισμό των δόλιων ενεργειών. Προτείνουμε τη χρήση αυτόματων κωδικοποιητών για την εκτέλεση επιλογής χαρακτηριστικών και την εκμάθηση της αναπαράστασης του κρυφού χώρου δεδομένων με βάση ένα μη γραμμικό μοντέλο βελτιστοποίησης. Στα προκαθορισμένα σημαντικά χαρακτηριστικά, εφαρμόζουμε στη συνέχεια ένα Συνελκτικό Νευρικό Δίκτυο (Convolutional Neural Network ή CNN) για τον εντοπισμό απάτης ή ένα Μοντέλο Μακράς Μνήμης με Βραχυπρόθεσμη Μνήμη (Long - Short Term Memory ή LSTM), συνδυάζοντας έτσι δύο διαφορετικά μπλοκ επεξεργασίας. Ο συνδυασμός που εφαρμόζουμε έχει ως στόχο την ανίχνευση απάτης σχετικά με την αποκαλυμμένη αναπαράσταση δεδομένων και όχι των αρχικών δεδομένων.

Ένας άλλος στόχος αυτής της διατριβής είναι να παρουσιάσει μια νέα βελτίωση του αλγο-

ρίθμου L.I.M.E. που χρησιμοποιείται για την ερμηνεία μοντέλων μηχανικής μάθησης. Εκτός από τη χρήση μοντέλων για την πρόβλεψη, η δυνατότητα ερμηνείας των αποτελεσμάτων που έχει αποκτήσει ένα μοντέλο λαμβάνει αυξημένη προσοχή στη σύγχρονη έρευνα. Δεν είναι ακόμη σαφές πώς συνδέεται το ευρύ φάσμα των προτεινόμενων μεθόδων ερμηνείας και ποιες κοινές έννοιες μπορούν να χρησιμοποιηθούν για την αξιολόγησή τους. Η τεχνική που προτείνουμε χρησιμοποιεί έναν αυτόματο κωδικοποιητή για τη δημιουργία νέων δεδομένων (neighborhood) γύρω από το data point που μας ενδιαφέρει να ερμηνεύσουμε, αντί να βασιστεί σε απλή δημιουργία δεδομένων με μια κανονική κατανομή γύρω από αυτό.

### **Λέξεις Κλειδιά**

Πιστωτικός Κίνδυνος ; Αυτοκωδικοποιητές; Συνελικτικά Νευρωνικά Δίκτυα; Μείωση Διαστάσεων; Ερμηνευσιμότητα; Τεχνικές Υπερδειγματοληψίας.

# PhD Dissertation

## **AI-Driven Application in Banking: Multistage Deep Learning for Fraud Detection and Model Interpretability**

**Georgios Zioviris**

### **Abstract**

The banking sector stands at the threshold of a profound transformation, and the driving force behind this revolution is Artificial Intelligence (AI). Innovative AI applications have already been put forth to address challenges across various domains, including credit scoring, risk assessment, enhancing client experiences, and portfolio management. Among the most pressing challenges in the banking sector is the detection of fraudulent activities within streams of transactions. Recently, deep learning models have emerged to tackle this specific issue, focusing on the identification and prediction of potential fraudulent events. The primary objective is twofold: first, to estimate the unknown distribution of normal and fraudulent transactions, and subsequently, to pinpoint anomalies that may signal potential fraud.

Within the pages of this dissertation, we delve into a groundbreaking multistage deep learning model designed to effectively handle incoming transaction streams and detect fraudulent activities. Our approach advocates the utilization of autoencoders for feature selection and the acquisition of latent data space representations based on nonlinear optimization models. On the carefully curated set of significant features, we then apply a deep Convolutional Neural Network (CNN) to detect instances of fraud. This amalgamation of techniques involves two distinct processing blocks and is tailored to identify fraud based on the exposed latent data representation rather than the initial data.

Another facet of this dissertation aims to introduce a novel enhancement to the LIME algorithm, a tool used for interpreting machine learning models. Beyond model prediction, the growing emphasis in contemporary research lies in the ability to decipher what a model has learned. The interconnectedness of the myriad interpretation methods proposed in the field and the identification of common evaluation concepts remain unclear. Our proposed technique employs an autoencoder to generate new instances around the data point under scrutiny, in contrast to the conventional method of simple data creation using a Gaussian

distribution centered around the point of interest.

### **Λέξεις Κλειδιά**

Fraud Detection; Autoencoder (AE); Variational Autoencoder (VAE); Convolutional Neural Network (CNN); Dimensionality Reduction; LSTM; L.I.M.E.; Interpretability; Oversampling techniques.

# Πίνακας περιεχομένων

<b>Περίληψη</b>	<b>xiii</b>
<b>Abstract</b>	<b>xv</b>
<b>Πίνακας περιεχομένων</b>	<b>xvii</b>
<b>Κατάλογος σχημάτων</b>	<b>xxiii</b>
<b>Κατάλογος πινάκων</b>	<b>xxv</b>
<b>Συνοτομογραφίες</b>	<b>xxvii</b>
<b>1 Εισαγωγή</b>	<b>1</b>
1.1 Η Πρόκληση της Ερμηνευσιμότητας . . . . .	3
1.2 Ερευνητικά Ερωτήματα και Συνεισφορές . . . . .	4
1.3 Δομή της Διατριβής . . . . .	4
<b>2 Θεωρητικό Υπόβαθρο και Συστημική Βιβλιογραφική Ανασκόπηση</b>	<b>7</b>
2.1 Εισαγωγή . . . . .	7
2.2 Υλικά Και Μέθοδοι Βιβλιογραφικής Ανασκόπησης . . . . .	9
2.3 Συζήτηση Βιβλιογραφικής Ανασκόπησης . . . . .	11
2.4 Επίλογος Βιβλιογραφικής Ανασκόπησης . . . . .	13
<b>3 Διαχείριση Ρίσκου στον Τραπεζικό Τομέα</b>	<b>17</b>
3.0.1 Πιστωτικός Κίνδυνος . . . . .	22
3.0.2 Κίνδυνος Αγοράς . . . . .	30
3.0.3 Κίνδυνος Ρευστότητας . . . . .	31
3.0.4 Λειτουργικός / Επιχειρησιακός Κίνδυνος . . . . .	31

<b>4</b>	<b>Machine Learning (ML, Μηχανική Μάθηση)</b>	<b>35</b>
4.1	Αλγόριθμοι Μηχανικής Μάθησης . . . . .	37
4.1.1	Ο αλγόριθμος K-Nearest Neighbors (K.N.N.) . . . . .	38
4.1.2	Πολυωνυμική παλινδρόμηση . . . . .	38
4.1.3	Support Vector Machine (S.V.M.) . . . . .	38
4.1.4	Artificial Neural Network (A.N.N.) . . . . .	39
<b>5</b>	<b>Deep Learning (DL, Βαθιά Μάθηση)</b>	<b>41</b>
5.1	Autoencoders . . . . .	41
5.2	Variational Autoencoders . . . . .	43
5.3	Convolutional Neural Networks, (CNNs) . . . . .	44
5.4	Long Short Term Memory . . . . .	46
<b>6</b>	<b>Πρωτότυπο θεωρητικό πλαίσιο και σχετικές υποθέσεις</b>	<b>49</b>
6.1	Επισκόπηση των συναφών οικονομικών θεωριών . . . . .	49
6.2	Πρότυπο Εννοιολογικό Πλαίσιο . . . . .	50
6.3	Ανάπτυξη υποθέσεων . . . . .	51
6.4	Επιλογή μοντέλων μηχανικής μάθησης και βαθιάς μάθησης . . . . .	52
6.5	Μετρήσεις αξιολόγησης και σύγκρισης μοντέλων . . . . .	53
<b>7</b>	<b>Ερευνητικά Κενά και Ευκαιρίες για Έρευνα</b>	<b>57</b>
7.1	Περιορισμένη εστίαση σε μακροοικονομικούς παράγοντες . . . . .	58
7.1.1	Σημασία των μακροοικονομικών παραγόντων . . . . .	59
7.1.2	Αλληλεπίδραση μεταξύ μακροοικονομικών συνθηκών και εταιρικού πιστωτικού κινδύνου . . . . .	59
7.1.3	Ενσωμάτωση μακροοικονομικών παραγόντων στην ανάλυση πιστωτικού κινδύνου . . . . .	59
7.1.4	Προκλήσεις στην ενσωμάτωση μακροοικονομικών παραγόντων . . . . .	60
7.1.5	Μελλοντικές ερευνητικές ευκαιρίες . . . . .	60
7.2	Ενσωμάτωση προηγμένων μοντέλων βαθιάς μάθησης . . . . .	61
7.2.1	Μέθοδοι ενσωμάτωσης μακροοικονομικών παραγόντων . . . . .	61
7.2.2	Προκλήσεις στην ενσωμάτωση μακροοικονομικών παραγόντων . . . . .	62
7.2.3	Μελλοντικές ερευνητικές ευκαιρίες . . . . .	62
7.3	Πολυφασικές και υβριδικές προσεγγίσεις . . . . .	63

7.3.1	Πιθανά οφέλη των πολυφασικών και υβριδικών προσεγγίσεων . . .	63
7.3.2	Προκλήσεις κατά την εφαρμογή πολυφασικών και υβριδικών προσεγγίσεων . . . . .	64
7.3.3	Μελλοντικές ερευνητικές ευκαιρίες . . . . .	65
7.4	Αντιμετώπιση μη ισορροπημένων συνόλων δεδομένων . . . . .	66
7.4.1	Επιπτώσεις των μη ισορροπημένων συνόλων δεδομένων . . . . .	66
7.4.2	Τεχνικές για την αντιμετώπιση ανισοβαρών συνόλων δεδομένων . .	66
7.4.3	Ερευνητικές ευκαιρίες . . . . .	67
7.5	Ερμηνευσιμότητα και επεξηγησιμότητα του μοντέλου . . . . .	68
7.5.1	Σημασία της ερμηνευσιμότητας και της επεξηγηματικότητας . . . . .	68
7.5.2	Προκλήσεις για την επίτευξη ερμηνευσιμότητας και επεξηγηματικότητας . . . . .	69
7.5.3	Ερευνητικές ευκαιρίες . . . . .	69
<b>8</b>	<b>Ανάλυση Δεδομένων και Αποτελέσματα</b>	<b>71</b>
8.1	Εκτίμηση Πιστωτικού Ρίσκου με τη χρήση Μακροοικονομικών Παραγόντων	71
8.1.1	Περιγραφικά στατιστικά στοιχεία . . . . .	71
8.1.2	Ανάλυση συσχέτισης . . . . .	72
8.1.3	Επιλογή χαρακτηριστικών και μείωση διαστάσεων . . . . .	74
8.1.4	Εκπαίδευση και επικύρωση μοντέλου . . . . .	76
8.1.5	Σύγκριση μοντέλων και αξιολόγηση επιδόσεων . . . . .	77
8.1.6	Έλεγχοι αξιοπιστίας και ανάλυση ευαισθησίας . . . . .	80
8.1.7	Δοκιμές προσομοίωσης ακραίων καταστάσεων υπό διαφορετικά μακροοικονομικά σενάρια . . . . .	81
<b>9</b>	<b>Καινοτομία Διατριβής</b>	<b>83</b>
9.1	Εντοπισμός fraud συναλλαγών με τη χρήση πολυδιάστατων μοντέλων βαθιάς μάθησης . . . . .	84
9.1.1	Εισαγωγή . . . . .	84
9.1.2	Υβριδικό Μοντέλο: Συνδυάζοντας Autoencoder και CNN . . . . .	85
9.1.3	Υβριδικό Μοντέλο: Συνδυάζοντας Autoencoder με LSTM . . . . .	91
9.2	Νέα βελτιωμένη τεχνική ερμηνευτικότητας με βάση την τεχνική L.I.M.E. .	96
9.2.1	Εισαγωγή . . . . .	96

9.2.2	Καινομία . . . . .	97
9.2.3	Πειραματικά Αποτελέσματα . . . . .	100
<b>10</b>	<b>Συζήτηση</b>	<b>103</b>
10.1	Ερμηνεία των αποτελεσμάτων . . . . .	103
10.2	Συνέπειες για τη θεωρία και την πράξη . . . . .	105
10.3	Περιορισμοί και μελλοντικές ερευνητικές κατευθύνσεις . . . . .	107
<b>11</b>	<b>Συμπέρασμα</b>	<b>111</b>
11.1	Σύνοψη των ευρημάτων . . . . .	111
11.2	Συνεισφορές και επιπτώσεις . . . . .	113
11.3	Τελικές παρατηρήσεις . . . . .	114
	<b>ΠΑΡΑΡΤΗΜΑΤΑ</b>	<b>117</b>
<b>A</b>	<b>Συμπληρωματικά δεδομένα και αποτελέσματα</b>	<b>119</b>
A.1	Συμπληρωματικές περιγραφικές στατιστικές . . . . .	119
A.2	Πρόσθετα μοντέλα μηχανικής μάθησης . . . . .	119
A.2.1	Random Forests . . . . .	120
A.2.2	Gradient Boosting . . . . .	120
A.3	Διευρυμένη ανάλυση επιλογής χαρακτηριστικών . . . . .	122
A.4	Ανάλυση ευαισθησίας με διαφορετικές τεχνικές επαναδειγματοληψίας . . . . .	124
A.5	Ανάλυση ερμηνευσιμότητας με χρήση SHapley Additive exPlanations (SHAP) . . . . .	125
<b>B</b>	<b>Κώδικας και τεκμηρίωση του μοντέλου εφαρμογής</b>	<b>129</b>
B.1	Τεκμηρίωση και περιγραφή του κώδικα . . . . .	129
B.2	Επιλογή χαρακτηριστικών και μείωση διαστάσεων . . . . .	131
B.3	Μοντέλα μηχανικής μάθησης και βαθιάς μάθησης . . . . .	134
B.4	Εκπαίδευση και επικύρωση μοντέλων . . . . .	137
B.5	Αξιολόγηση και σύγκριση μοντέλων . . . . .	139
B.6	Έλεγχοι αξιοπιστίας και ανάλυση ευαισθησίας . . . . .	142
B.7	Ερμηνευσιμότητα και ερμηνευσιμότητα του μοντέλου . . . . .	145
<b>Γ</b>	<b>Δεοντολογία &amp; Συγκατάθεση</b>	<b>149</b>
Γ.1	Έντυπα έγκρισης δεοντολογίας και συγκατάθεσης . . . . .	149

---

Γ.2	Έγκριση δεοντολογίας . . . . .	149
Γ.3	Έντυπα συγκατάθεσης . . . . .	150
<b>Δ</b>	<b>Επικαιροποιήσεις βιβλιογραφίας (2023)</b>	<b>151</b>
Δ.1	Νευρωνικά δίκτυα γραφημάτων (GNNs) . . . . .	151
Δ.2	Εξηγήσιμη τεχνητή νοημοσύνη (XAI) . . . . .	152
Δ.3	Μάθηση μεταφοράς και προσαρμογή στον τομέα . . . . .	152
Δ.4	Transformer-based Models . . . . .	152
Δ.5	Μετρικές αμεροληψίας . . . . .	152
<b>Z</b>	<b>Πίνακας Συστημικής Ανασκόπησης</b>	<b>155</b>
<b>E</b>	<b>Βιβλιογραφία</b>	<b>157</b>



# Κατάλογος σχημάτων

3.1	Ταξινόμηση των κινδύνων . . . . .	20
3.2	Μέθοδοι και εργαλεία διαχείρισης κινδύνων . . . . .	21
4.1	Τυπική αρχιτεκτονική A.N.N. . . . .	40
5.1	Τυπική αρχιτεκτονική Autoencoder . . . . .	42
5.2	Τυπική αρχιτεκτονική Variational Autoencoder . . . . .	44
5.3	Τυπική αρχιτεκτονική Convolutional Neural Network . . . . .	46
5.4	Τυπική αρχιτεκτονική Convolutional Neural Network . . . . .	46
5.5	Τυπική αρχιτεκτονική A.N.N. . . . .	48
Z.1	Πίνακας Συστημικής Ανασκόπησης . . . . .	156



# Κατάλογος πινάκων

8.1	Μακροοικονομικοί δείκτες . . . . .	71
8.2	Περιγραφικά στατιστικά στοιχεία χρηματοοικονομικών δεδομένων επιχειρήσεων . . . . .	72
8.3	Περιγραφικά στατιστικά στοιχεία του πιστωτικού κινδύνου: Αποτελέσματα	72
8.4	Πίνακας συσχέτισης μεταβλητών . . . . .	73
9.1	3 Folds Cross Validation Experiments . . . . .	87
9.2	4 Folds Cross Validation Experiments . . . . .	88
9.3	5 Folds Cross Validation Experiments . . . . .	89
9.4	10 Folds Cross Validation Experiments . . . . .	90
9.5	Σύγκριτική αξιολόγηση με [167], AE & RBM . . . . .	90
9.6	Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166] χωρίς τη χρήση AdaBoost	91
9.7	Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166] με τη χρήση AdaBoost	92
9.8	Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166] . . . . .	93
9.9	Group of Experiments Implementing a Simple Autoencoder . . . . .	94
9.10	<b>Group of Experiments Implementing a Variational Autoencoder . . . . .</b>	<b>94</b>
9.11	<b>Group of Experiments Implementing Conventional Machine Learning Algorithms . . . . .</b>	<b>94</b>
9.12	Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [168], LSTM . . . . .	96
9.13	<b>Average F1 of trustworthiness for different explainers on a collection of classifiers regarding the Books Dataset Ribeiro . . . . .</b>	<b>102</b>
9.14	<b>Average F1 of trustworthiness for different explainers on a collection of classifiers regarding the DVDs Dataset Ribeiro . . . . .</b>	<b>102</b>



# Συντομογραφίες

βλπ	βλέπε
κ.α.	και άλλοι
κ.λπ.	και λοιπά
κ.ο.κ	και ούτω καθεξής
XI	Χρηματοπιστωτικά ιδρύματα
TN	Τεχνητή Νοημοσύνη
ML	Machine Learning
AI	Artificial Intelligence
C.N.N.	Convolution Neural Network
L.S.T.M.	Long-Short Term Memory
S.M.O.T.E.	Synthetic Minority Over-sampling Technique
M.L.	Machine Learning
D.L.	Deep Learning
A.E.	Autoencoder
V.A.E.	Variational Autoencoder
L.I.M.E.	Local Interpretable Model-agnostic Explanations
S.V.M.	Support Vector Machine
K.N.N.	K-Nearest Neighbor
R.F.	Random Forest
C4.5	Decision Tree
A.N.N.	Artificial Neural Network
P.L.S.	Partial Least Squares
MARS	Multivariate Adaptive Regression Splines
CART	Classification and Regression Trees
VAR	Vector Autoregression

---

VaR	Value at Risk
EL	Expected Loss
LGD	Loss Given Default
PD	Probability of Default
EAD	Exposure at Default
SEC	Securities and Exchange Commission
E.M.A.	Environmental Management Assistant
E.L.M.	Extreme Machine Learning
GARCH	Generalized Autoregressive Conditional Heteroskedastic
GELF-RBF	Generalized Extreme Learning Machine - Radial Basis Function
RBF	Radial Basis Function
SDE	Stochastic Differential Equations
B.N.	Bayesian Networks
NN	Numerical Scores
N.N.	Neural Network
LR	Linear Regression
S.V.R.	Support Vector Regression
VAE	Variational Autoencoders
NLP	Natural Language Processing
ReLU	REctified Linear Unit
RNN	Recurrent Neural Network
AUC	Area Under Curve
ROC	Receiver Operating Characteristic
RFE	Recursive Feature Elimination
PCA	Principal Components Analysis
RBM	Restricted Boltzmann Machine

# Κεφάλαιο 1

## Εισαγωγή

Η διάδοση των ψηφιακών συναλλαγών και η παγκόσμια εξάρτηση από τις πληρωμές μέσω πιστωτικών καρτών έχουν μετασηματίσει τον τρόπο που διεξάγουμε οικονομικές συναλλαγές στον σύγχρονο κόσμο. Η ευκολία χρήσης των πιστωτικών καρτών έχει διευκολύνει αναμφισβήτητα τη διαδικασία αγοράς αγαθών και υπηρεσιών, τόσο σε φυσικά καταστήματα όσο και στον ευρύ χώρο του ηλεκτρονικού εμπορίου. Ωστόσο, αυτή η ευκολία συνοδεύεται από έναν πανταχού παρόντα κίνδυνο απάτης μέσω πιστωτικών καρτών. Καθώς η ψηφιοποίηση των οικονομικών συναλλαγών συνεχίζει να προχωρά, το ίδιο κάνουν και οι μέθοδοι που χρησιμοποιούνται από επιτήδους για να εκμεταλλευτούν τα αδύναμα σημεία των συστημάτων πληρωμών, καθιστώντας το έργο της ανίχνευσης και πρόληψης της απάτης μια διαρκή και κρίσιμη πρόκληση.

Τις τελευταίες δεκαετίες, ο τομέας της ανίχνευσης απάτης μέσω πιστωτικών καρτών βίωσε σημαντικές εξελίξεις, ιδίως με την εμφάνιση τεχνικών μηχανικής μάθησης (M.L.) και βαθιάς μάθησης (D.L.). Αυτές οι προσεγγίσεις έχουν επανασχεδιάσει τον τρόπο που οι οικονομικοί φορείς και οι επιχειρήσεις καταπολεμούν τις απάτες. Χρησιμοποιώντας τη δύναμη της ανάλυσης δεδομένων και της αναγνώρισης προτύπων, οι αλγόριθμοι M.L. και D.L. έχουν τη δυνατότητα ανίχνευσης απάτης με υψηλό βαθμό ακρίβειας και αποδοτικότητας. Ωστόσο, αυτή η αποτελεσματικότητα συχνά συνοδεύεται από έλλειψη ερμηνευσιμότητας, ένα κρίσιμο στοιχείο κάθε συστήματος ανίχνευσης απάτης.

Η ερμηνευσιμότητα στη μηχανική μάθηση είναι η δυνατότητα να κατανοήσουμε και να εξηγήσουμε πώς ένα μοντέλο λαμβάνει τις αποφάσεις ή κάνει τις προβλέψεις του. Είναι απαραίτητη σε πολλούς τομείς, συμπεριλαμβανομένης της υγείας, της χρηματοοικονομίας και του δικαίου, όπου οι συνέπειες μιας ενδεχόμενης εσφαλμένης απόφασης μπορούν να είναι

σημαντικές. Στο πλαίσιο της ανίχνευσης απάτης μέσω πιστωτικών καρτών, η ερμηνευσιμότητα είναι εξίσου κρίσιμη. Οι οικονομικοί φορείς, οι ρυθμιστικοί φορείς και οι πελάτες οι ίδιοι χρειάζονται να κατανοήσουν γιατί μια συγκεκριμένη συναλλαγή έχει χαρακτηριστεί ως μια πιθανή απάτη. Η ερμηνευσιμότητα όχι μόνο εξασφαλίζει διαφάνεια και εμπιστοσύνη, αλλά επιτρέπει επίσης την βελτίωση των συστημάτων ανίχνευσης απάτης με το πέρασ του χρόνου.

Αυτή η διδακτορική διατριβή αφιερώνεται στην εξερεύνηση του συνδυασμού της ανίχνευσης απάτης μέσω πιστωτικών καρτών και της ερμηνευσιμότητας στο πλαίσιο των σύγχρονων μεθόδων M.L. και D.L.. Θα εξετάσουμε τις κύριες προκλήσεις και ευκαιρίες σε αυτόν τον τομέα, με στόχο την ανάπτυξη νέων υβριδικών αλγορίθμων που επιτυγχάνουν ισορροπία ανάμεσα στην ακρίβεια και την ερμηνευσιμότητα των μοντέλων ανίχνευσης απάτης. Στόχος μας είναι η βελτίωση της συνολικής αποτελεσματικότητας των συστημάτων ανίχνευσης απάτης μέσω πιστωτικών καρτών, παρέχοντας ταυτόχρονα στους εμπλεκόμενους μια σαφή κατανόηση γιατί ορισμένες συναλλαγές κατηγοριοποιούνται ως απάτη.

Σε αυτήν την εισαγωγή, θα παρουσιάσουμε τη δομή και το πεδίο εφαρμογής αυτής της διατριβής, θα δώσουμε μια επισκόπηση της τρέχουσας κατάστασης της ανίχνευσης απάτης μέσω πιστωτικών καρτών και θα συζητήσουμε τη σημασία της ερμηνευσιμότητας στη μηχανική μάθηση. Επιπλέον, θα επισημάνουμε τα ερευνητικά ερωτήματα, τους στόχους και τις συνεισφορές αυτής της εργασίας, θέτοντας το πλαίσιο για τα κεφάλαια που ακολουθούν.

Η Τρέχουσα Κατάσταση στον Τομέα της Ανίχνευσης Απάτης μέσω Πιστωτικών Καρτών

Ο τομέας της ανίχνευσης απάτης μέσω πιστωτικών καρτών έχει υποστεί σημαντικές εξελίξεις τα τελευταία χρόνια. Πριν την ανάπτυξη του κλάδου της Μηχανικής και Βαθιάς Μάθησης, οι παραδοσιακές στατιστικές τεχνικές ήταν αυτές που χρησιμοποιούνταν κατά κανόνα. Ενώ αυτά τα συστήματα ήταν εν μέρει αποτελεσματικά, αντιμετώπιζαν δυσκολίες στο να προσαρμοστούν στις διαρκώς μεταβαλλόμενες τακτικές που χρησιμοποιούν οι απατεώνες για να εκμεταλλευτούν τις αδυναμίες των συστημάτων πληρωμών. Ως αποτέλεσμα, οι οικονομικοί φορείς και οι επιχειρήσεις στράφηκαν σε πιο προηγμένες τεχνικές.

Τα προαναφερθέντα στατιστικά μοντέλα, όπως η λογιστική παλινδρόμηση, ήταν ανάμεσα στα πρώτα που εφαρμόστηκαν στην ανίχνευση απάτης. Αυτά τα μοντέλα χρησιμοποιούν ιστορικά δεδομένα συναλλαγών για να ανιχνεύσουν πρότυπα που είναι ενδεικτικά της απάτης. Παρόλο που προσέφεραν μεγαλύτερη ακρίβεια σε σχέση με τα συστήματα που βασίζονταν σε κανόνες, είχαν περιορισμένη δυνατότητα να εντοπίσουν πολύπλοκες, μη γραμμ-

μικές σχέσεις στα δεδομένα.

Η εμφάνιση της μηχανικής μάθησης επέφερε μια παραδειγματική αλλαγή στην ανίχνευση απάτης. Οι αλγόριθμοι M.L., όπως τα δέντρα απόφασης (Decision Trees), τα τυχαία δάση (Random Forests) και οι μηχανές υποστήριξης (Support Vector Machines), έδειξαν υψηλή απόδοση αναγνωρίζοντας αυτόματα πολύπλοκα πρότυπα από τα δεδομένα. Οι μέθοδοι συνδυασμού όπως η βελτιστοποίηση κλίσης βελτίωσαν ακόμη περισσότερο την ακρίβεια συνδυάζοντας πολλά μοντέλα. Αυτές οι εξελίξεις επέτρεψαν στους οικονομικούς φορείς να ανιχνεύουν την απάτη με μεγαλύτερη αποτελεσματικότητα, μειώνοντας τις ψευδείς θετικές ενδείξεις και ελαχιστοποιώντας τις οικονομικές απώλειες.

## 1.1 Η Πρόκληση της Ερμηνευσιμότητας

Παρόλο που οι αλγόριθμοι μηχανικής μάθησης προσέφεραν σχετική ακρίβεια στην ανίχνευση απάτης, συχνά λειτουργούσαν ως "μαύρα κουτιά". Αυτό σημαίνει ότι οι αποφάσεις που λαμβάνονταν ήταν δυσνόητες, δημιουργώντας ανησυχίες για τη διαφάνεια, την ευθύνη και τη συμμόρφωση με τη νομοθεσία. Οι εμπλεκόμενοι, συμπεριλαμβανομένων των πελατών, των ρυθμιστικών αρχών και των εσωτερικών ομάδων ελέγχου, χρειάζονταν περισσότερα από μια δυαδική απόφαση απάτης/όχι απάτης. Χρειάζονταν εξηγήσεις.

Η έλλειψη ερμηνευσιμότητας στα μοντέλα μηχανικής μάθησης οφείλεται στην πολυπλοκότητά τους. Τα βαθιά νευρωνικά δίκτυα (Deep Neural Networks), για παράδειγμα, αποτελούνται από πολλά διασυνδεδεμένα επίπεδα, καθένα από τα οποία περιέχει πληθώρα νευρώνων. Οι περίπλοκες αυτές μετασχηματίσεις που συμβαίνουν στα εσωτερικά των αυτών των δικτύων καθιστούν δύσκολο το να ανιχνευτεί πώς μια συγκεκριμένη είσοδος οδηγεί σε μια συγκεκριμένη έξοδο. Ως αποτέλεσμα, οι οικονομικοί φορείς βρέθηκαν σε δίλημμα: είτε να θυσιάσουν την ακρίβεια για τη ερμηνευτικότητα, είτε να εισάγουν πολύπλοκα μοντέλα με περιορισμένη ερμηνευσιμότητα.

Αυτή η σχέση μεταξύ ακρίβειας και ερμηνευσιμότητας βρίσκεται στο επίκεντρο της έρευνάς μας. Ο στόχος μας είναι να γεφυρώσουμε το χάσμα μεταξύ αυτών των δύο φαινομενικά αντιφατικών στόχων, αναπτύσσοντας υβριδικούς αλγορίθμους που συνδυάζοντας τα καλύτερα στοιχεία από τους δύο αυτούς κόσμους. Θέλουμε να επιτύχουμε την υψηλή ακρίβεια των αλγορίθμων M.L. και D.L., ενώ παράλληλα να παρέχουμε ερμηνευσιμότητα στις αποφάσεις τους, επιτρέποντας στους χρήστες να κατανοήσουν τον λόγο πίσω από κάθε

απόφαση.

## 1.2 Ερευνητικά Ερωτήματα και Συνεισφορές

Αυτή η διδακτορική διατριβή αναλύει τα εξής κύρια ερευνητικά ερωτήματα:

- Πώς μπορούμε να βελτιώσουμε την ακρίβεια των αλγορίθμων ανίχνευσης απάτης μέσω πιστωτικών καρτών χωρίς να θυσιάζουμε την ερμηνευσιμότητά τους; Εξερευνούμε τη χρήση υβριδικών μοντέλων που συνδυάζουν τεχνικές M.L. και D.L. με μεθόδους ερμηνευσιμότητας, όπως τη διαφάνεια του μοντέλου, τη γενίκευση των αποφάσεων και την αναγνώριση των προτύπων.
- Πώς μπορούμε να δημιουργήσουμε εργαλεία και πλατφόρμες που επιτρέπουν την ερμηνευσιμότητα των μοντέλων ανίχνευσης απάτης στον πραγματικό κόσμο; Αναπτύσσουμε πρακτικές εφαρμογές και εργαλεία που επιτρέπουν στους χρήστες να αντιληφθούν τη λειτουργία των μοντέλων και τις αποφάσεις που λαμβάνουν.
- Πώς μπορούμε να διαχειριστούμε την είσοδο δεδομένων σε πραγματικό χρόνο στον τομέα της ανίχνευσης απάτης μέσω πιστωτικών καρτών; Εξετάζουμε τις τεχνικές αντιμετώπισης του όγκου και της ταχύτητας των δεδομένων, καθώς και την αναπτυξιακή επεκτασιμότητα των μοντέλων.

Στο πλαίσιο αυτής της διατριβής, προτείνουμε νέους αλγορίθμους, αναπτύσσουμε εργαλεία και παρουσιάζουμε πειραματικά αποτελέσματα που δείχνουν την αποτελεσματικότητά τους. Η συνεισφορά μας αναμένεται να βοηθήσει στην εξέλιξη της ανίχνευσης απάτης μέσω πιστωτικών καρτών προσφέροντας πιο αξιόπιστες, διαφανείς και αποτελεσματικές λύσεις.

## 1.3 Δομή της Διατριβής

Τα επόμενα κεφάλαια της διατριβής έχουν οργανωθεί ως εξής:

Στο Κεφάλαιο 2, παρουσιάζουμε μια λεπτομερή ανασκόπηση της βιβλιογραφίας αναφορικά με την ανίχνευση απάτης μέσω πιστωτικών καρτών και την ερμηνευσιμότητα στη μηχανική μάθηση. Το Κεφάλαιο 3 εξετάζει τις βασικές έννοιες στον τραπεζικό κλάδο με ιδιαίτερη έμφαση στην έννοια του ρίσκου και την κατηγοριοποίηση αυτού. Στο Κεφάλαιο 4, παρουσιάζουμε τις προτεινόμενες τεχνικές και αλγορίθμους που συνδυάζουν ακρίβεια

και ερμηνευσιμότητα στην ανίχνευση απάτης μέσω πιστωτικών καρτών που ανήκουν στον κλάδο της μηχανικής μάθησης (Machine Learning). Στο Κεφάλαιο 5, παρουσιάζουμε τις προτεινόμενες τεχνικές και αλγορίθμους που συνδυάζουν ακρίβεια και ερμηνευσιμότητα στην ανίχνευση απάτης μέσω πιστωτικών καρτών που ανήκουν στον κλάδο της βαθιάς μάθησης (Deep Learning). Στο Κεφάλαιο 6 παρουσιάζουμε το Θεωρητικό πλαίσιο και τις σχετικές υποθέσεις που υιοθετήθηκαν. Στο Κεφάλαιο 7 παρουσιάζονται τα ερευνητικά κενά και οι μελλοντικές ερευνητικές ευκαιρίες στον εν λόγω κλάδο. Στο Κεφάλαιο 8 περιγράφουμε τα πειραματικά αποτελέσματα που προκύπτουν από την εφαρμογή των προτεινομένων μοντέλων σε πραγματικά δεδομένα ανίχνευσης απάτης μέσω πιστωτικών καρτών. Στο Κεφάλαιο 9 παρουσιάζεται εκτενώς η καινοτομία της διατριβής, ενώ στα Κεφάλαια 10 και 11 πραγματοποιείται συζήτηση των συμπερασμάτων που αναλύθηκαν εκτενώς στα προηγούμενα κεφάλαια.



## Κεφάλαιο 2

# Θεωρητικό Υπόβαθρο και Συστημική Βιβλιογραφική Ανασκόπηση

### 2.1 Εισαγωγή

Η μηχανική μάθηση έχει ολοένα και μεγαλύτερη επιρροή στις επιχειρηματικές εφαρμογές, με πολλές λύσεις να έχουν ήδη εφαρμοστεί και πολλές ακόμη να διερευνώνται. Μετά την παγκόσμια χρηματοπιστωτική κρίση, η διαχείριση των κινδύνων στις τράπεζες έχει αποκτήσει μεγαλύτερη σημασία και υπάρχει συνεχής εστίαση γύρω από τον τρόπο με τον οποίο ανιχνεύονται, μετρώνται, αναφέρονται και διαχειρίζονται οι κίνδυνοι. Σημαντική έρευνα στον ακαδημαϊκό και βιομηχανικό χώρο έχει επικεντρωθεί στις εξελίξεις στον τραπεζικό τομέα, στη διαχείριση κινδύνων και στις τρέχουσες και αναδυόμενες προκλήσεις. Η παρούσα εργασία, διατρέχοντας παράλληλα τη διαθέσιμη βιβλιογραφία, επιδιώκει να αναλύσει και να αξιολογήσει τις τεχνικές μηχανικής μάθησης που έχουν ερευνηθεί στο πλαίσιο της διαχείρισης τραπεζικών κινδύνων και να εντοπίσει τομείς ή προβλήματα στη διαχείριση κινδύνων που δεν έχουν διερευνηθεί επαρκώς και αποτελούν δυνητικά πεδία για περαιτέρω έρευνα. Η ανασκόπηση έδειξε ότι η έχει διερευνηθεί η εφαρμογή της μηχανικής μάθησης στη διαχείριση τραπεζικών κινδύνων, όπως ο πιστωτικός κίνδυνος, ο κίνδυνος αγοράς, ο λειτουργικός κίνδυνος και ο κίνδυνος ρευστότητας - ωστόσο, δεν έχει διερευνηθεί επαρκώς. Στη διαχείριση των τραπεζικών κινδύνων παραμένει μεγάλος αριθμός τομέων που θα μπορούσαν να επωφεληθούν σημαντικά από τη σε βάθος μελέτη του τρόπου με τον οποίο η μηχανική μάθηση θα μπορούσε να εφαρμοστεί για την αντιμετώπιση συγκεκριμένων προβλημάτων.

Μετά την παγκόσμια χρηματοπιστωτική κρίση, η διαχείριση των κινδύνων στις τράπεζες

έχει αποκτήσει μεγαλύτερη σημασία και δίνεται διαρκώς έμφαση στον τρόπο με τον οποίο ανιχνεύονται, μετρώνται, αναφέρονται και διαχειρίζονται οι κίνδυνοι. Σημαντικές έρευνες<sup>123 4 5</sup>, τόσο στον ακαδημαϊκό χώρο όσο και στον κλάδο, έχουν επικεντρωθεί στις εξελίξεις στον τραπεζικό τομέα και στη διαχείριση κινδύνων και στις τρέχουσες και αναδυόμενες προκλήσεις. Παράλληλα, έχει αυξηθεί η επιρροή της μηχανικής μάθησης στις επιχειρηματικές εφαρμογές, με πολλές λύσεις να έχουν ήδη εφαρμοστεί και πολλές ακόμη να διερευνώνται.

Η McKinsey & Co υπογράμμισε ότι οι λειτουργίες κινδύνου στις τράπεζες, έως το 2025<sup>6</sup>, θα πρέπει να είναι ριζικά διαφορετικές απ' ό,τι είναι σήμερα. Η διεύρυνση και η εμπάθυνση των κανονισμών, οι εξελισσόμενες προσδοκίες των πελατών και η εξέλιξη των τύπων κινδύνου αναμένεται να οδηγήσουν στην αλλαγή πλαισίου της διαχείρισης κινδύνων. Τα νέα προϊόντα, οι υπηρεσίες και οι τεχνικές διαχείρισης κινδύνων ενεργοποιούνται μέσω της εφαρμογής των εξελισσόμενων τεχνολογιών και των προηγμένων αναλύσεων. Η μηχανική μάθηση, η οποία προσδιορίζεται ως μία από τις σημαντικά αναπτυσσόμενες τεχνολογίες στη διαχείριση κινδύνων, μπορεί να επιτρέψει τη δημιουργία ακριβέστερων μοντέλων κινδύνου με τον εντοπισμό σύνθετων, μη γραμμικών προτύπων μέσα σε μεγάλα σύνολα δεδομένων. Η προβλεπτική ικανότητα αυτών των μοντέλων μπορεί να αυξάνεται με κάθε πληροφορία που προστίθεται, ενισχύοντας έτσι την προβλεπτική τους ικανότητα με την πάροδο του χρόνου. Αναμένεται ότι η μηχανική μάθηση θα εφαρμοστεί σε πολλούς τομείς εντός του οργανισμού κινδύνου μιας τράπεζας. Η μηχανική μάθηση έχει επίσης προταθεί ως πρωτοβουλία που θα μπορούσε να βοηθήσει στον μετασχηματισμό της λειτουργίας διαχείρισης κινδύνων στις τράπεζες.

Το παρόν κεφάλαιο επιδιώκει να μελετήσει το βαθμό στον οποίο η μηχανική μάθηση, η οποία έχει επισημανθεί ως αναδυόμενος επιχειρηματικός παράγοντας, έχει ερευνηθεί στο πλαίσιο της διαχείρισης κινδύνων στον τραπεζικό κλάδο και, στη συνέχεια, να προσδιορίσει πιθανά πεδία για περαιτέρω έρευνα. Σκοπός της παρούσας διατριβής είναι να αξιολογήσει, να αναλύσει και να αξιολογήσει τις τεχνικές μηχανικής μάθησης που έχουν εφαρμοστεί στη διαχείριση τραπεζικών κινδύνων και να εντοπίσει τομείς ή προβλήματα στη διαχείριση

---

<sup>1</sup>Van Liebergen 2017

<sup>2</sup>Deloitte University Press 2017

<sup>3</sup>Helbekkmo et al. 2013

<sup>4</sup>MetricStream 2018

<sup>5</sup>Oliver Wyman 2017

<sup>6</sup>[https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/the-future-of-bank-risk-](https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/the-future-of-bank-risk-management)

κινδύνων που δεν έχουν διερευνηθεί επαρκώς και να διατυπώσει προτάσεις για περαιτέρω έρευνα.

Για τον προσδιορισμό των κινδύνων που αφορούν κυρίως τις τράπεζες, παράλληλα με την αξιοποίηση της υπάρχουσας βιβλιογραφίας, η παρούσα εργασία παρέχει μια ταξινόμηση των κινδύνων που αναπτύσσονται με βάση την επισκόπηση των ετήσιων εκθέσεων των τραπεζών. Πραγματοποιήθηκε ανάλυση της διαθέσιμης βιβλιογραφίας για την αξιολόγηση των τομέων της διαχείρισης τραπεζικών κινδύνων, οι οποίοι έχουν υποβληθεί σε τεχνικές μηχανικής μάθησης. Η έρευνα αξιολόγησε τους τομείς κινδύνου στους οποίους έχει εφαρμοστεί η μηχανική μάθηση και τη συγκεκριμένη μεθοδολογία κινδύνου που αντιμετώπισαν. Η ανάλυση προσδιόρισε επίσης τους αλγόριθμους μηχανικής μάθησης που χρησιμοποιούνται, τόσο για συγκεκριμένους τομείς όσο και σε ευρύτερο πλαίσιο.

## 2.2 Υλικά Και Μέθοδοι Βιβλιογραφικής Ανασκόπησης

Για να πραγματοποιηθεί η ανασκόπηση της βιβλιογραφίας που ερευνά την εφαρμογή της μηχανικής μάθησης στη διαχείριση τραπεζικών κινδύνων, χρησιμοποιήθηκαν δύο ομάδες λέξεων-κλειδιών για την αναζήτηση σχετικών εργασιών. Η αναζήτηση των εργασιών έγινε με τη χρήση των βάσεων δεδομένων scholar.google.com, SSRN και ProQuest. Η αναζήτηση επικεντρώθηκε σε μεγάλο βαθμό σε εργασίες μετά το 2007 για να καταγραφούν οι εξελίξεις μετά την παγκόσμια χρηματοπιστωτική κρίση- ωστόσο, συμπεριλήφθηκαν και εργασίες πριν από αυτή την περίοδο, εφόσον γινόταν αναφορά σε άλλες πρόσφατες εργασίες.

Η πρώτη ομάδα λέξεων ήταν η "μηχανική μάθηση", σύμφωνα με το θέμα. Η δεύτερη ομάδα περιλάμβανε όρους που εντοπίστηκαν από την εξέταση των ετήσιων εκθέσεων των τραπεζών. Αυτό περιλαμβάνει τύπους κινδύνου, όπως αναφέρονται στην ταξινόμηση κινδύνου και εργαλεία ή μεθόδους διαχείρισης κινδύνου που εντοπίστηκαν από τις ετήσιες εκθέσεις των τραπεζών. Η ταξινόμηση είναι όπως φαίνεται στο Σχήμα 2.1 και οι μέθοδοι όπως στο Σχήμα 2.2.

Η αναζήτηση και η ανασκόπηση περιορίστηκε σε δημοσιεύσεις σε συνέδρια, άρθρα σε περιοδικά και επιλεγμένες εργασίες (μεταπτυχιακές ή διδακτορικές διατριβές). Η ανασκόπηση δεν έλαβε υπόψη άρθρα, λευκά έγγραφα, έγγραφα πωλητών ή διαδικτυακά άρθρα που απλώς έκαναν αναφορά στη μηχανική μάθηση χωρίς να παρέχουν λεπτομέρειες για τον τρόπο, ή που έκαναν αναφορές στην εφαρμογή κάποιου συγκεκριμένου αλγορίθμου, αν και

πολλά τέτοια άρθρα προέκυψαν κατά την αναζήτηση. Ειδικότερα, υπάρχει μεγάλος αριθμός άρθρων (διαδικτυακών και περιοδικών) και δημοσιεύσεων που περιλαμβάνουν τη μηχανική μάθηση ως λύση ή ως γενική σύσταση χωρίς να παρέχουν περαιτέρω λεπτομέρειες για το πώς μπορεί να αντιμετωπιστεί ένα συγκεκριμένο πρόβλημα. Η ανασκόπηση εξέτασε μόνο τα άρθρα που έχουν αναλύσει το θέμα με ένα επίπεδο βάθους, δηλαδή κάνοντας αναφορές σε συγκεκριμένους αλγόριθμους ή παρέχοντας ένα σχέδιο ή μοντέλο για το πώς μπορεί να εφαρμοστεί η μηχανική μάθηση. Άρθρα ή εργασίες ή πρακτικά συνεδρίων που έκαναν μόνο μια επιφανειακή ή γενική αναφορά στην εφαρμογή της ML στο χώρο της διαχείρισης κινδύνων δεν ελήφθησαν υπόψη για την παρούσα έρευνα. Σημειώνεται ότι υπάρχουν πολλές διαθέσιμες αναφορές όπου οι συγγραφείς ή οι ομιλητές έχουν προτείνει ότι η ML ή η TN μπορεί να εφαρμοστεί στη διαχείριση κινδύνων- ωστόσο, πολλές από αυτές δεν παρέχουν σαφήνεια σχετικά με το ποιους αλγόριθμους ή δεν παρέχουν παραδείγματα για το πώς η ML/AI έχει εφαρμοστεί σε μια δοκιμαστική ή βιομηχανική εγκατάσταση.

Το μεθοδολογικό πλαίσιο για την παρούσα έρευνα καθορίστηκε με την ανάλυση των διαφόρων προβληματικών περιοχών που σχετίζονται με τη μηχανική μάθηση και τη διαχείριση κινδύνων στις τράπεζες. Τα άρθρα ταξινομήθηκαν για να γίνει κατανοητό: i) ο τομέας κινδύνου στον οποίο εστίαζαν, ii) το εργαλείο διαχείρισης κινδύνου ή το στοιχείο του πλαισίου διαχείρισης κινδύνου στο οποίο στόχευαν, ή iii) οι αλγόριθμοι που εφαρμόστηκαν/μελετήθηκαν/προτάθηκαν. Η έρευνα επιδίωκε επίσης να εξετάσει άρθρα που επικεντρώνονταν περισσότερο στην αξιολόγηση και τη μέτρηση του κινδύνου.

Τομείς κινδύνου, όπως η ασφάλεια στον κυβερνοχώρο και ο κίνδυνος απάτης, έχουν εξεταστεί ευρέως- ωστόσο, η παρούσα ανασκόπηση επικεντρώθηκε μόνο σε περιπτώσεις όπου σχετίζονται συγκεκριμένα με περιπτώσεις χρήσης διαχείρισης τραπεζικών κινδύνων. Δεν ελήφθησαν υπόψη έγγραφα που εστιάζουν την έρευνα σε λειτουργικά θέματα, όπως λύσεις διαχείρισης πιστωτικού κινδύνου που αφορούν τη λειτουργική διαδικασία πιστωτικού ελέγχου και έγκρισης, ή εργαλεία που επικεντρώνονται στην υποστήριξη των εμπόρων και των διαχειριστών κινδύνου συναλλαγών στη διαδικασία διαχείρισης εντολών και συναλλαγών. Επιπλέον, δεν έχουν ερευνηθεί λύσεις διαχείρισης λειτουργικού κινδύνου που εντάσσονται στη λειτουργική διαδικασία για τον μετριασμό των λειτουργικών συμβάντων/περιστατικών (π.χ. αυτοματοποίηση ρομποτικών διαδικασιών, ανίχνευση ανωμαλιών, κ.ά.).

## 2.3 Συζήτηση Βιβλιογραφικής Ανασκόπησης

Η βαθμολόγηση της πιστοληπτικής ικανότητας περιλαμβάνει την απόδοση μιας αριθμητικής τιμής στην επιχείρηση (ή τον πελάτη), η οποία υποδεικνύει εάν η επιχείρηση (ή ο πελάτης) είναι πιθανό να αθετήσει ή όχι τις υποχρεώσεις της. Το μεγαλύτερο μέρος της έρευνας έχει επικεντρωθεί στην αντιμετώπιση αυτού του τομέα με τον χειρισμό του ως προβλήματος ταξινόμησης, προβλέποντας έναν υποψήφιο πελάτη ως "καλό" ή "κακό", ώστε να διευκολύνεται η λήψη πιστωτικών αποφάσεων και η διαχείριση του πιστωτικού κινδύνου. Ως εκ τούτου, κυριαρχούν οι αλγόριθμοι που σχετίζονται με την ταξινόμηση. Πολλές εργασίες έχουν προχωρήσει πέρα από την απλή ταξινόμηση των πιστωτών και έχουν ασχοληθεί με τον τρόπο πρόβλεψης της πιθανότητας αθέτησης ή του ποσοστού ανάκτησης, δηλαδή με την εκτίμηση της πιθανότητας αθέτησης (PD), της ζημίας λόγω αθέτησης (LGD), της έκθεσης σε αθέτηση (EAD). Θα είχε σημαντική αξία για τις τράπεζες και τις λειτουργίες διαχείρισης κινδύνου τους να διαθέτουν έρευνα και μοντέλα για την πρόβλεψη ή την εκτίμηση της PD, LGD, EAD, και συνεπώς την εκτίμηση της έκθεσης σε πιστωτικές ζημιές.

Οι τεχνικές μηχανικής μάθησης έχουν αποδειχθεί ότι αποδίδουν καλύτερα από τις παραδοσιακές στατιστικές τεχνικές, τόσο στην ταξινόμηση όσο και στην ακρίβεια πρόβλεψης. Η μηχανή διανύσματος υποστήριξης θεωρείται ότι είναι μια ευρέως δοκιμασμένη και αποδεδειγμένη προσέγγιση μηχανικής μάθησης. Πολλές εμπειρικές εργασίες βασίζονται σε δεδομένα παρατήρησης. Οι μηχανισμοί επιλογής θα μπορούσαν να οδηγήσουν σε μη τυχαία δείγματα, είτε λόγω του σχεδιασμού του δείγματος είτε λόγω της συμπεριφοράς των δειγματοληπτικών μονάδων. Στην πρώτη περίπτωση, τα δεδομένα συνήθως λείπουν, ενώ στη δεύτερη περίπτωση, υπάρχει αυτοεπιλογή των μονάδων του δείγματος. Αυτό οδηγεί σε μεροληψία επιλογής δείγματος (Arezzo and Guagnano 2018). Όταν τα δεδομένα του δείγματος που μελετάται έχουν αναλογική εκπροσώπηση ορισμένων αποτελεσμάτων εξαρτημένης μεταβλητής (π.χ. αθέτηση πληρωμών, δόλια συναλλαγή) διαφορετική από την αναλογική τους εκπροσώπηση στον πληθυσμό από τον οποίο αντλούνται, λέγεται ότι είναι "βασισμένα σε επιλογές". Αυτή η δειγματοληψία "βασισμένη στην επιλογή" προκαλεί μεροληψία στην εκτίμηση (Greene 1992). Καθώς η μηχανική μάθηση βασίζεται σε μεγάλο μέρος της μοντελοποίησης στη μάθηση από τα διαθέσιμα δεδομένα, θα μπορούσε να είναι επιρρεπής στα ίδια προβλήματα και μεροληψίες που επηρεάζουν τις παραδοσιακές στατιστικές μεθόδους. Καθώς οι μέθοδοι μηχανικής μάθησης συγκρίνονται με τις παραδοσιακές στατιστικές τεχνικές, θα ήταν χρήσιμο να αξιολογηθεί και να κατανοηθεί πώς αντιμετωπίζονται τα προβλήματα

που είναι εγγενή στις παραδοσιακές στατιστικές μεθόδους έρευνας όταν αντιμετωπίζονται από τις τεχνικές μηχανικής μάθησης.

Η εφαρμογή της μηχανικής μάθησης για τη διαχείριση του κινδύνου της αγοράς δεν φαίνεται επίσης να έχει μελετηθεί επαρκώς. Πολλές εργασίες έχουν ερευνήσει τη μεταβλητότητα της αγοράς ή τον κίνδυνο της αγοράς από την οπτική γωνία της διαχείρισης χαρτοφυλακίου ή επενδυτικού κινδύνου. Ωστόσο, από την άποψη της διαχείρισης τραπεζικού κινδύνου, οι εργασίες εμφανίζονται περιορισμένες. Αξίζει να γίνει περαιτέρω έρευνα για το θέμα αυτό, ιδίως από την άποψη των δοκιμών αντοχής ή της σύλληψης του κινδύνου στο τέλος της αγοράς. Για τον κίνδυνο ρευστότητας, ο οποίος, μετά τη χρηματοπιστωτική κρίση, έχει προσελκύσει μεγάλη προσοχή από τις ρυθμιστικές αρχές, έχουν ερευνηθεί λίγες περιπτώσεις χρήσης.

Στον τομέα του λειτουργικού κινδύνου, οι μελέτες έχουν επικεντρωθεί κυρίως στην ανίχνευση απάτης και ύποπτων συναλλαγών - προβλήματα που συνήθως αντιμετωπίζονται με αλγορίθμους ταξινόμησης. Η ανάλυση ομαδοποίησης, τα δίκτυα Bayes, τα δέντρα απόφασης, τα δέντρα ταξινόμησης, τα S.V.M. σημειώνονται συνήθως στην εφαρμογή αλγορίθμων μηχανικής μάθησης. Τα νευρωνικά δίκτυα έχουν επίσης αναφερθεί ως μια πολύ διαδεδομένη και εξέχουσα τεχνική στην ανίχνευση απάτης μέσω πιστωτικών καρτών. Πολλοί από τους άλλους τομείς της διαχείρισης μη χρηματοοικονομικών κινδύνων, της διαχείρισης κινδύνου χώρας, της διαχείρισης κινδύνου συμμόρφωσης -εκτός από τις χρήσεις που σχετίζονται με το ξέπλυμα χρήματος- δεν έχουν διερευνηθεί επαρκώς. Ο κίνδυνος συμπεριφοράς, ο οποίος έχει καταστεί βασικός κίνδυνος και έχει αποκτήσει υψηλή προτεραιότητα για τις ρυθμιστικές αρχές και τις τράπεζες λόγω της έξαρσης των ζητημάτων συμπεριφοράς στην Ευρώπη, τις ΗΠΑ και την Ασία-Ειρηνικό, φαίνεται επίσης να λείπει από τις ερευνητικές εργασίες, αν και η μηχανική μάθηση προσφέρεται ως λύση για τη διαχείριση του κινδύνου συμπεριφοράς (Oliver Wyman 2017). Επιπλέον, πολλά από τα τυποποιημένα εργαλεία διαχείρισης κινδύνου, όπως οι αξιολογήσεις κινδύνου, η παρακολούθηση κινδύνου και η υποβολή εκθέσεων κινδύνου, εμφανίζονται ως τομείς που θα μπορούσαν να επωφεληθούν σημαντικά από περαιτέρω έρευνα.

Οι πηγές δεδομένων που χρησιμοποιούν οι ερευνητές για τις μελέτες τους ποικίλλουν, με ορισμένους να λαμβάνουν δεδομένα από εμπορικές τράπεζες ή από βάσεις δεδομένων που παρέχονται από παρόχους χρηματοπιστωτικών υπηρεσιών (π.χ. Moody's), και με ορισμένους ερευνητές να διεξάγουν τις μελέτες τους χρησιμοποιώντας δημόσια διαθέσιμα δεδομένα. Τα προβλήματα που επισημάνθηκαν είναι ότι τα δεδομένα μπορεί να είναι ιδιαίτερα στρεβλά

(π.χ. χαμηλά ποσοστά αθέτησης), η μη πληρότητα των δεδομένων και η ακεραιότητα των δεδομένων (π.χ. μπορεί να μην έχουν επισημανθεί σωστά ή όπως αναμενόταν (Khandani et al. 2010)). Μια ευρύτερη διαθεσιμότητα συνόλων δεδομένων του πραγματικού κόσμου θα ενθάρρυνε σίγουρα περισσότερη έρευνα για την αξιολόγηση των πολλών προβλημάτων που αντιμετωπίζουν οι λειτουργίες διαχείρισης κινδύνου.

## 2.4 Επίλογος Βιβλιογραφικής Ανασκόπησης

Το μέλλον της μηχανικής μάθησης στον τραπεζικό και χρηματοπιστωτικό τομέα είναι καλά αναγνωρισμένο και αναμένεται ότι ο τομέας της διαχείρισης κινδύνων θα επιδιώξει επίσης να εφαρμόσει τεχνικές μηχανικής μάθησης για να ενισχύσει τις δυνατότητές του. Παρά την κριτική που δέχεται για το γεγονός ότι λειτουργεί σαν μαύρο κουτί, η ικανότητα των τεχνικών μηχανικής μάθησης να αναλύουν όγκους δεδομένων χωρίς να περιορίζονται από υποθέσεις κατανομής και να παρέχουν μεγάλη αξία στη διερευνητική ανάλυση, την ταξινόμηση και την προγνωστική ανάλυση, είναι σημαντική. Αυτό προσφέρει τη δυνατότητα μετασχηματισμού του τομέα της διαχείρισης κινδύνων. Η μηχανική μάθηση, η οποία αναγνωρίστηκε ως μία από τις τεχνολογίες με σημαντικές επιπτώσεις στη διαχείριση κινδύνων, μπορεί να επιτρέψει τη δημιουργία ακριβέστερων μοντέλων κινδύνου με τον εντοπισμό σύνθετων, μη γραμμικών προτύπων σε μεγάλα σύνολα δεδομένων. Η παρούσα διατριβή παρουσιάζει μια εκτίμηση, ανάλυση και αξιολόγηση της έρευνας γύρω από την εφαρμογή της μηχανικής μάθησης στη διαχείριση κινδύνων στον τραπεζικό κλάδο. Το μεγαλύτερο μέρος της έρευνας φαίνεται να επικεντρώνεται στη διαχείριση του πιστωτικού κινδύνου. Αυτό θα μπορούσε να αποδοθεί στο γεγονός ότι οι πιστωτικοί κίνδυνοι θεωρούνται ο σημαντικότερος κίνδυνος για έναν τραπεζικό οργανισμό. Πιο συγκεκριμένα, από άποψη μεθοδολογίας, τα προβλήματα διαχείρισης πιστωτικού κινδύνου που έχουν ερευνηθεί ήταν γύρω από την πιστωτική βαθμολόγηση- θα ήταν πολύ σημαντικό να ερευνηθεί πώς η μηχανική μάθηση μπορεί να εφαρμοστεί σε ποσοτικούς τομείς για καλύτερους υπολογισμούς της έκθεσης σε πιστωτικό κίνδυνο με την πρόβλεψη των πιθανοτήτων αθέτησης, της ζημίας λόγω αθέτησης, δεδομένων των πολλών περιπλοκών και των ποικίλων παραγόντων που εμπλέκονται. Στον κίνδυνο αγοράς έχει γίνει κάποια έρευνα με τη μηχανική μάθηση να χρησιμοποιείται για την πρόβλεψη της μεταβλητότητας, των καμπυλών επιτοκίων και της αλλαγής του καθεστώτος της αγοράς. Ο κίνδυνος ρευστότητας, παρά την αυξημένη προσοχή του κλάδου

μετά την ανησυχία των ρυθμιστικών αρχών, έχει τύχει περιορισμένης έρευνας. Λαμβάνοντας υπόψη τις επιπτώσεις στην κερδοφορία και τη φερεγγυότητα μιας τράπεζας ως συνέπεια της υλοποίησης ενός γεγονότος κινδύνου ρευστότητας, ο κίνδυνος ρευστότητας θα ήταν ένας πολύ καλός υποψήφιος για εκτεταμένη έρευνα, και πιο συγκεκριμένα για έρευνα γύρω από την πρόβλεψη γεγονότων κινδύνου ρευστότητας μεμονωμένα ή ως δίκτυο παραγόντων ή γεγονότων. Ο λειτουργικός κίνδυνος έχει επίσης ερευνηθεί πολύ περιορισμένα. Η έρευνα έχει επικεντρωθεί στην εφαρμογή της μηχανικής μάθησης για την πρόληψη ή τον εντοπισμό συμβάντων λειτουργικού κινδύνου- ωστόσο, υπάρχει πολύ περιορισμένη εφαρμογή σε πρακτικές διαχείρισης του λειτουργικού κινδύνου, ιδίως στους τομείς του εντοπισμού, της αξιολόγησης, της παρακολούθησης και της υποβολής εκθέσεων. Δεδομένου του τεράστιου όγκου των διαθέσιμων επιχειρησιακών δεδομένων (εσωτερικά σε μια τράπεζα), η μηχανική μάθηση θα μπορούσε να εφαρμοστεί στην αποτελεσματική ανάπτυξη των δυνατοτήτων διαχείρισης του επιχειρησιακού κινδύνου, η οποία βασίζεται κυρίως σε ποιοτικούς παράγοντες για τη μέτρηση, την αναφορά και τη διαχείριση του κινδύνου. Η επισκόπηση έδειξε ότι οι τομείς των δοκιμών αντοχής, της καταγραφής του κινδύνου ουράς, της ανάλυσης σεναρίων - τομείς που βασίζονται στην προγνωστική ανάλυση μεγάλου όγκου δεδομένων - έχουν επίσης δει μόνο περιορισμένη έρευνα. Τα πλεονεκτήματα και τα μειονεκτήματα των διαφόρων τεχνικών μηχανικής μάθησης στην επίλυση συγκεκριμένων προβλημάτων διαχείρισης κινδύνου μπορούν επίσης να αξιολογηθούν και να μελετηθούν περαιτέρω για τη μεγιστοποίηση της αξίας.

Η ανασκόπηση έδειξε ότι έχει διερευνηθεί η εφαρμογή της μηχανικής μάθησης στη διαχείριση τραπεζικών κινδύνων, όπως ο πιστωτικός κίνδυνος, ο κίνδυνος αγοράς, ο λειτουργικός κίνδυνος και ο κίνδυνος ρευστότητας- ωστόσο, δεν φαίνεται να είναι ανάλογη με το σημερινό επίπεδο εστίασης του κλάδου τόσο στη διαχείριση κινδύνων όσο και στη μηχανική μάθηση. Στους τομείς του κινδύνου αγοράς, του λειτουργικού κινδύνου και του κινδύνου ρευστότητας η έρευνα φαίνεται να είναι ελλιπής και υπάρχει σημαντικό δυναμικό για περαιτέρω μελέτη. Η εφαρμογή της μηχανικής μάθησης θα μπορούσε να διερευνηθεί περαιτέρω για ορισμένους τομείς όπου απαιτείται ανάλυση ή μοντελοποίηση σε όγκους δεδομένων με πολύπλοκους και μη γραμμικούς υπολογισμούς. Η μέτρηση και η αναφορά του τεχνολογικού κινδύνου είναι ακόμη ένας νέος τομέας και θα μπορούσε να διερευνηθεί περαιτέρω, ιδίως καθώς ο κίνδυνος αυτός ανεβαίνει στους πίνακες και τα ανώτερα στελέχη και οι υπεύθυνοι κινδύνου στις τράπεζες αρχίζουν να αναζητούν περισσότερη εικόνα για το ποιος είναι ο

τεχνολογικός κίνδυνος. Καθώς οι τράπεζες επιδιώκουν να ωριμάσουν τις δυνατότητες διαχείρισης επιχειρηματικών κινδύνων, θα ήταν ωφέλιμο να μελετήσουν τον τρόπο με τον οποίο η μηχανική μάθηση μπορεί να εφαρμοστεί στη συγκέντρωση των κινδύνων και στην ενίσχυση των δυνατοτήτων αναφοράς κινδύνων. Ενώ θα μπορούσαν επίσης να διερευνηθούν τομείς όπως ο κίνδυνος συμπεριφοράς, σημειώνεται ότι οι τομείς αυτοί θα επωφελούνταν περισσότερο από την εφαρμογή στον επιχειρησιακό τομέα, όπως η παρακολούθηση της συμπεριφοράς και η παρακολούθηση της δραστηριότητας. Αν και αυτά συμβάλλουν στη διαχείριση των κινδύνων (μετριασμός κινδύνων, ανίχνευση κινδύνων) στην τράπεζα, δεν αποτελούν τα συστήματα διαχείρισης κινδύνων (μέτρηση κινδύνων, αξιολόγηση κινδύνων) στα οποία επικεντρώνεται η παρούσα έρευνα. Εν κατακλείδι, ενώ έχει γίνει έρευνα σχετικά με την εφαρμογή της μηχανικής μάθησης στη διαχείριση κινδύνων με την πάροδο των ετών, εξακολουθεί να υπολείπεται και να μην είναι ισότιμη σε όλους τους διάφορους τομείς της διαχείρισης κινδύνων ή των μεθοδολογιών κινδύνου. Εξακολουθεί να υπάρχει μεγάλος αριθμός τομέων, όπως τονίστηκε παραπάνω, στη διαχείριση τραπεζικών κινδύνων που θα μπορούσαν να επωφεληθούν σημαντικά από τη μελέτη του τρόπου εφαρμογής της μηχανικής μάθησης για την αντιμετώπιση συγκεκριμένων προβλημάτων.



## Κεφάλαιο 3

# Διαχείριση Ρίσκου στον Τραπεζικό Τομέα

Η επιδίωξη της διοίκησης μιας τράπεζας να αυξήσει τις αποδόσεις για τους μετόχους της έχει ως κόστος την αύξηση του κινδύνου. Οι τράπεζες αντιμετωπίζουν διάφορους κινδύνους - κίνδυνο επιτοκίου, κίνδυνο αγοράς, πιστωτικό κίνδυνο, κίνδυνο εκτός ισολογισμού, τεχνολογικό και λειτουργικό κίνδυνο, συναλλαγματικό κίνδυνο, κίνδυνο χώρας ή κράτους, κίνδυνο ρευστότητας, κίνδυνο ρευστότητας και κίνδυνο αφερεγγυότητας. Η αποτελεσματική διαχείριση αυτών των κινδύνων είναι το κλειδί για την απόδοση μιας τράπεζας. Επίσης, δεδομένων αυτών των κινδύνων και του ρόλου που διαδραματίζουν οι τράπεζες στα χρηματοπιστωτικά συστήματα, υπόκεινται σε ρυθμιστική προσοχή [59]. Οι ρυθμιστικές αρχές απαιτούν από τις τράπεζες να διαθέτουν κεφάλαια για τους πολλούς κινδύνους που προκύπτουν και μεταφέρονται λόγω των ποικίλων δραστηριοτήτων μιας τράπεζας. Τα πρότυπα της Βασιλείας για τον προσδιορισμό των κεφαλαιακών απαιτήσεων αναπτύχθηκαν το 1998 και έκτοτε έχουν εξελιχθεί και εξελίσσεται. Απαιτείται κεφάλαιο για κάθε έναν από τους κύριους τύπους κινδύνου. Ο πιστωτικός κίνδυνος είναι παραδοσιακά ο μεγαλύτερος κίνδυνος που αντιμετωπίζουν οι τράπεζες και συνήθως αυτός που απαιτεί τα περισσότερα κεφάλαια. Ο κίνδυνος αγοράς προκύπτει κυρίως από τις συναλλαγές μιας τράπεζας, ενώ ο λειτουργικός κίνδυνος είναι ο κίνδυνος ζημιών από αστοχίες εσωτερικών συστημάτων ή εξωτερικά γεγονότα. Εκτός από τον υπολογισμό του εποπτικού κεφαλαίου, οι περισσότερες μεγάλες τράπεζες υπολογίζουν επίσης το οικονομικό κεφάλαιο, το οποίο βασίζεται στα μοντέλα μιας τράπεζας και όχι στις προδιαγραφές των ρυθμιστικών αρχών [34]. Οι κυριότεροι κίνδυνοι που αντιμετωπίζουν οι τράπεζες είναι ο πιστωτικός, ο κίνδυνος αγοράς και ο λειτουργικός κίνδυνος, ενώ άλλα είδη κινδύνων περιλαμβάνουν τον κίνδυνο ρευστότητας, τον επιχειρηματικό κίνδυνο και τον κίνδυνο φήμης. Οι τράπεζες ασχολούνται ενεργά με τη διαχείριση

κινδύνων για την παρακολούθηση, τη διαχείριση και τη μέτρηση αυτών των κινδύνων [3].

Ο κίνδυνος αγοράς μπορεί να οριστεί ως ο κίνδυνος ζημιών "λόγω μεταβολών στο επίπεδο ή τη μεταβλητότητα των τιμών της αγοράς" [37]. Ο κίνδυνος αγοράς περιλαμβάνει τον κίνδυνο επιτοκίου, τον κίνδυνο μετοχών, τον συναλλαγματικό κίνδυνο και τον κίνδυνο εμπορευμάτων. Ο κίνδυνος επιτοκίου μπορεί να οριστεί ως η δυνητική ζημία που οφείλεται σε μεταβολές των επιτοκίων. Ο μετοχικός κίνδυνος μπορεί να οριστεί ως η δυνητική ζημία που προκύπτει από μια δυσμενή μεταβολή στην τιμή μιας μετοχής. Ο συναλλαγματικός κίνδυνος μπορεί να οριστεί ως ο κίνδυνος μεταβολής της αξίας των στοιχείων ενεργητικού ή παθητικού μιας τράπεζας λόγω διακυμάνσεων της συναλλαγματικής ισοτιμίας. Ο κίνδυνος εμπορευμάτων μπορεί να οριστεί ως η πιθανή ζημία λόγω δυσμενούς μεταβολής της τιμής των εμπορευμάτων που κατέχει. Το πλαίσιο κινδύνου αγοράς της συμφωνίας της Βασιλείας αποτελείται από μια προσέγγιση εσωτερικών υποδειγμάτων και μια τυποποιημένη προσέγγιση. Για την καλύτερη καταγραφή του κινδύνου ουράς, στο αναθεωρημένο πλαίσιο έγινε επίσης μετατόπιση του μέτρου του κινδύνου υπό πίεση από την αξία κινδύνου (VaR) στην αναμενόμενη ζημία (EL) <sup>1</sup>.

Η πίστωση μπορεί να οριστεί ως ο κίνδυνος δυνητικής ζημίας για την τράπεζα εάν ένας δανειολήπτης δεν εκπληρώσει τις υποχρεώσεις του (τόκοι, ποσά κεφαλαίου). Ο πιστωτικός κίνδυνος είναι ο μεγαλύτερος κίνδυνος που αντιμετωπίζουν οι τράπεζες [3]. Το Σύμφωνο της Βασιλείας επιτρέπει στις τράπεζες να εφαρμόζουν την προσέγγιση των εσωτερικών διαβαθμίσεων για τον πιστωτικό κίνδυνο. Οι τράπεζες μπορούν να αναπτύξουν εσωτερικά τα δικά τους μοντέλα πιστωτικού κινδύνου για τον υπολογισμό της αναμενόμενης ζημίας. Οι βασικές παράμετροι κινδύνου που πρέπει να εκτιμηθούν είναι η πιθανότητα αθέτησης (PD), η ζημία λόγω αθέτησης (LGD) και η έκθεση σε αθέτηση (EAD).

$$EL = PD * LGD * EAD \quad (3.1)$$

<sup>2</sup>.

Ο κίνδυνος ρευστότητας, ο οποίος αντιμετωπίζεται χωριστά από τους άλλους κινδύνους, λαμβάνει δύο μορφές - τον κίνδυνο ρευστότητας περιουσιακών στοιχείων και τον κίνδυνο ρευστότητας χρηματοδότησης. Μια τράπεζα εκτίθεται σε κίνδυνο ρευστότητας περιουσιακών στοιχείων όταν μια συναλλαγή δεν μπορεί να εκτελεστεί στις τρέχουσες τιμές της αγο-

<sup>1</sup> Επιτροπή Τραπεζικής Εποπτείας της Βασιλείας 2006

<sup>2</sup> Επιτροπή Τραπεζικής Εποπτείας της Βασιλείας 2005α, 2005β

ράς, γεγονός που μπορεί να είναι συνέπεια του μεγέθους της θέσης σε σχέση με το κανονικό μέγεθος της παρτίδας διαπραγμάτευσης. Ο κίνδυνος ρευστότητας χρηματοδότησης αναφέρεται στην αδυναμία εκπλήρωσης των ταμειακών υποχρεώσεων και είναι επίσης γνωστός ως κίνδυνος ταμειακών ροών [37]. Οι τράπεζες υποχρεούνται να θεσπίσουν ένα ισχυρό πλαίσιο διαχείρισης του κινδύνου ρευστότητας που θα διασφαλίζει τη διατήρηση επαρκούς ρευστότητας, συμπεριλαμβανομένης της ικανότητας να αντέχουν σε μια σειρά γεγονότων πίεσης. Θα πρέπει να εφαρμόζεται μια υγιής διαδικασία για τον εντοπισμό, τη μέτρηση, την παρακολούθηση και τον έλεγχο του κινδύνου ρευστότητας <sup>3</sup>.

Ο λειτουργικός κίνδυνος ορίζεται από την BCBS ως ο κίνδυνος ζημίας που προκύπτει από "ανεπαρκείς ή αποτυχημένες εσωτερικές διαδικασίες, ανθρώπους και συστήματα ή από εξωτερικά γεγονότα" και αποτελεί "θεμελιώδες στοιχείο της διαχείρισης κινδύνων" στις τράπεζες. Ο ορισμός αυτός περιλαμβάνει τον νομικό κίνδυνο, αλλά αποκλείει τον στρατηγικό κίνδυνο και τον κίνδυνο φήμης. Θεωρείται εγγενής σε όλα τα τραπεζικά προϊόντα, δραστηριότητες, διαδικασίες και συστήματα <sup>4</sup>. Στις ετήσιες εκθέσεις, ο λειτουργικός κίνδυνος παρουσιάστηκε ποικιλοτρόπως και περιλάμβανε διάφορους επιμέρους κινδύνους, ενώ θα μπορούσε να αναφέρεται περισσότερο ως μη χρηματοοικονομικός κίνδυνος. Περιελάμβανε, μεταξύ πολλών άλλων, τον κίνδυνο απάτης, την ασφάλεια στον κυβερνοχώρο, τα προϊόντα πελατών και τις επιχειρηματικές πρακτικές, τον κίνδυνο πληροφοριών και ανθεκτικότητας, τον κίνδυνο νομιμοποίησης εσόδων από παράνομες δραστηριότητες και τον κίνδυνο οικονομικού εγκλήματος, τους κινδύνους προμηθευτών και εξωτερικής ανάθεσης, τον τεχνολογικό κίνδυνο, τους κινδύνους διακοπής της λειτουργίας. Σε ορισμένες περιπτώσεις, οι τράπεζες ανέφεραν τους κινδύνους συμμόρφωσης και τους νομικούς κινδύνους επίσης στο πλαίσιο του λειτουργικού κινδύνου.

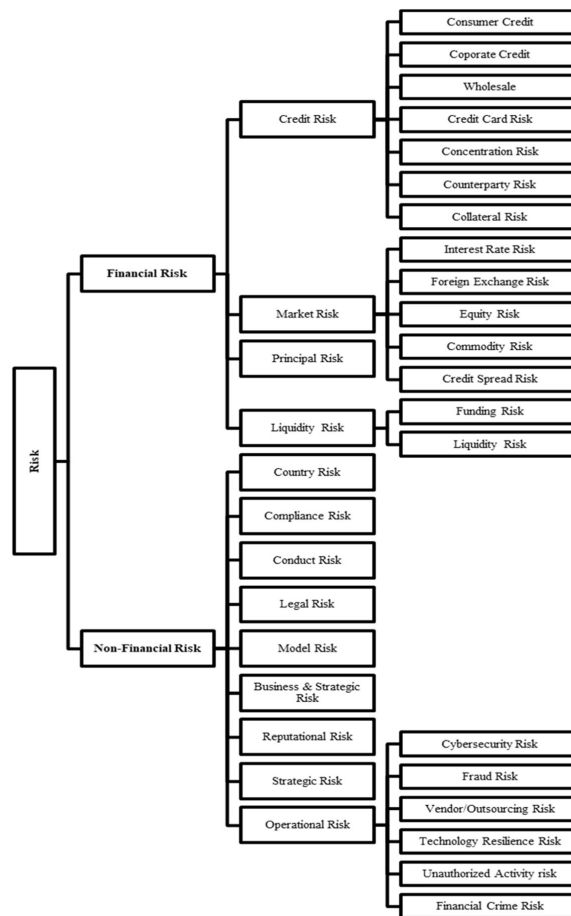
Για να προσδιοριστούν οι κίνδυνοι που αφορούν ειδικά τις τράπεζες, συμπληρωματικά με την υπάρχουσα βιβλιογραφία, έγινε ανασκόπηση των ετήσιων εκθέσεων των τραπεζών. Με βάση την ανασκόπηση, χαρτογραφήθηκε μια ταξινόμηση των διαφόρων τύπων κινδύνων που οι τράπεζες συνήθως επιδιώκουν να διαχειριστούν στο πλαίσιο των δραστηριοτήτων τους και των μεθοδολογιών και εργαλείων που χρησιμοποιούν. Εξετάστηκαν οι ετήσιες εκθέσεις 10 κορυφαίων τραπεζών για να προσδιοριστούν οι τομείς κινδύνου που αναφέρονται συγκεκριμένα από τις τράπεζες αυτές. Η επισκόπηση περιλάμβανε επίσης τον προσδιορισμό των συγκεκριμένων εργαλείων, μεθοδολογιών ή στοιχείων του πλαισίου διαχείρισης κινδύνων που

<sup>3</sup>Επιτροπή της Βασιλείας για την Τραπεζική Εποπτεία 2008

<sup>4</sup>Επιτροπή της Βασιλείας για την Τραπεζική Εποπτεία 2011

χρησιμοποιούνταν. Για να επιτευχθεί ευρύτερη κάλυψη, ο κατάλογος των τραπεζών περιλάμβανε έναν εκπρόσωπο από κάθε περιοχή των Η.Π.Α., τράπεζες που δραστηριοποιούνται κυρίως σε παγκόσμιο επίπεδο, ευρωπαϊκές τράπεζες, καθώς και μια ασιατική τράπεζα. Επίσης, οι τράπεζες αυτές συμπεριελάμβαναν ένα ευρύ φάσμα τραπεζικών δραστηριοτήτων - επενδυτική τραπεζική, διαπραγμάτευση τίτλων, καταναλωτική ή λιανική τραπεζική και εταιρική τραπεζική. Ενώ υπήρχαν διαφορές στον τρόπο με τον οποίο συζητούνταν και παρουσιάζονταν οι κίνδυνοι, συμπεριλαμβανομένων των επιμέρους κινδύνων, οι κορυφαίοι κίνδυνοι ήταν σε μεγάλο βαθμό οι ίδιοι και περιλάμβαναν τη διαχείριση πιστωτικού κινδύνου, τη διαχείριση κινδύνου αγοράς, τον κίνδυνο ρευστότητας, τον λειτουργικό κίνδυνο.

Παρακάτω περιλαμβάνεται ένα διάγραμμα (Σχήμα 2.1)<sup>5</sup> που απεικονίζει την ταξινόμηση των διαφόρων τύπων κινδύνου που συζητούνται στις ετήσιες εκθέσεις των τραπεζών, καθώς και τις διάφορες μεθοδολογίες ή εργαλεία (Σχήμα 2.2)<sup>6</sup> που εφαρμόζονται για τη διαχείριση αυτών των κινδύνων.



Σχήμα 3.1: Ταξινόμηση των κινδύνων

<sup>5</sup><https://epubl.ktu.edu/object/elaba:6937111/>

<sup>6</sup><https://epubl.ktu.edu/object/elaba:6937111/>

	Market Risk	Credit Risk	Liquidity Risk	Non-Financial Risk (Operational Risk)
<b>Risk Management Tools</b>				
Risk Limits	√	√	√	
Credit Risk limits		√		
Value at Risk	√			
Earnings at Risk	√			
Expected Shortfall	√			
Economic Value Stress Testing	√			
Economic Capital	√	√	√	√
Risk Sensitivities	√			
Risk Assessment (RCSA)				√
Operational Risk Losses				√
Loss Distribution Approach				√
Scenario Analysis	√	√	√	√
Tail Risk Capture	√	√	√	√
Stress Testing	√	√	√	√
Scoring Models		√		
Rating Models		√		
Exposure				
- Probability of Default		√		
- Loss Given Default				
- Exposure at Default				
Back Testing	√	√	√	
<b>Risk Management Framework Components</b>				
Risk Appetite	√	√	√	√
Risk Identification	√	√	√	√
Risk Assessment	√	√	√	√
Risk Measurement	√	√	√	√
Risk Testing	√	√	√	√
Risk Monitoring	√	√	√	√
Risk reporting	√	√	√	√
Risk Oversight	√	√	√	√
Capital Management (calculation and allocation)				
- CCAR	√	√	√	√
- ICAAP				

Σχήμα 3.2: Μέθοδοι και εργαλεία διαχείρισης κινδύνων

Ο επικεφαλής κινδύνου έχει πρόσβαση σε γνώσεις και πληροφορίες σχετικά με τους κινδύνους που είχαν περισσότερο αναδρομικό χαρακτήρα, όπως αναλύσεις περιστατικών που επικεντρώνονται στην κατανόηση του τι συνέβη και γιατί. Τώρα, ολοένα και περισσότερο, προσανατολίζονται με εργαλεία που επιτρέπουν μια ματιά προς τα εμπρός που διευκολύνει την πρόβλεψη πιθανών περιστατικών κινδύνου. Η εξόρυξη δεδομένων, η μοντελοποίηση σεναρίων και η πρόβλεψη είναι ενσωματωμένα χαρακτηριστικά των περισσότερων λύσεων διαχείρισης κινδύνων. Η γνωστική (αναγνώριση μοτίβων με την οπτικοποίηση και τον εντοπισμό εμφανών και μεταγενέστερων τάσεων σε ιστορικά δεδομένα) και αλγοριθμική (καθιέρωση αιτιωδών σχέσεων μεταξύ διαφορετικών γεγονότων και συνόλων δεδομένων) νοημοσύνη δίνει τη θέση της στην επαυξημένη (επεξεργασία φυσικής γλώσσας και μηχανική εκμάθηση) και υποστηρικτική (εικονική ευφυής βοήθεια με βάση το πλαίσιο) νοημοσύνη

που επαυξάνει και επιταχύνει τη λήψη αποφάσεων [48].

### 3.0.1 Πιστωτικός Κίνδυνος

Η αξιολόγηση του πιστωτικού κινδύνου παραμένει ένα σημαντικό και δύσκολο ερευνητικό θέμα στον τομέα της χρηματοοικονομικής. Οι αρχικές προσπάθειες αξιολογούνται από τον περασμένο αιώνα. Με αφορμή τα γεγονότα της παγκόσμιας χρηματοπιστωτικής κρίσης και την επακόλουθη αυξημένη εστίαση των ρυθμιστικών αρχών, η διαδικασία αξιολόγησης του πιστωτικού κινδύνου έχει παρουσιάσει αυξημένο ενδιαφέρον στην ακαδημαϊκή και επιχειρηματική κοινότητα. Η γενική προσέγγιση για την αξιολόγηση του πιστωτικού κινδύνου ήταν η εφαρμογή μιας τεχνικής ταξινόμησης σε δεδομένα πελατών του παρελθόντος, συμπεριλαμβανομένων των πελατών τα δάνεια των οποίων βρίσκονται σε καθεστώς καθυστέρησης, για την ανάλυση και την αξιολόγηση της σχέσης μεταξύ των χαρακτηριστικών ενός πελάτη και της πιθανής αποτυχίας του. Αυτό θα μπορούσε να χρησιμοποιηθεί για τον προσδιορισμό ταξινομητών που μπορούν να εφαρμοστούν στην κατηγοριοποίηση νέων αιτούντων ή υφιστάμενων πελατών ως καλών ή κακών [69].

Η αξιολόγηση του πιστωτικού κινδύνου κατέχει σημαντική θέση στο πλαίσιο της διαχείρισης κινδύνων. Τεχνικές όπως η λογιστική παλινδρόμηση και η ανάλυση διάκρισης χρησιμοποιούνται παραδοσιακά στην πιστωτική βαθμολόγηση για τον προσδιορισμό της πιθανότητας αθέτησης. Οι μηχανές διανυσμάτων υποστήριξης είναι επιτυχείς στην ταξινόμηση των πελατών πιστωτικών καρτών που αθετούν τις υποχρεώσεις τους. Διαπιστώθηκε επίσης ότι είναι ανταγωνιστικές στην ανακάλυψη των χαρακτηριστικών που είναι πιο σημαντικά για τον προσδιορισμό του κινδύνου αθέτησης, όταν δοκιμάστηκαν και συγκρίθηκαν με τις παραδοσιακές τεχνικές [16]. Η μοντελοποίηση του πιστωτικού κινδύνου για τον υπολογισμό της έκθεσης σε πιστωτικές ζημίες περιλαμβάνει την εκτίμηση της πιθανότητας αθέτησης (PD), της έκθεσης σε αθέτηση (EAD) και της ζημίας σε περίπτωση αθέτησης (LGD). Αυτό τονίζεται από τη συμφωνία της Βασιλείας II. Οι επικρατέστερες μέθοδοι για την ανάπτυξη μοντέλων για την PD είναι η ταξινόμηση και η ανάλυση επιβίωσης, με την τελευταία να περιλαμβάνει την εκτίμηση του κατά πόσον ο πελάτης θα αθετήσει τις υποχρεώσεις του και πότε θα μπορούσε να συμβεί η αθέτηση. Διαπιστώθηκε ότι οι αλγόριθμοι ταξινόμησης αποδίδουν σημαντικά ακριβέστερα από την τυπική λογιστική παλινδρόμηση στην πιστωτική βαθμολόγηση. Επίσης, διαπιστώθηκε ότι προηγμένες μέθοδοι αποδίδουν εξαιρετικά καλά σε σύνολα δεδομένων πιστωτικής βαθμολόγησης, όπως τα τεχνητά νευρωνικά δίκτυα, με καλύτερες

επιδόσεις από τις μηχανές ακραίας μάθησης [44].

Μέσω των απαιτήσεων της συμφωνίας της Βασιλείας, η ανάγκη κατανομής του κεφαλαίου με αποτελεσματικό και κερδοφόρο τρόπο οδήγησε τα ΑΠΙ στη δημιουργία μοντέλων πιστωτικής βαθμολόγησης για την αξιολόγηση του κινδύνου αθέτησης των πελατών τους. Και πάλι, η S.V.M. έχει αποδειχθεί ότι αποδίδει σημαντικά καλύτερα αποτελέσματα στην πιστωτική βαθμολόγηση [69]. Μια ακριβής πρόβλεψη της εκτιμώμενης πιθανότητας αθέτησης παρέχει μεγαλύτερη αξία στη διαχείριση κινδύνου σε σύγκριση με μια δυαδική ταξινόμηση των πελατών ως αξιόπιστων ή μη αξιόπιστων. Στην πιστωτική βαθμολόγηση χρησιμοποιούνται διάφορες τεχνικές, όπως η διακριτική ανάλυση, η λογιστική παλινδρόμηση, ο ταξινομητής Bayes, ο πλησιέστερος γείτονας, τα τεχνητά νευρωνικά δίκτυα και τα δέντρα ταξινόμησης. Έχει αποδειχθεί ότι τα τεχνητά νευρωνικά δίκτυα εκτελούν ταξινομήσεις με μεγαλύτερη ακρίβεια από τις άλλες πέντε μεθόδους [67]. Μέθοδοι και μοντέλα αναπτύσσονται συνεχώς για την αντιμετώπιση ενός σημαντικού ζητήματος στις τράπεζες, δηλαδή τη σωστή ταξινόμηση των πελατών και την εκτίμηση του πιστωτικού κινδύνου. Οι διάφορες προσεγγίσεις που εφαρμόζονται σε αυτές τις μεθόδους επιδιώκουν να αυξήσουν την ακρίβεια των προβλέψεων της πιστοληπτικής ικανότητας που θα μπορούσαν να οδηγήσουν σε ένα μεγαλύτερο και κερδοφόρο χαρτοφυλάκιο δανείων. Τα νευρωνικά δίκτυα έχουν αποδειχθεί ότι έχουν σημαντική αξία στη διαδικασία λήψης αποφάσεων για τον πιστωτικό κίνδυνο, ενώ η εφαρμογή τους σε προβλέψεις δυσπραγίας επιχειρήσεων αναφέρθηκε ότι είναι επωφελής στην αξιολόγηση του πιστωτικού κινδύνου [71].

Ενώ ο πιστωτικός κίνδυνος είναι ο πιο ερευνημένος και αξιολογημένος τομέας κινδύνου για την εφαρμογή της μηχανικής μάθησης, αυτό δεν είναι ένα νέο φαινόμενο. Από το 1994, ο Altman και οι συνεργάτες του διεξήγαγαν μια ανάλυση που συνέκρινε τις παραδοσιακές στατιστικές μεθόδους πρόβλεψης της δυσχέρειας και της πτώχευσης με εναλλακτικό αλγόριθμο νευρωνικών δικτύων και κατέληξαν στο συμπέρασμα ότι μια συνδυασμένη προσέγγιση των δύο βελτιώνει σημαντικά την ακρίβεια [97].

Οι Hand και Henley (1997) υποστήριξαν ότι "η πιστωτική βαθμολόγηση είναι ο όρος που χρησιμοποιείται για να περιγράψει τις επίσημες στατιστικές μεθόδους που χρησιμοποιούνται για την ταξινόμηση των αιτούντων πίστωση σε "καλές" και "κακές" κατηγορίες κινδύνου". Τα μοντέλα πιστωτικής βαθμολόγησης είναι πολυμεταβλητά στατιστικά μοντέλα που εφαρμόζονται σε οικονομικούς και χρηματοοικονομικούς δείκτες για την πρόβλεψη του κινδύνου αθέτησης πληρωμών ατόμων ή επιχειρήσεων. Στους δείκτες αυτούς αποδίδεται ένα βάρος

σχετικό με τη σημασία τους στις προβλέψεις και τροφοδοτούνται ως δεδομένα εισόδου για να προκύψει ένας δείκτης πιστοληπτικής ικανότητας. Αυτή η αριθμητική βαθμολογία χρησιμεύει ως μέτρο της πιθανότητας αθέτησης του δανειολήπτη. Η τεχνική της μηχανής διανυσμάτων υποστήριξης κατέληξε στο συμπέρασμα ότι είναι η πιο διαδεδομένη στις αξιολογήσεις πιστωτικού κινδύνου. Προτάθηκαν υβριδικά μοντέλα S.V.M. για τη βελτίωση της απόδοσης με την προσθήκη μεθόδων για τη μείωση του υποσυνόλου χαρακτηριστικών. Αυτά, ωστόσο, μόνο ταξινομούν και δεν παρέχουν εκτίμηση της πιθανότητας αθέτησης [40].

Η δραματική αύξηση της καταναλωτικής πίστης έχει αυξήσει τη σημασία των μοντέλων πιστωτικής βαθμολόγησης. Ο κύριος όγκος της έρευνας φαίνεται να επικεντρώνεται στις τεχνικές πιστωτικής βαθμολόγησης, όπως φαίνεται από τον αριθμό των εργασιών που επικεντρώνονται σε αυτόν τον τομέα [1][2][3][16][19][28][33][40][44][47][65][67][69][71]. Κατά κύριο λόγο, η έμφαση δίνεται στην ταξινόμηση και στην εφαρμογή αλγορίθμων που την καθιστούν δυνατή. Πολλές εργασίες αξιολογούν τους διάφορους αλγορίθμους σε μια προσπάθεια να εντοπιστεί ο πιο αποτελεσματικός και ακριβής αλγόριθμος πρόβλεψης. Οι εργασίες διατυπώνουν την άποψη ότι η μηχανική μάθηση παρέχει συγκρίσιμη ακρίβεια και είναι καλύτερα εξοπλισμένα για να καταγράψει μη γραμμικές σχέσεις που είναι κοινές στον πιστωτικό κίνδυνο [7][29][78][79].

Οι Zhou και Wang [80] προτείνουν την κατανομή βαρών στα δέντρα αποφάσεων για καλύτερη πρόβλεψη. Προτείνουν έναν βελτιωμένο αλγόριθμο τυχαίου δάσους για προβλέψεις. Ο αλγόριθμος, κατά τη διάρκεια της συνάθροισης, κατανέμει βάρη τα οποία υπολογίζονται με βάση τα σφάλματα εκτός bagging κατά την εκπαίδευση στα δέντρα απόφασης στο δάσος. Προσπαθούν να αντιμετωπίσουν το πρόβλημα της δυαδικής ταξινόμησης και το πείραμά τους δείχνει ότι ο προτεινόμενος αλγόριθμος υπερτερεί του αρχικού τυχαίου δάσους (R.F.) και άλλων δημοφιλών αλγορίθμων ταξινόμησης (S.V.M., K.N.N., C4.5) όσον αφορά τις ισορροπημένες και συνολικές μετρήσεις ακρίβειας.

Ορισμένες εργασίες στοχεύουν, σε αντίθεση με τις παραδοσιακές στατιστικές μεθόδους, να αναδείξουν την αποτελεσματικότητα στην εφαρμογή αλγορίθμων μηχανικής μάθησης. Οι Galindo και Tamayo [25] ερευνούν, μέσω μιας συγκριτικής ανάλυσης στατιστικών τεχνικών ταξινόμησης και τεχνικών μηχανικής μάθησης, τα πιστωτικά χαρτοφυλάκια ιδρυμάτων για να κάνουν ακριβείς προβλέψεις του ατομικού κινδύνου. Κατασκεύασαν περισσότερα από 9000 μοντέλα στο πλαίσιο της μελέτης και κατέταξαν την απόδοση των διαφόρων αλγορίθμων. Δείχνουν ότι τα μοντέλα δέντρων αποφάσεων CART παρείχαν τις καλύτερες εκτιμήσεις

για την αθέτηση, ενώ τα νευρωνικά δίκτυα ήρθαν δεύτερα. Οι Hamori κ.ά.[29] μελέτησαν και συνέκριναν την ακρίβεια πρόβλεψης και την ικανότητα ταξινόμησης των μεθόδων συσσωρευτικού δείγματος (bagging), τυχαίου δάσους (R.F.), ενισχυτικής μαθήσεως (boosting) με νευρωνικά δίκτυα στην ανάλυση δεδομένων αθέτησης πληρωμών. Διαπίστωσαν ότι η μέθοδο ενισχυτικής μαθήσεως (boosting) ήταν ανώτερο μεταξύ των μελετημένων μεθόδων μηχανικής μάθησης.

Ορισμένοι ερευνητές έχουν ακόμα αξιολογήσει την εφαρμογή υβριδικών τεχνικών και μεθόδων συνόλου για τη μελέτη της πιστωτικής βαθμολόγησης [14][29][56]. Σε ένα υβριδικό σύστημα, μια τεχνική χρησιμοποιείται για την τελική πρόβλεψη μετά τη χρήση πολλών ετερογενών τεχνικών στην ανάλυση [33]. Για την αντιμετώπιση προβλημάτων πιστωτικής βαθμολόγησης, μπορεί να εφαρμοστεί η μάθηση συνόλου, με χρήση κανονικοποιημένης λογιστικής παλινδρόμησης. Διαπιστώθηκε ότι μια μέθοδος εφαρμογής αλγορίθμων ομαδοποίησης και συσσωρευτικού δείγματος (bagging) για την εξισορρόπηση και τη διαφοροποίηση των δεδομένων, ακολουθούμενη από σύνολο λογαριθμικής παλινδρόμησης lasso-logistic για την αξιολόγηση των πιστωτικών κινδύνων, υπερτερεί έναντι πολλών δημοφιλών μοντέλων αξιολόγησης πιστοληπτικής ικανότητας [70].

Οι Khandani κ.ά.[41], για να βελτιώσουν τα ποσοστά ταξινόμησης των καθυστερήσεων και των αθετήσεων των κατόχων πιστωτικών καρτών, κατασκεύασαν ένα μη γραμμικό, μη παραμετρικό μοντέλο πρόβλεψης. Το μοντέλο κινδύνου καταναλωτικής πίστωσης ήταν σε θέση να εντοπίσει λεπτές μη γραμμικές σχέσεις σε μαζικά σύνολα δεδομένων. Αυτές οι σχέσεις ήταν συνήθως δύσκολο να εντοπιστούν κατά τη χρήση τυποποιημένων μοντέλων καταναλωτικής πίστης-μηδενικού κινδύνου, όπως logit, διακριτική ανάλυση ή πιστωτικές βαθμολογίες. Αυτό επιτρέπει τη διαχείριση του κινδύνου πιστωτικών γραμμών, την πρόβλεψη των συνολικών καθυστερήσεων καταναλωτικών πιστώσεων και την πρόβλεψη του κύκλου καταναλωτικών πιστώσεων.

Οι Yu κ.ά. [76] προτείνουν ένα νέο πολυβάθμιο βαθύ δίκτυο πεποιθήσεων βασισμένο στην ακραία μηχανική μάθηση ως πολλά υποσχόμενο εργαλείο για την αξιολόγηση του πιστωτικού κινδύνου. Το πλαίσιο των παραδειγμάτων μάθησης συνόλου πολλών σταδίων, που λειτουργούν σε τρία στάδια, αποδεικνύεται ότι υπερτερεί έναντι των τυπικών τεχνικών ταξινόμησης μεμονωμένων κατηγοριών και παρόμοιων παραδειγμάτων μάθησης συνόλου πολλών σταδίων με υψηλή ακρίβεια πρόβλεψης.

Το S.V.M. είναι ένας αλγόριθμος μηχανικής μάθησης με επίβλεψη, και ενώ χρησιμο-

ποιείται ευρέως σε προβλήματα ταξινόμησης, είναι σχετικά νέος στην πιστωτική βαθμολόγηση. Σε αυτόν τον αλγόριθμο, κάθε στοιχείο δεδομένων σχεδιάζεται ως σημείο σε χώρο n-διαστάσεων, η τιμή κάθε χαρακτηριστικού είναι η τιμή μιας συγκεκριμένης συντεταγμένης (n-ο αριθμός των χαρακτηριστικών). Η ταξινόμηση πραγματοποιείται με την εύρεση του υπερεπιπέδου που είναι το σύνορο που διαχωρίζει δύο κλάσεις [57]. Το S.V.M. έχει εφαρμοστεί ως έχει ή σε κάποια παραλλαγμένη μορφή για τον σχεδιασμό ενός μοντέλου αξιολόγησης πιστωτικού κινδύνου και βαθμολόγησης πιστώσεων [16][19][43][65][138]. Ο Harris [31] συγκρίνει μοντέλα πιστωτικής βαθμολόγησης με βάση S.V.M. χρησιμοποιώντας ευρεία (<90 ημέρες καθυστέρησης) και στενή (>90 ημέρες καθυστέρησης) παρακολούθηση, με τη δεύτερη να αποτελεί την πιο παραδοσιακή προσέγγιση. Διαπιστώθηκε ότι τα μοντέλα που κατασκευάστηκαν με τη χρήση ευρύτερου ορισμού ήταν πιο ακριβή, επιτρέποντας βελτιώσεις στην ακρίβεια πρόβλεψης.

Οι Wang κ.ά. [70] προτείνουν μια νέα "ασαφή μηχανή διανυσμάτων υποστήριξης". Ο αλγόριθμος προσπαθεί να διακρίνει τους καλούς πιστωτές από τους κακούς μέσω μεγαλύτερης γενίκευσης, διατηρώντας παράλληλα την ικανότητα της ασαφούς S.V.M. να μην αντιδρά στις ακραίες τιμές. Παρουσιάζουν μια διμερώς σταθμισμένη ασαφή S.V.M. με αποτελέσματα που δείχνουν πολλά υποσχόμενη εφαρμογή στην πιστωτική ανάλυση. Οι Huang κ.ά. [30] κατασκεύασαν ένα μοντέλο πιστωτικής βαθμολόγησης για την αξιολόγηση της πιστωτικής βαθμολογίας ενός αιτούντος από χαρακτηριστικά εισόδου με βάση ένα υβριδικό S.V.M. που κατασκευάστηκε με τη χρήση τριών στρατηγικών.

Οι Yeh και Lien [75], στην εργασία τους, αναγνώρισαν ότι η πρόβλεψη της πιθανότητας αθέτησης (PD) είναι μια πρόκληση που αντιμετωπίζουν οι επαγγελματίες και οι ερευνητές και χρήζει περισσότερης μελέτης. Ορισμένες εργασίες έχουν ως στόχο να προχωρήσουν πέρα από την απλή ταξινόμηση, ώστε να προβλέπουν την πιθανότητα αθέτησης (PD) ή τα ποσοστά ανάκτησης (RR) [14][26]. Η τεχνική της διανυσματικής μηχανής ελαχίστων τετραγώνων, όταν ενσωματώνεται σε ένα μοντέλο δύο καταστάσεων, ακολουθούμενο από ένα βήμα παλινδρόμησης, για την πρόβλεψη των ποσοστών ανάκτησης, αναφέρθηκε επίσης ότι παρουσιάζει βελτίωση σε σύγκριση με τα παραδοσιακά μοντέλα στατιστικής παλινδρόμησης [74]. Οι τεχνικές διανυσματικής παλινδρόμησης υποστήριξης θα μπορούσαν επίσης να εφαρμοστούν για την αύξηση της προβλεπτικής ικανότητας της ζημίας λόγω αθέτησης για τα εταιρικά ομόλογα σε σχέση με τα στατιστικά μοντέλα [73]. Αυτές οι εργασίες δεν σταματούν να είναι σε θέση να παρέχουν ποσοτική αξία, καθώς φαίνεται να προσεγγίζουν περισσότερο

από την άποψη της ταξινόμησης.

Οι Raei κ.ά. [26] ερευνούν ένα νέο υβριδικό μοντέλο για την εκτίμηση της πιθανότητας αθέτησης των εταιρικών πελατών σε μια εμπορική τράπεζα. Παρουσιάζουν τον υβριδικό τρόπο λειτουργίας ως έναν που μπορεί να αντιμετωπίσει την κριτική των νευρωνικών δικτύων ως "μαύρο κουτί" -το μοντέλο που λαμβάνεται δεν είναι κατανοητό από την άποψη των παραμέτρων. Η έρευνα συνδυάζει μια προσέγγιση δύο σταδίων, δηλαδή τον συνδυασμό της κατανοητότητας των μοντέλων logit με την προβλεπτική ικανότητα μη γραμμικών τεχνικών όπως τα νευρωνικά δίκτυα. Η συνολική ακρίβεια αυτού του υβριδικού μοντέλου αποδείχθηκε ότι υπερέρχει και των δύο βασικών μοντέλων. Τα χαρτοφυλάκια χαμηλής αθέτησης (LDP) είναι εκείνα που θεωρούνται πολύ χαμηλού κινδύνου. Τα LDP έχουν πρόβλημα ανισορροπίας τάξεων, καθώς, σε μια τάξη αθετών, περιέχονται σε μικρότερο αριθμό από ό,τι σε μια τάξη καλών πληρωτών. Διαπιστώθηκε ότι οι ταξινομητές Gradient boosting και οι ταξινομητές τυχαίου δάσους είχαν καλή απόδοση στην αντιμετώπιση δειγμάτων που παρουσίαζαν πρόβλημα ανισορροπίας κλάσεων [18].

Οι τράπεζες επιδιώκουν να αναπτύξουν αποδοτικά μοντέλα που μπορούν να εκτιμήσουν την πιθανότητα αθέτησης των υποχρεώσεων των αντισυμβαλλομένων. Οι Barboza κ.ά. [18] δοκιμάζουν μοντέλα μηχανικής μάθησης για την πρόβλεψη της πτώχευσης ένα έτος πριν από το γεγονός συγκρίνοντας τις επιδόσεις με τα αποτελέσματα των παραδοσιακών μεθόδων. Αναφέρουν ότι επιτυγχάνεται σημαντική ακρίβεια πρόβλεψης και επίσης υποδεικνύουν ότι οι τεχνικές ML μπορούν εύκολα να εφαρμοστούν προσφέροντας σημαντική ακρίβεια ταξινόμησης σε σύγκριση με τους παραδοσιακούς μηχανισμούς. Παρά τις ανησυχίες αναφορικά με την επεξηγηματική ικανότητα του μοντέλου, δεδομένης της πολυπλοκότητας των μοντέλων πτώχευσης, η μηχανική μάθηση θα μπορούσε να αποδειχθεί σημαντική βοήθεια. Οι Yang κ.ά. [72] διερευνούν επίσης μια νέα μέθοδο για την πρόβλεψη της πτώχευσης, προτείνοντας μια συνδυασμένη μέθοδο επιλογής χαρακτηριστικών βασισμένη στα μερικά ελάχιστα τετράγωνα (P.L.S.) με S.V.M. για τη συγχώνευση πληροφοριών. Μια τράπεζα θα μπορούσε να επωφεληθεί από την ικανότητα του μοντέλου να επιλέγει τους χρηματοοικονομικούς δείκτες που είναι πιο σημαντικοί στη διαδικασία πρόβλεψης, καθώς και από το υψηλό επίπεδο ακρίβειας πρόβλεψης.

Υπάρχουν επίσης μερικές εργασίες που ερευνούν τον τομέα των δοκιμών αντοχής στη διαχείριση του πιστωτικού κινδύνου [35]. Οι δοκιμές προσομοίωσης ακραίων καταστάσεων απαιτούν τη μοντελοποίηση της σχέσης μεταξύ των μακροοικονομικών εξελίξεων και των

τραπεζικών μεταβλητών για τον προσδιορισμό των επιπτώσεων ακραίων σεναρίων σε μια τράπεζα. Συχνότερα, χρησιμοποιούνται προσεγγίσεις από κάτω προς τα πάνω, όπου οι προβλέψεις για τα μελλοντικά κέρδη/ζημιές γίνονται σε επίπεδα χαρτοφυλακίου ως επί το πλείστον διαχωρισμένα, γεγονός που καθιστά εντατική την επεξεργασία δεδομένων και δύσκολο τον ακριβή προσδιορισμό των κινητήριων δυνάμεων των απωλειών. Οι προβλέψεις σε ένα συγκεντρωτικό χαρτοφυλάκιο με τη χρήση μιας μεθόδου από πάνω προς τα κάτω μπορούν να συμπληρώσουν αυτή τη διαδικασία. Ένας αλγόριθμος μάθησης με επίβλεψη που δεν χρειάζεται ένα προκαθορισμένο μοντέλο είναι η μέθοδος Lasso (Least Absolute Shrinkage and Selection Operator). Μια πιο εξειδικευμένη έκδοση του Lasso είναι το Adaptive Lasso, το οποίο διαθέτει ελκυστικές ιδιότητες σύγκλισης. Το Adaptive Lasso μπορεί να χρησιμοποιηθεί ελλείψει θεωρητικών μοντέλων, όπως στην περίπτωση του top-down stress testing, για την ανακάλυψη ενός λιτού top-down μοντέλου από ένα σύνολο χιλιάδων πιθανών προδιαγραφών. Αποδείχθηκε ότι δίνει κατά προσέγγιση αμερόληπτες λύσεις, αναζητώντας μεταβλητές που περιγράφουν καλύτερα τη συμπεριφορά των ποσοστών πιστωτικών ζημιών με αποτέλεσμα μια φειδωλή περιγραφή της σχέσης μεταξύ μακροοικονομίας και ποσοστών πιστωτικών ζημιών. Ένα βασικό ζήτημα είναι η ανάγκη για σημαντικές ποσότητες δεδομένων για την εκπαίδευση ενός μοντέλου [17].

Η επιλογή μοντέλου και η πρόβλεψη έχουν καταστεί πρόκληση καθώς τα σενάρια προσομοίωσης ακραίων καταστάσεων γίνονται όλο και πιο ολοκληρωμένα, περιλαμβάνοντας όλο και περισσότερες πρωτογενείς μεταβλητές. Οι τεχνικές μηχανικής μάθησης για τον εντοπισμό μοτίβων και σχέσεων μεταξύ των δεδομένων μπορούν να διευκολύνουν την επιλογή μοντέλων και την πρόβλεψη. Οι τεχνικές αυτές δεν φαίνεται να εφαρμόζονται ευρέως στις δοκιμές αντοχής. Όταν υπάρχει μεγάλος αριθμός δυνητικών συνδιαλλαγών και ο αριθμός των παρατηρήσεων είναι μικρός, οι παλινδρομήσεις lasso κρίνονται κατάλληλες για τη δημιουργία μοντέλων πρόβλεψης. Είναι πιθανό να υπερτερούν των παραδοσιακών στατιστικών μοντέλων στην πρόβλεψη των δεικτών απόδοσης που απαιτούνται στον εφαρμοσμένο έλεγχο αντοχής. Ένα πλεονέκτημα των εκτιμητών τύπου Lasso είναι ότι μπορούν να χειριστούν τις επιπλοκές που προκύπτουν από την υψηλής διάστασης φύση των δοκιμών αντοχής [20]. Οι Multivariate Adaptive Regression Splines (MARS), μια τεχνική μηχανικής μάθησης, μπορεί να θεωρηθεί ως γενίκευση της βηματικής γραμμικής παλινδρόμησης της μεθόδου δέντρων ταξινόμησης και παλινδρόμησης (CART). Ο έλεγχος αντοχής στατιστικών μοντέλων παλινδρόμησης, όπως η διανυσματική αυτοπαλινδρόμηση (VAR), είναι μια κοινή προσέγ-

γιση μοντελοποίησης, η οποία είναι γνωστό ότι δεν μπορεί να εξηγήσει το φαινόμενο των κατανομών με παχιά ουρά. Μια εμπειρική δοκιμή αυτών των μοντέλων διαπίστωσε ότι το μοντέλο MARS παρουσίασε μεγαλύτερη ακρίβεια στη δοκιμή μοντέλου και ανώτερη απόδοση εκτός δείγματος, με το MARS να παράγει πιο λογικές προβλέψεις [36]. Τα πιθανοτικά γραφήματα μπορούν να χρησιμοποιηθούν για τη μοντελοποίηση και την αξιολόγηση του κινδύνου πιστωτικής συγκέντρωσης με ένα δενδροειδώς ενισχυμένο δίκτυο Bayes να παρέχει καλύτερη κατανόηση του κινδύνου. Αυτό κρίθηκε επίσης κατάλληλο για αναλύσεις προσομοίωσης ακραίων καταστάσεων (stress testing), με τη δυνατότητα να παρέχει εκτιμήσεις του κινδύνου απωλειών ως συνέπεια αλλαγών στη χρηματοοικονομική κατάσταση ενός δανειολήπτη [52]. Το E.M.A. workbench είναι μια εργαλειοθήκη λογισμικού, που αναπτύχθηκε από μια ομάδα στο Technische Universiteit Delft. Χρησιμοποιούνται συγκεκριμένοι αλγόριθμοι μηχανικής μάθησης και προηγμένα εργαλεία οπτικοποίησης για την εκτέλεση πολλαπλών πειραμάτων και την ανάλυση των αποτελεσμάτων, παρέχοντας τη δυνατότητα διερεύνησης πιθανών αβεβαιοτήτων και εντοπισμού αιτιών με βάση τις εισροές.

Τα τελευταία χρόνια, υπήρξε σημαντική αύξηση στις προσπάθειες έρευνας που στοχεύουν στη βελτίωση των μεθόδων ανίχνευσης απάτης μέσω πιστωτικών καρτών. Αυτές οι προσπάθειες έχουν εξερευνήσει διάφορες τεχνικές και προσεγγίσεις για την αντιμετώπιση των εξελισσόμενων προκλήσεων που προκύπτουν από τις απάτες στον χρηματοοικονομικό τομέα. Πολλές μελέτες, συμπεριλαμβανομένων των "CATCHM: A novel network-based credit card fraud detection method using node representation learning" του R. Van Belle et al.[156], "A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection" του H. Fanai και H. Abbasimehr [157], και "The effect of feature extraction and data sampling on credit card fraud detection" του Z. Salekshahrezaee et al [158], έχουν επικεντρωθεί στην αξιοποίηση τεχνικών μηχανικής μάθησης και βαθιάς μάθησης για βελτιωμένη ανίχνευση απάτης. Άλλα έργα, όπως το "Semi-supervised credit card fraud detection via attribute-driven graph representation" των S. Xiang et al [146], "NUS: Noisy-Sample-Removed Undersampling Scheme for Imbalanced Classification and Application to Credit Card Fraud Detection." των H. Zhu et al. [147], και "Comparative analysis of binary and one-class classification techniques for credit card fraud data" του JL Leevy et al [148], έχουν εστιάσει σε προηγμένες μεθόδους όπως η μοντελοποίηση βάσει γράφων και τα σχήματα ανίχνευσης απάτης με ανισορροπημένα δεδομένα. Επιπλέον, μελέτες όπως "Oppositional Cat Swarm Optimization-Based Feature Selection Approach for Credit Card Fraud Detection"

του N. Prabhakaran και R. Nedunchelian [149], έχουν εξερευνήσει νέες τεχνικές επιλογής χαρακτηριστικών για την αύξηση της ακρίβειας της ανίχνευσης. Αυτές οι έρευνες συνεισφέρουν συλλογικά στη συνεχή ανάπτυξη πιο αξιόπιστων και αποτελεσματικών συστημάτων ανίχνευσης απάτης μέσω πιστωτικών καρτών, αντιμετωπίζοντας την κρίσιμη ανάγκη για τη διασφάλιση των χρηματοοικονομικών συναλλαγών και τη διατήρηση της ακεραιότητας του χρηματοοικονομικού οικοσυστήματος.

### 3.0.2 Κίνδυνος Αγοράς

Ο κίνδυνος μπορεί να μετρηθεί με την τυπική απόκλιση των απροσδόκητων αποτελεσμάτων, που ονομάζεται επίσης μεταβλητότητα. Η αξία σε κίνδυνο (VaR) υπολογίζει τη χειρότερη ζημία σε έναν ορίζοντα-στόχο που δεν θα ξεπεραστεί με δεδομένο επίπεδο εμπιστοσύνης και αποτυπώνει τη συνδυασμένη επίδραση της υποκείμενης μεταβλητότητας και της έκθεσης σε χρηματοοικονομικούς κινδύνους [37]. Η πρόβλεψη της μεταβλητότητας στις χρηματοπιστωτικές αγορές είναι σημαντική, μεταξύ άλλων, στους τομείς της διαχείρισης κινδύνου και της τιμολόγησης περιουσιακών στοιχείων. Με τη χρήση μοντέλων NN, η απόδοση της μεθόδου εκτίμησης της μεταβλητότητας μπορεί να βελτιωθεί [49].

Οι Zhang κ.ά. [78][79] προτείνουν ένα μοντέλο που βασίζεται στο μοντέλο Generalized Autoregressive Conditional Heteroskedastic (GARCH) και στον αλγόριθμο Extreme Machine Learning (E.L.M.) για την εκτίμηση της μεταβλητότητας. Το μοντέλο προβλέπει τη μεταβλητότητα των χρονοσειρών-στόχων χρησιμοποιώντας το GELM-RBF και η παρέκταση των προβλεπόμενων μεταβλητοτήτων επιτρέπει τον υπολογισμό της VaR με βελτιωμένη απόδοση όσον αφορά την ακρίβεια και την αποτελεσματικότητα. Το μοντέλο χρησιμοποιεί μια μέθοδο στοχαστικής χαρτογράφησης που δεν απαιτεί την πιθανότητα Gauss για την εκτίμηση και είναι ένα μη γραμμικό μοντέλο οδηγού δεδομένων.

Ο κίνδυνος αγοράς περιλαμβάνει επίσης τον κίνδυνο επιτοκίου και τον κίνδυνο μετοχών. Οι καμπύλες επιτοκίου, οι οποίες είναι η σχέση μεταξύ του επιτοκίου και του χρόνου μέχρι τη λήξη του χρέους για έναν δεδομένο δανειολήπτη σε ένα δεδομένο νόμισμα, χρησιμοποιούνται ευρέως στη χρηματοοικονομική μηχανική και τη διαχείριση του κινδύνου αγοράς. Μια μέθοδος ομαδοποίησης που ονομάζεται "Gaussian Mixture Model" μπορεί να χρησιμοποιηθεί για την ανάπτυξη μη γραμμικών μοντέλων της εξέλιξης των παραμέτρων και στη συνέχεια για την πρόβλεψη των καμπυλών επιτοκίων. Αυτό μπορεί να επιτρέψει την καλύτερη απεικόνιση των επιτοκίων [38]. Οι μέθοδοι ομαδοποίησης μηχανικής μάθησης που

έχουν σχεδιαστεί για την αντιμετώπιση της στοχαστικής διαφορικής εξίσωσης (SDE) μπορούν να εφαρμοστούν για την ανάπτυξη προβλέψιμων μοντέλων VAR που στοχεύουν να αποτελέσουν κορυφαίο μέτρο κινδύνου για την αλλαγή καθεστώτος της αγοράς. Αυτό μπορεί να αντιμετωπίσει εν μέρει κάποια από την πολυπλοκότητα που εισάγεται από το δύσκολο ρυθμιστικό περιβάλλον, όπως η συνοχή των σεναρίων [45].

### 3.0.3 Κίνδυνος Ρευστότητας

Ορισμένα προβλήματα κινδύνου ρευστότητας μπορούν να επιλυθούν με τη χρήση της μηχανικής μάθησης. Η μέτρηση του κινδύνου ρευστότητας, η ανάλυση των βασικών παραγόντων, συμπεριλαμβανομένης της μελέτης των διασυνδέσεων μεταξύ των παραγόντων, μπορούν να επιτευχθούν μέσω της χρήσης της μηχανικής μάθησης. Για τους σκοπούς της εκτίμησης ενός μέτρου κινδύνου μπορούν να εφαρμοστούν τεχνητά νευρωνικά δίκτυα (A.N.N.), ένας γενετικός αλγόριθμος. Τα A.N.N. μπορούν να χρησιμοποιηθούν για την προσέγγιση της γενικής τάσης κινδύνου και τον προσδιορισμό των παραγόντων με τη μεγαλύτερη επιρροή. Η πιθανότητα εμφάνισης ενός γεγονότος κινδύνου ρευστότητας μπορεί να εκτιμηθεί με την εφαρμογή Μπεϋζιανών Δικτύων (B.N.). Οι εφαρμογές A.N.N. και B.N. ήταν ικανές να διακρίνουν τους πιο κρίσιμους παράγοντες κινδύνου ρευστότητας που μετρούν τον κίνδυνο μέσω μιας λειτουργικής προσέγγισης και μιας εκτίμησης κατανομής, αντίστοιχα [63].

### 3.0.4 Λειτουργικός / Επιχειρησιακός Κίνδυνος

Η μηχανική μάθηση εφαρμόζεται επίσης σε επιχειρησιακούς τομείς που επιτρέπουν τον μετριασμό των κινδύνων, δηλαδή την ανίχνευση και/ή την πρόληψη των κινδύνων. Στον τομέα του λειτουργικού κινδύνου, εκτός από τις περιπτώσεις ασφάλειας στον κυβερνοχώρο, η μηχανική μάθηση επικεντρώνεται κυρίως σε προβλήματα που σχετίζονται με την ανίχνευση απάτης και την ανίχνευση ύποπτων συναλλαγών.

Οι Khrestina κ.ά. [42], στην εργασία τους, προτείνουν ένα πρωτότυπο για τη δημιουργία μιας αναφοράς που επιτρέπει την ανίχνευση ύποπτων συναλλαγών. Το πρωτότυπο χρησιμοποιεί έναν αλγόριθμο λογιστικής παλινδρόμησης. Αξίζει να σημειωθεί ότι συμπεριέλαβαν επίσης μια έρευνα έξι λύσεων λογισμικού που εφαρμόζονται σήμερα σε διάφορες τράπεζες για την αυτοματοποίηση των διαδικασιών ανίχνευσης και παρακολούθησης ύποπτων συναλλαγών. Αν και οι συγγραφείς κάνουν αναφορά σε αλγόριθμους, δεν είναι σαφές αν τα προϊόντα αυτά εφαρμόζουν τεχνικές μηχανικής μάθησης, και αν ναι, με ποιους αλγόριθμους. Δεν

έγινε περαιτέρω έρευνα για τα προϊόντα αυτά, καθώς αυτό δεν ήταν στο πεδίο εφαρμογής της διατριβής.

Ένας τέτοιος τομέας όπου ένα ευφύες σύστημα που βασίζεται στη μηχανική μάθηση είναι γνωστό ότι μπορεί να προσθέσει αξία είναι η άμυνα κατά των spammers, όπου οι τεχνικές των επιτιθέμενων εξελίσσονται. Οι απώλειες από τα ανεπιθύμητα μηνύματα περιλαμβάνουν ενδεχομένως απώλεια παραγωγικότητας, διακοπή των επικοινωνιών, επιθέσεις κακόβουλου λογισμικού και κλοπή δεδομένων, συμπεριλαμβανομένης της οικονομικής απώλειας. Η τεχνολογία MLX της Proofpoint χρησιμοποιεί προηγμένες τεχνικές μηχανικής μάθησης για να παρέχει ολοκληρωμένη ανίχνευση spam που προφυλάσσει από την απειλή του spam. Εκατομμύρια μηνύματα μπορούν να αναλυθούν από την τεχνολογία, η οποία επίσης βελτιώνει αυτόματα τον αλγόριθμο ανίχνευσης για τον εντοπισμό και την ανίχνευση νεότερων απειλών (Proofpoint 2010). Αν και δεν ήταν στο πεδίο εφαρμογής της παρούσας έρευνας, καθώς πρόκειται περισσότερο για επιχειρησιακό έλεγχο για τη διαχείριση του κινδύνου, έχει επισημανθεί ως παράδειγμα του τρόπου με τον οποίο η μηχανική μάθηση χρησιμοποιείται στη διαχείριση των κινδύνων κυβερνοασφάλειας.

Στο ξέπλυμα χρήματος, οι εγκληματίες δρομολογούν χρήματα μέσω διαφόρων συναλλαγών, τις οποίες συνδυάζουν με νόμιμες συναλλαγές για να αποκρύψουν την πραγματική πηγή των κεφαλαίων. Τα κεφάλαια προέρχονται συνήθως από εγκληματικές ή παράνομες δραστηριότητες και μπορούν να χρησιμοποιηθούν περαιτέρω σε άλλες παράνομες δραστηριότητες, συμπεριλαμβανομένης της χρηματοδότησης τρομοκρατικών δραστηριοτήτων. Έχει γίνει εκτεταμένη έρευνα για τον εντοπισμό οικονομικών εγκλημάτων με τη χρήση παραδοσιακών στατιστικών μεθόδων και πιο πρόσφατα με τη χρήση τεχνικών μηχανικής μάθησης. Οι αλγόριθμοι ομαδοποίησης εντοπίζουν πελάτες με παρόμοια μοτίβα συμπεριφοράς και μπορούν να βοηθήσουν στον εντοπισμό ομάδων ανθρώπων που συνεργάζονται για τη διάπραξη νομιμοποίησης εσόδων από παράνομες δραστηριότητες [62]. Μια σημαντική πρόκληση για τις τράπεζες, δεδομένου του μεγάλου όγκου συναλλαγών ανά ημέρα και του ανομοιομορφου χαρακτήρα πολλών από αυτές, είναι να είναι σε θέση να ταξινομήσουν όλες τις συναλλαγές και να εντοπίζουν εκείνες που είναι ύποπτης φύσης. Τα χρηματοπιστωτικά ιδρύματα χρησιμοποιούν συστήματα καταπολέμησης της νομιμοποίησης εσόδων από παράνομες δραστηριότητες για να φιλτράρουν και να ταξινομήσουν τις συναλλαγές με βάση τους βαθμούς υποψίας. Απαιτούνται δομημένες διαδικασίες και ευφυή συστήματα για να καταστεί δυνατός ο εντοπισμός αυτών των συναλλαγών νομιμοποίησης εσόδων από παράνομες δραστηριότητες [39].

Η νομιμοποίηση εσόδων από παράνομες δραστηριότητες είναι ένας άλλος τομέας που αποτελεί σημαντική πρόκληση για τα χρηματοπιστωτικά ιδρύματα, δεδομένου του μεγάλου όγκου και της πολυπλοκότητας των συναλλαγών σε συνδυασμό με τη δυναμική και ταχέως εξελισσόμενη φύση των οικονομικών εγκλημάτων, καθώς και της ανάγκης να γίνεται αυτό σε σύνολα δεδομένων σε πραγματικό χρόνο. Στον τομέα της ανίχνευσης οικονομικών εγκλημάτων, έχει πραγματοποιηθεί σημαντικός όγκος έρευνας στην εφαρμογή της στατιστικής μάθησης και της εξόρυξης δεδομένων για την ανάπτυξη μοντέλων ταξινόμησης ώστε να επισημανθούν ύποπτες συναλλαγές. Εφαρμόστηκε ένας αλγόριθμος C5.0 για την πρόβλεψη των επιπέδων κινδύνου με βάση τους διάφορους παράγοντες δυνητικού κινδύνου των πελατών, ώστε να δημιουργηθεί το σύνολο των κανόνων για την κατανομή σε ομάδες. Οι βασικοί παράγοντες χρησιμοποιήθηκαν για τον χαρακτηρισμό των προφίλ των συναλλαγών. Το μοντέλο φέρεται να παρείχε ποσοστό ορθής ταξινόμησης 99,6% στα δεδομένα δοκιμής. Ο αριθμός των προειδοποιημένων περιπτώσεων φέρεται να μειώθηκε από το σχεδόν 30% των συναλλαγών σε λιγότερο από 1% [68].

Η απάτη με πιστωτικές κάρτες αυξάνεται σημαντικά ετησίως, ζημιώνοντας τους καταναλωτές και τον κλάδο κατά δισεκατομμύρια δολάρια. Οι απατεώνες βρίσκουν συνεχώς νεότερες τεχνικές για να διαπράττουν οικονομικά εγκλήματα. Προκειμένου να διαχειριστούν τον αυξανόμενο κίνδυνο απάτης και να ελαχιστοποιήσουν τις απώλειες, οι τράπεζες διαθέτουν συστήματα ανίχνευσης απάτης. Τα συστήματα είναι προσανατολισμένα στην αύξηση του ποσοστού ανίχνευσης, ενώ ελαχιστοποιούν το ποσοστό ψευδώς θετικών αποτελεσμάτων. Τα μοντέλα εκτιμώνται με βάση δείγματα δόλιων και νόμιμων συναλλαγών στις μεθόδους ανίχνευσης με επίβλεψη, ενώ στις μεθόδους ανίχνευσης χωρίς επίβλεψη οι ακραίες ή ασυνήθιστες συναλλαγές αναγνωρίζονται ως πιθανές περιπτώσεις απάτης. Και οι δύο επιδιώκουν να προβλέψουν την πιθανότητα απάτης σε μια δεδομένη συναλλαγή. Ορισμένες αναφερθείσες προκλήσεις στην ανίχνευση απάτης μέσω πιστωτικών καρτών είναι η μη διαθεσιμότητα πραγματικών συνόλων δεδομένων, τα μη ισορροπημένα σύνολα δεδομένων, το μέγεθος των συνόλων δεδομένων και η δυναμική συμπεριφορά των απατεώνων. Οι αλγόριθμοι Bayes, K.N.N., S.V.M. και ο ταξινομητής συνόλου bagging που βασίζεται σε δέντρο αποφάσεων (D.T.) έχουν χρησιμοποιηθεί ποικιλοτρόπως σε συστήματα ανίχνευσης απάτης. Μια συγκριτική αξιολόγηση έδειξε ότι ο ταξινομητής bagging ensemble που βασίζεται σε αλγορίθμους δέντρων απόφασης λειτουργεί αποτελεσματικά, καθώς είναι ανεξάρτητος από τις τιμές των χαρακτηριστικών και είναι επίσης σε θέση να χειριστεί την ανισορροπία των

κλάσεων [77]. Οι ψευδείς συναγερμοί, δηλαδή οι συναλλαγές που χαρακτηρίζονται ως δόλιες και στην πραγματικότητα είναι νόμιμες, είναι σημαντικοί, προκαλώντας ανησυχίες στους πελάτες και καθυστερώντας την ανίχνευση των πραγματικών συναλλαγών απάτης. Οι μεγάλες καναδικές τράπεζες βασίζονται σε μεγάλο βαθμό στις βαθμολογίες NN, οι οποίες κυμαίνονται από 1 έως 999, με το 1 να αποτελεί τη χαμηλότερη πιθανότητα μιας απατηλής συναλλαγής, η οποία καθορίζεται από αλγορίθμους νευρωνικών δικτύων. Σύμφωνα με πληροφορίες, το 20% των συναλλαγών με βαθμολογία NN μεγαλύτερη ή ίση με 990 είναι δόλιες, με αποτέλεσμα οι αναλυτές απάτης να ξοδεύουν αναποτελεσματικά χρόνο για τη διερεύνηση νόμιμων συναλλαγών. Ένας μετα-ταξινομητής (τεχνική εκμάθησης πολλαπλών αλγορίθμων) που εφαρμόζεται σε ένα μετα-νευρωνικό δίκτυο αποδείχθηκε ότι παρέχει ποσοτικά μετρήσιμες βελτιώσεις στην εξοικονόμηση πόρων με μεγαλύτερο ποσοστό απατηλών συναλλαγών που συλλαμβάνονται [55]. Επίσης, στους τομείς του επιχειρησιακού κινδύνου, υπάρχουν ορισμένες εργασίες σχετικά με την ανίχνευση κινδύνου απάτης στις πιστωτικές κάρτες και την ηλεκτρονική τραπεζική. Αφορούν την ανίχνευση απάτης μέσω πιστωτικών καρτών σε τομείς που δεν σχετίζονται ειδικά με τη διαχείριση τραπεζικού κινδύνου ή τον τραπεζικό κλάδο. Θα πρέπει να σημειωθεί ότι οι αλγόριθμοι στους οποίους αναφέρονται ήταν οι S.V.M., K.N.N., Naïve Bayes Classifier, Bagging ensemble classifier βασισμένο σε D.T.[22][55][64].

## Κεφάλαιο 4

# Machine Learning (ML, Μηχανική Μάθηση)

Η μηχανική μάθηση βρίσκεται στο σημείο τομής της επιστήμης των υπολογιστών, της μηχανικής και της στατιστικής. Έχει επισημανθεί ως ένα εργαλείο που μπορεί να εφαρμοστεί σε διάφορα προβλήματα, ιδίως σε τομείς που απαιτούν την ερμηνεία και την ανάληψη δράσης σε δεδομένα [5]. Η μηχανική μάθηση παρέχει τη δυνατότητα ανίχνευσης ουσιαστικών μοτίβων στα δεδομένα και έχει γίνει ένα κοινό εργαλείο για σχεδόν κάθε εργασία που αντιμετωπίζει την απαίτηση εξαγωγής ουσιαστικών πληροφοριών από σύνολα δεδομένων. Όταν ένας προγραμματιστής έρχεται αντιμέτωπος με την απαίτηση εξαγωγής ουσιαστικών πληροφοριών από δεδομένα και τη συνακόλουθη πολυπλοκότητα των προς μελέτη μοτίβων, μπορεί να μην είναι σε θέση να παράσχει σαφείς και λεπτομερείς προδιαγραφές για τη διαδικασία εκτέλεσης. Η μηχανική μάθηση αντιμετωπίζει αυτή την πρόκληση ” εφοδιάζοντας τα προγράμματα” με την ικανότητα να ”μαθαίνουν και να προσαρμόζονται”. Τα προγράμματα μηχανικής μάθησης μαθαίνουν και βελτιώνονται και μπορούν να εφαρμοστούν όταν το πρόβλημα που πρέπει να αντιμετωπιστεί έχει τη διπλή πρόκληση της πολυπλοκότητας και της ανάγκης προσαρμοστικότητας [50].

Τα εργαλεία μηχανικής μάθησης που οδηγούν τις εξελίξεις στις μηχανές αναζήτησης και στα αυτοκινούμενα αυτοκίνητα μπορούν να υιοθετηθούν και να εφαρμοστούν στον χρηματοπιστωτικό τομέα. Διάφορες τεχνολογικές εξελίξεις έχουν συμβάλει ώστε ο χρηματοπιστωτικός τομέας να είναι σε θέση να διερευνήσει και να εξορύξει μια ογκώδη υποδομή δεδομένων που περιλαμβάνει ποικίλα σύνολα μη δομημένων μορφών οικονομικών δεδομένων σχετικά με τις αγορές και τους καταναλωτές. Οι οικονομολόγοι υιοθετούν όλο και περισσότερο τη

μηχανική μάθηση, σε συνδυασμό με άλλα εργαλεία και τεχνογνωσία για την αξιολόγηση σύνθετων σχέσεων, παρά τους περιορισμούς της μηχανικής μάθησης όσον αφορά τη δυνατότητα προσδιορισμού της αιτιότητας. Η υιοθέτηση της μηχανικής μάθησης έχει υποκινηθεί από τις δυνητικές ευκαιρίες για μείωση του κόστους, βελτίωση της παραγωγικότητας και βελτίωση της διαχείρισης κινδύνων. Οι νέοι κανονισμοί έχουν επίσης ωθήσει τις τράπεζες στην αυτοματοποίηση καθώς υπάρχει ανάγκη για αποτελεσματική κανονιστική συμμόρφωση <sup>1</sup>.

Οι αλγόριθμοι μηχανικής μάθησης που βασίζονται στα δεδομένα και στους υπολογισμούς, στηρίζονται λιγότερο σε υποθέσεις σχετικά με τα δεδομένα, συμπεριλαμβανομένης της κατανομής. Ενώ θεωρούνται πιο ισχυροί και καλύτεροι στην αντιμετώπιση πολύπλοκων μη γραμμικών σχέσεων, θεωρείται παράλληλα ότι είναι δύσκολο να ερμηνευθούν [25].

Τα τελευταία χρόνια παρατηρήθηκε αύξηση του όγκου των δεδομένων που συγκεντρώνονται στα χρηματοπιστωτικά ιδρύματα (XI). Η μεγάλη ώθηση προς την ψηφιοποίηση των υπηρεσιών και οι αυξημένες κανονιστικές απαιτήσεις υποβολής εκθέσεων είχαν ως αποτέλεσμα τη δημιουργία και/ή τη συλλογή μεγάλου όγκου μη δομημένων δεδομένων με μεγάλη συχνότητα. Τα δεδομένα αυτά προέρχονται από διάφορες πηγές, όπως εφαρμογές για καταναλωτές, αλληλεπιδράσεις με πελάτες, μεταδεδομένα και άλλες εξωτερικές πηγές δεδομένων. Η επιθυμία να ενισχύσουν τις αναλυτικές τους ικανότητες και να αυτοματοποιήσουν όλους τους επιχειρηματικούς τομείς, συμπεριλαμβανομένης της διαχείρισης κινδύνων, με τη διαχείριση και την εξόρυξη αυτών των αυξημένων όγκων και μιας ποικιλίας δεδομένων έχει οδηγήσει τα χρηματοπιστωτικά ιδρύματα να διερευνήσουν ισχυρές και αναλυτικές λύσεις, συνέπεια των οποίων είναι η αύξηση του ενδιαφέροντος και η δημοτικότητα της μηχανικής μάθησης και της τεχνητής νοημοσύνης στην κοινότητα των χρηματοπιστωτικών ιδρυμάτων [66]. Η μηχανική μάθηση θεωρείται ευρέως στον τομέα των χρηματοπιστωτικών υπηρεσιών ότι έχει τη δυνατότητα να προσφέρει την αναλυτική ικανότητα που επιθυμούν τα Χρηματοπιστωτικά Ιδρύματα. Η μηχανική μάθηση είναι ικανή να επηρεάσει κάθε πτυχή του επιχειρηματικού μοντέλου των XI - βελτιώνοντας την εικόνα των προτιμήσεων των πελατών, τη διαχείριση κινδύνων, την ανίχνευση απάτης, την παρακολούθηση της συμπεριφοράς, την αυτοματοποίηση της υποστήριξης πελατών και ακόμη και την αυτοματοποιημένη επαλήθευση ταυτότητας, όταν συνδυάζεται με βιομετρικά στοιχεία.

Ο Van Liebergen [66] εισάγει τον τομέα και, μέσω συζητήσεων με το Institute of International Finance καθώς και με άλλα τεχνολογικά εγχειρήματα, εξηγεί τις περιπτώσεις χρήσης

---

<sup>1</sup>Financial Stability Board 2017

εντός των χρηματοπιστωτικών ιδρυμάτων. Συζητά εφαρμογές στον τομέα της μοντελοποίησης του πιστωτικού κινδύνου, της ανίχνευσης της απάτης μέσω πιστωτικών καρτών και της νομιμοποίησης εσόδων από παράνομες δραστηριότητες και της επιτήρησης των παραβιάσεων συμπεριφοράς στα χρηματοπιστωτικά ιδρύματα. Επισημαίνει επίσης ότι η μηχανική μάθηση επιδιώκει να κάνει προβλέψεις "εκτός δείγματος", ενώ μαθαίνει τις "διαπιστωμένες εντός δείγματος" (παρελθοντικές) συσχετίσεις, ενώ υπολείπεται της παροχής εξήγησης για την αναλυόμενη σχέση. Αυτό θα μπορούσε να δημιουργήσει πολυπλοκότητες γύρω από την ανάπτυξη και την αξιολόγηση των μοντέλων. Η μηχανική μάθηση παίζει επίσης ρόλο στην Επιτροπή Κεφαλαιαγοράς (SEC) στη διαδικασία αξιολόγησης κινδύνου για τον εντοπισμό παραπτώματων. Ενώ αυτό έχει εφαρμογή από εποπτική άποψη και για την εποπτεία των συστημικών κινδύνων, μπορεί επίσης να χρησιμεύσει ως οδηγός για μια τράπεζα σχετικά με το πώς παρόμοιες τεχνικές μηχανικής μάθησης μπορούν να εφαρμοστούν σε αξιολογήσεις κινδύνου για τον εντοπισμό παραπτώματων (εσωτερικών ή εξωτερικών), συμπεριλαμβανομένων των αξιολογήσεων κινδύνου για εταιρικούς εκδότες ή αντισυμβαλλόμενους [15]. Στην υπολογιστική χρηματοοικονομική, η μηχανική μάθηση έχει μεγάλες δυνατότητες και θα μπορούσε να χρησιμοποιηθεί ποικιλοτρόπως, από την ολοκληρωμένη διερευνητική ανάλυση δεδομένων έως την παρουσίαση/οπτικοποίηση των αποτελεσμάτων της μοντελοποίησης [38]. Ορισμένα από τα μειονεκτήματα της μηχανικής μάθησης, όπως υποστηρίζεται, είναι ότι είναι πιο "μαύρο κουτί" στη φύση, με τα αποτελέσματα κατά καιρούς να είναι δύσκολο να ερμηνευθούν. Υποστηρίζεται επίσης ότι είναι ευαίσθητη στις ακραίες τιμές, με αποτέλεσμα την υπερβολική προσαρμογή των δεδομένων και τις αντιφατικές προβλέψεις. Υποστηρίζεται επίσης ότι έχουν τα πλεονεκτήματα ότι μπορούν να προσαρμόζονται καλύτερα σε μη γραμμικές σχέσεις μεταξύ των επεξηγηματικών μεταβλητών και των εξηγούμενων μεταβλητών, καθώς και ότι η δυνατότητά τους να εφαρμόζουν ένα ευρύτερο σύνολο μεταβλητών τείνει να βελτιώνει την ακρίβεια [7].

## 4.1 Αλγόριθμοι Μηχανικής Μάθησης

Οι αλγόριθμοι ML που χρησιμοποιήθηκαν στην παρούσα εργασία ήταν οι αλγόριθμοι K.N.N., πολυωνυμική παλινδρόμηση, S.V.M. και A.N.N. Κάθε αλγόριθμος είναι εν συντομία επιθυμητός στις ακόλουθες υποενότητες .

### 4.1.1 Ο αλγόριθμος K-Nearest Neighbors (K.N.N.)

Ο αλγόριθμος K-Nearest Neighbors (K.N.N.) [170] είναι ένας κλασικός αλγόριθμος μηχανικής μάθησης με επίβλεψη. Ο KNN χρησιμοποιείται για εργασίες ταξινόμησης καθώς και παλινδρόμησης. Στην παρούσα εργασία, ο αλγόριθμος KNN χρησιμοποιήθηκε ως παλινδρομητής. Ο KNN παραμετροποιείται από την παράμετρο K η οποία αντιπροσωπεύει τον αριθμό των πλησιέστερων γειτόνων για την παλινδρόμηση. Στην παρούσα εργασία χρησιμοποιήσαμε  $K = 1$ . Δεδομένου του δείγματος δοκιμής και του συνόλου εκπαίδευσης, αναζητήθηκε ο closest neighbor του δείγματος στο σύνολο εκπαίδευσης για πρόβλεψη και πρόγνωση.

Ο KNN είναι ένας κλασικός αλγόριθμος μηχανικής μάθησης ο οποίος χρησιμοποιείται για εργασίες ταξινόμησης καθώς και παλινδρόμησης.

### 4.1.2 Πολυωνυμική παλινδρόμηση

Η πολυωνυμική παλινδρόμηση [171] είναι μια τεχνική μηχανικής μάθησης με επίβλεψη για την παλινδρόμηση. Στην πολυωνυμική παλινδρόμηση, η τιμή της εξαρτημένης μεταβλητής (δηλαδή η πρόβλεψη ή η πρόβλεψη) λαμβάνεται με την προσαρμογή του πολυωνύμου n-οστού βαθμού στις ανεξάρτητες μεταβλητές. Οι παράμετροι του πολυωνύμου n-οστού βαθμού εκτιμώνται με τη χρήση του συνόλου εκπαίδευσης. Στη συνέχεια, οι εκτιμώμενες παράμετροι χρησιμοποιούνται για την πρόβλεψη της εξόδου ή της πρόβλεψης με βάση τις εισόδους ή τις ανεξάρτητες μεταβλητές. Στην εργασία χρησιμοποιήθηκαν πολυώνυμα 1ου, 2ου, 3ου, 4ου και 5ου βαθμού. Από την άλλη πλευρά, οι ερευνητές χρησιμοποίησαν επίσης γραμμική παλινδρόμηση για την πρόβλεψη μακροοικονομικών δραστηριοτήτων. Οι γραμμικές παλινδρομήσεις (LR) διακρίνονται σε απλές και πολλαπλές παλινδρομήσεις, ανάλογα με τον εμπλεκόμενο αριθμό ανεξάρτητων μεταβλητών. Επιπλέον, η πολλαπλή παλινδρόμηση δηλώνει μια πρόβλεψη βάσει μοντέλου με την ανάλυση της συσχέτισης μεταξύ τουλάχιστον δύο ανεξάρτητων μεταβλητών και εξαρτημένων μεταβλητών [133].

### 4.1.3 Support Vector Machine (S.V.M.)

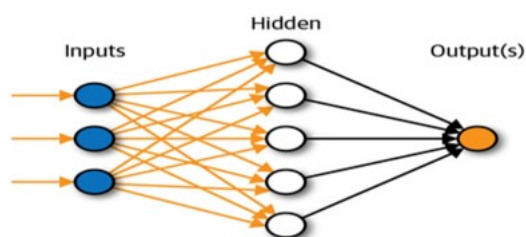
Η S.V.M. είναι μια τεχνική μάθησης με επίβλεψη [172]. Χρησιμοποιείται τόσο για παλινδρόμηση όσο και για ταξινόμηση. Η S.V.M. είναι μια παραμετρική μέθοδος που προσπαθεί να βρει υπερεπίπεδα για να περικλείσει τα μέγιστα σημεία μέσα στα υπερεπίπεδα για την παλινδρόμηση. Η S.V.M. χρησιμοποιείται ως παλινδρομητής στην παρούσα εργασία και

οι συναρτήσεις Kernel που χρησιμοποιούνται είναι γραμμικές και συναρτήσεις (RBF). Τα S.V.M. εισήχθησαν από τον Vapnik και ένα γραμμικό μοντέλο χρησιμοποιείται για την υλοποίηση μη γραμμικών ορίων κλάσεων με μη γραμμική απεικόνιση των διανυσμάτων εισόδου σε έναν χώρο χαρακτηριστικών υψηλής διάστασης και με τη χρησιμοποίηση των δεδομένων έργου πολλών κατηγοριών σε ένα υπερεπίπεδο για παλινδρόμηση ή ταξινόμηση. [140,141]. Τα S.V.M. χρησιμοποιούνται κυρίως για την ταξινόμηση δεδομένων- περιλαμβάνουν την ταξινόμηση ενός επιπέδου που χωρίζει ένα σύνολο δεδομένων σε δύο υποσύνολα, ενώ το S.V.R. συνεπάγεται την επεξεργασία σταθερών δεδομένων για τον εντοπισμό μιας συνάρτησης χαρτογράφησης που μπορεί να χρησιμοποιηθεί για την ακριβή πρόβλεψη δεδομένων. Η αρχή του SVR είναι να μετασχηματίζει τα ακατέργαστα δεδομένα σε έναν ιδιοχώρο υψηλής διάστασης για να εκτελέσει LR σε αυτόν τον ιδιοχώρο μετά την εφαρμογή μιας μη γραμμικής συνάρτησης χαρτογράφησης [134]. Ως εκ τούτου, όντας ένας αλγόριθμος μάθησης με επίβλεψη, ο S.V.M. χρειάζεται επισημασμένα δεδομένα για να εκπαιδευτεί- επομένως, τα δεδομένα που χρησιμοποιούνται ταξινομούνται και διαχωρίζονται σε διαφορετικές ομάδες. Μετά από πολλαπλές διαδικασίες εκπαίδευσης, ο αλγόριθμος προσπαθεί να διαχωρίσει τα δεδομένα. Ο S.V.M. χρησιμοποιείται ευρέως για την αναγνώριση πλαστών πιστωτικών καρτών και μπορεί να μάθει να αναγνωρίζει τις ταξινομήσεις γραφής ή εικόνας.

#### 4.1.4 Artificial Neural Network (A.N.N.)

Τα τελευταία χρόνια, τα A.N.N.s έχουν προσελκύσει την προσοχή πολλών ερευνητών, καθώς διαθέτουν την ικανότητα να προσαρμόζουν μη γραμμικά δεδομένα για πρόβλεψη και πρόγνωση. Το A.N.N. είναι ένα μαθηματικό μοντέλο το οποίο εισήχθη για να μιμηθεί τις δομές των νευρωνικών δικτύων και τις λειτουργίες στους εγκεφάλους των ανθρώπων- περιλαμβάνει πολλαπλούς αλληλοσυνδεόμενους νευρώνες/κόμβους. Οι πληροφορίες παράγονται μέσω εξωτερικής διέγερσης ή στους κόμβους και υποβάλλονται σε επεξεργασία μέσω ενεργοποίησης- τα σήματα της διαδικασίας μεταφέρονται σε άλλους κόμβους ή εξόδους [119]. Το A.N.N. περιέχει τα ακόλουθα τρία βασικά είδη στρωμάτων: στρώματα εισόδου, κρυφά στρώματα και στρώματα εξόδου. Έχουν εξοικειωθεί με διάφορα είδη του A.N.N., και αρκετά έχουν δομή πολυεπίπεδου perceptron [126]. Οι τιμές που εξάγονται από τους κρυμμένους νευρώνες εισάγονται στη συνέχεια στο επόμενο στρώμα με την έξοδο των νευρώνων. Η ακρίβεια του μοντέλου A.N.N. ενδεχομένως βελτιώνεται με τη βελτιστοποίηση των κρίσιμων παραμέτρων ενός A.N.N., συμπεριλαμβανομένου του αριθμού των κρυφών

στρωμάτων, του αριθμού των κόμβων στο κρυφό στρώμα, της χρησιμοποιούμενης συνάρτησης ενεργοποίησης και των βαρών [132]. Επιπλέον, η συνάρτηση ενεργοποίησης είναι κατά κύριο λόγο κρίσιμη, καθώς είναι μια μη γραμμική συνάρτηση και ενδεχομένως απεικονίζει την τιμή από μια γραμμική συνάρτηση σε έναν εναλλακτικό χώρο για να βελτιώσει τη συνολική απόδοση του μοντέλου και την ικανότητα δεδομένων προκειμένου να αποτυπώσει τις πραγματικές/πραγματικές καταστάσεις [131]. Τα A.N.N. χρησιμοποιούνται επίσης σε διάφορους τομείς για σκοπούς πρόβλεψης (Gurjar, Naik, Mujumdar, Vaidya, & Technology, 2018- Qeethara & Review, 2011). Η απεικόνιση του μοντέλου A.N.N. εξηγείται παρακάτω:



Σχήμα 4.1: Τυπική αρχιτεκτονική A.N.N.

# Κεφάλαιο 5

## Deep Learning (DL, Βαθιά Μάθηση)

### 5.1 Autoencoders

Οι αυτόματοι κωδικοποιητές (Autoencoders) είναι Νευρωνικά Δίκτυα που χρησιμοποιούνται παραδοσιακά για τη μείωση της διαστατικότητας με σκοπό τη δημιουργία αναπαραστάσεων των δεδομένων που μπορεί να βελτιώσουν την απόδοση ενώ καταναλώνουν λιγότερη μνήμη και χρόνο εκτέλεσης. Οι αυτόματοι κωδικοποιητές εκπαιδεύονται να είναι σε θέση να αναπαράγουν τα δεδομένα εισόδου τους ως εξόδους. Υιοθετούν ένα κρυφό επίπεδο  $h(x')$  που εκπαιδεύεται να απεικονίζει την παρεχόμενη είσοδο. Ένας αυτόματος κωδικοποιητής μπορεί να θεωρηθεί ως ένα μοντέλο που περιλαμβάνει δύο μέρη: (i) τη συνάρτηση κωδικοποιητή  $f(x)$  ( $x$  είναι η είσοδος στον αυτόματο κωδικοποιητή) και (ii) ένα σχήμα αποκωδικοποίησης που παράγει μια ανακατασκευή της αρχικής εισόδου  $g(h)$ . Ισχύουν οι παρακάτω εξισώσεις:

$$f(x) \rightarrow h(x') \quad (5.1)$$

$$h(x') \rightarrow g(h) = x \quad (5.2)$$

$$f(x), h(x') = \operatorname{argmin}_{f(x), h(x')} |x - (f(x) \circ h(x'))x|^2 \quad (5.3)$$

Η συνάρτηση κωδικοποίησης, που συμβολίζεται ως  $f(x)$ , απεικονίζει τα αρχικά δεδομένα  $x$  σε έναν κρυφό χώρο  $h(x')$  που βρίσκεται στο κρυφό επίπεδο πριν ανακατασκευαστούν από τον αποκωδικοποιητή. Η συνάρτηση αποκωδικοποιητή, που συμβολίζεται ως  $g(h)$ , απεικονίζει τον κρυφό χώρο  $h(x')$  στο κρυφό επίπεδο στην έξοδο, η οποία είναι ίδια

με την είσοδο. Το δίκτυο ενός κωδικοποιητή μπορεί να αναπαρασταθεί από μια τυπική συνάρτηση N.N. που μεταφέρεται μέσω μιας συνάρτησης ενεργοποίησης, όπου  $l$  είναι η κρυφή διάσταση, δηλαδή ( $W$  και  $b$  είναι οι συντελεστές και οι βιάζσεις των επιπέδων),

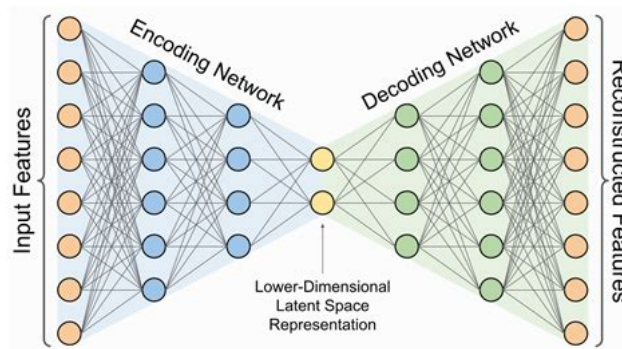
$$l = \sigma(Wx + b) \quad (5.4)$$

Επίσης, ο αποκωδικοποιητής μπορεί να απεικονιστεί με τον ίδιο τρόπο, αλλά με διαφορετικά weights, biases και, ενδεχομένως, συναρτήσεις ενεργοποίησης. Η φάση αποκωδικοποίησης μπορεί να αναπαρασταθεί από την ακόλουθη εξίσωση ( $W'$  και  $b'$  είναι τα βάρη και οι βιάζσεις του κρυφού επιπέδου):

$$x' = \sigma'(W'l + b') \quad (5.5)$$

Στον προτεινόμενο αυτόματο κωδικοποιητή, υιοθετούμε μια συνάρτηση απώλειας που γράφεται σε όρους των προαναφερθεισών συναρτήσεων. Η συνάρτηση απώλειας χρησιμοποιείται για την εκπαίδευση του N.N. μέσω του αλγορίθμου αντίστροφης διάδοσης. Μέσω της αντίστροφης διάδοσης, τα μοντέλα μας προσπαθούν συνεχώς να περιορίσουν το σφάλμα μεταξύ των υπολογισμένων τιμών και των επιθυμητών εξόδων. Αυτό αναγκάζει το κρυφό επίπεδο να χρησιμοποιεί τη μείωση της διαστατικότητας και να εξαλείφει θόρυβο κατά την ανακατασκευή των εισόδων, ειδικά όταν ο αριθμός των νευρώνων στο κρυφό επίπεδο είναι χαμηλός. Οι ακόλουθες εξισώσεις ισχύουν, ενώ στο σχήμα 5.1. αποτυπώνεται μια τυπική αρχιτεκτονική ενός Autoencoder:

$$L(x, x') = |x - x'|^2 = |x - \sigma'(W'(\sigma(Wx + b)) + b')|^2 \quad (5.6)$$



Σχήμα 5.1: Τυπική αρχιτεκτονική Autoencoder

## 5.2 Variational Autoencoders

Αντίθετα από τους τυπικούς αυτόματους κωδικοποιητές, οι πολυπολικοί αυτόματοι κωδικοποιητές (VAEs) είναι generative μοντέλα που παρουσιάζουν διαφορετικές μαθηματικές διατυπώσεις σε σύγκριση με τους αυτόματους κωδικοποιητές. Οι VAEs επικεντρώνονται σε πιθανοτικά γραφικά μοντέλα με τις εκ των υστέρων πιθανότητες να προσεγγίζονται από ένα N.N., διαμορφώνοντας έτσι την αρχιτεκτονική ενός αυτόματου κωδικοποιητή. Οι VAEs προσπαθούν να προσομοιώσουν τον τρόπο που δημιουργούνται τα δεδομένα για να αποκαλύψουν τις υποκείμενες αιτιώδεις σχέσεις. Αυτή η προσέγγιση διαφέρει από τα διακριτικά μοντέλα που στοχεύουν στο να μάθουν έναν προβλέποντα δείκτη δεδομένων. Οι VAEs βασίζονται σε ισχυρές υποθέσεις για την κατανομή των κρυφών χαρακτηριστικών χρησιμοποιώντας μια ποικιλιακή προσέγγιση. Αυτή η προσέγγιση οδηγεί σε επιπλέον συνιστώσες απώλειας και έναν συγκεκριμένο εκτιμητή για τους σκοπούς εκπαίδευσης, δηλαδή τον εκτιμητή SGVB. Η υπόθεση είναι ότι τα δεδομένα δημιουργούνται από ένα κατευθυνόμενο γραφικό μοντέλο, δηλαδή  $p_\theta(\mathbf{x}|\mathbf{h})$  και ότι ο κωδικοποιητής μαθαίνει την ακόλουθη προσέγγιση  $q_\phi(\mathbf{h}|\mathbf{x})$  στην κατανομή της μετενεργειακής κατανομής  $p_\theta(\mathbf{h}|\mathbf{x})$ , όπου τα  $\phi$  και  $\theta$  δηλώνουν τις παραμέτρους του κωδικοποιητή και του αποκωδικοποιητή, αντίστοιχα. Η πιθανοτική κατανομή του κρυφού διανύσματος ενός VAE συνήθως ταιριάζει πολύ περισσότερο με εκείνη των δεδομένων εκπαίδευσης από ό,τι ένας τυπικός αυτόματος κωδικοποιητής. Ο στόχος του VAE έχει την ακόλουθη μορφή:

$$\mathcal{L}(\phi, \theta, \mathbf{x}) = D_{\text{KL}}(q_\phi(\mathbf{h}|\mathbf{x})|p_\theta(\mathbf{h})) - \mathbb{E}q_\phi(\mathbf{h}|\mathbf{x})(\log p_\theta(\mathbf{x}|\mathbf{h})) \quad (5.7)$$

Στην παραπάνω εξίσωση, το  $D_{\text{KL}}$  αναπαριστά την απόκλιση Kullback-Leibler (KL divergence) [173]. Η προτεινόμενη κατανομή για τα κρυφά χαρακτηριστικά συνήθως ταιριάζει πολύ καλύτερα με αυτήν των δεδομένων εκπαίδευσης από έναν τυπικό αυτόματο κωδικοποιητή. Ο στόχος του VAE μπορεί να εκφραστεί ως:

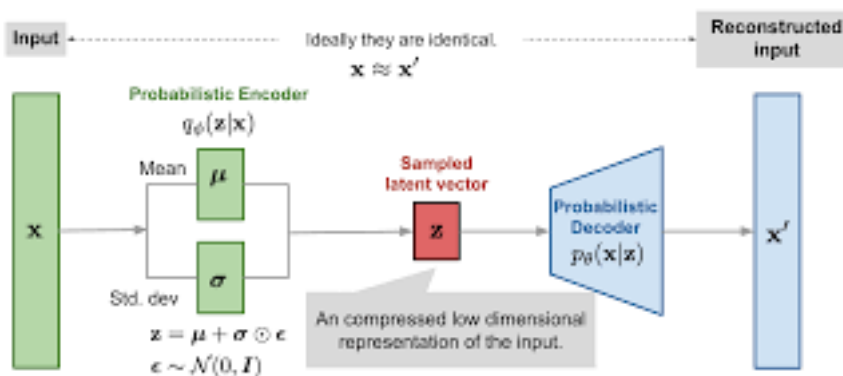
$$p_\theta(\mathbf{h}) = \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad (5.8)$$

Συνήθως, οι μορφές των προσομοιωτικών και των πιθανοτήτων επιλέγονται έτσι ώστε να είναι παραγοντοποιημένες γκαουσιανές κατανομές:

$$q_\phi(\mathbf{h}|\mathbf{x}) = \mathcal{N}(\boldsymbol{\rho}(\mathbf{x}), \boldsymbol{\omega}^2(\mathbf{x})\mathbf{I}) \quad (5.9)$$

$$p_{\theta}(\mathbf{x}|\mathbf{h}) = \mathcal{N}(\boldsymbol{\mu}(\mathbf{h}), \boldsymbol{\sigma}^2(\mathbf{h})\mathbf{I}) \quad (5.10)$$

όπου  $\boldsymbol{\rho}(\mathbf{x})$  και  $\boldsymbol{\omega}^2(\mathbf{x})$  είναι οι έξοδοι του κωδικοποιητή, ενώ  $\boldsymbol{\mu}(\mathbf{h})$  και  $\boldsymbol{\sigma}^2(\mathbf{h})$  είναι οι έξοδοι του αποκωδικοποιητή. Αυτές οι διατυπώσεις δικαιολογούνται από τον σκοπό της απλούστευσης των τελικών αποτελεσμάτων κατά την αξιολόγηση τόσο του KL divergence όσο και του όρου πιθανοφάνειας στο ποικιλιακό αντικείμενο που καθορίζεται παραπάνω. Στο παρακάτω σχήμα (5.2) παρουσιάζεται η τυπική αρχιτεκτονική ενός VAE.



Σχήμα 5.2: Τυπική αρχιτεκτονική Variational Autoencoder

### 5.3 Convolutional Neural Networks, (CNNs)

Τα Συνελεκτικά Νευρωνικά Δίκτυα (CNNs) προτάθηκαν αρχικά στην πρωτοποριακή εργασία των Yann Lecun, Le'on Bottou, Yoshua Bengio, and Parick Haffner [169] ως η τεχνολογία που χρησιμοποιείται για τη διαχείριση των δεδομένων ως ενός "σχεδίου" στον διδιάστατο χώρο. Τα CNNs έχουν εφαρμοστεί σε διάφορους τομείς, όπως η αναγνώριση εικόνων, η ανίχνευση αντικειμένων, η ταξινόμηση, η αναγνώριση βίντεο, οι συστάσεις, η ανάλυση ιατρικών εικόνων, η Επεξεργασία Φυσικής Γλώσσας (NLP) και η ανάλυση χρονοσειρών στην οικονομική. Ένα CNN χρησιμοποιεί μια ειδική λειτουργία που ονομάζεται συνέλιξη, η οποία είναι μια εξειδικευμένη μορφή γραμμικής επεξεργασίας αντί για γενική πολλαπλασιαστική πίνακα. Σε ένα CNN, τα κρυφά επίπεδα περιλαμβάνουν ένα σύνολο συνελκτικών επεξεργασιών που συνελίσσονται με πολλαπλασιασμό, άλλες πράξεις προϊόντος και υπολογισμούς αντίστασης. Αυτό είναι σημαντικό για τους δείκτες εντός του πίνακα δε-

δομένων, επηρεάζοντας έτσι τον τρόπο που τίθενται τα βάρη σε ένα επιλεγμένο σημείο <sup>1</sup>. Η συνάρτηση ReLU είναι μια ευρέως υιοθετημένη συνάρτηση ενεργοποίησης, ακολουθούμενη από επιπλέον συνελίξεις όπως επίπεδα συγκέντρωσης, πλήρως συνδεδεμένα επίπεδα και επίπεδα κανονικοποίησης. Γενικά, η συνέλιξη είναι μια λειτουργία που περιγράφεται από μια λειασμένη εκτίμηση της εισόδου συνάρτησης  $x(t)$ ,

$$s(t) = \int x(a)w(t-a)da \quad (5.11)$$

όπου  $w(a)$  είναι η συνάρτηση πυρήνα στη μορφή μιας έγκυρης πυκνότητας πιθανοφάνειας και  $s(t)$  είναι το αποτέλεσμα. Η λειτουργία της συνέλιξης απεικονίζεται συνήθως με έναν αστερίσκο:

$$s(t) = (x*w)(t) \quad (5.12)$$

Εναλλακτικά, μπορούμε να ορίσουμε τη διακριτική συνέλιξη όπως απεικονίζεται στην ακόλουθη εξίσωση:

$$s(t) = (x * w)(t) = \sum x(a)w(t-a) \quad (5.13)$$

Συνήθως, χρησιμοποιούμε συνελίξεις για πολλούς άξονες σε ένα συγκεκριμένο χρόνο. Έτσι, οι παραπάνω συναρτήσεις μπορούν να εκφραστούν ως:

$$S(i, j) = (I * K)(i, j) = \sum \sum I(m, n)K(i-m, j-n) \quad (5.14)$$

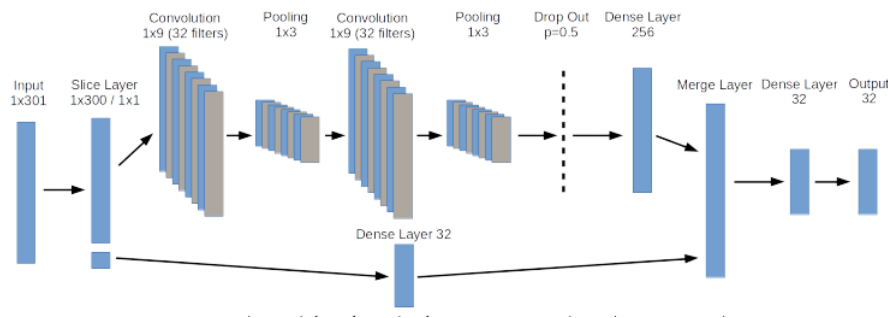
όπου  $K$  και  $I$  είναι οι χρησιμοποιημένοι πυρήνες και η παρεχόμενη είσοδος αντίστοιχα. Εξίσου, μπορούμε να εκμεταλλευτούμε την ανταλλαγή της συνέλιξης:

$$S(i, j) = (K * I)(i, j) = \sum \sum I(i-m, j-n)K(m, n) \quad (5.15)$$

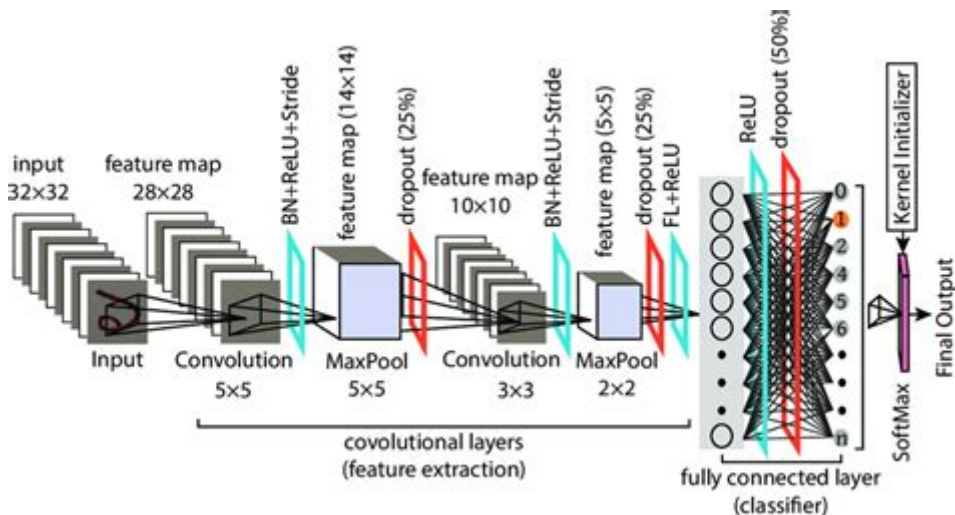
Το κίνητρο πίσω από τη χρήση αυτής της τεχνικής είναι ότι η συνέλιξη εκμεταλλεύεται τρεις σημαντικές ιδέες που μπορούν να βοηθήσουν στη βελτίωση ενός συστήματος μηχανικής μάθησης, δηλαδή (i) αραιές αλληλεπιδράσεις, (ii) κοινή χρήση παραμέτρων και (iii)

<sup>1</sup>Ian Goodfellow, Yoshua Bengio και Aaron Courville. Deep Learning. MIT Press, 2016. <http://www.deeplearningbook.org>.

ισοδύναμες αναπαραστάσεις <sup>2</sup>. Στα παρακάτω σχήματα (5.3 & 5.4) παρουσιάζονται τυπικές αρχιτεκτονικές ενός C.N.N.



Σχήμα 5.3: Τυπική αρχιτεκτονική Convolutional Neural Network



Σχήμα 5.4: Τυπική αρχιτεκτονική Convolutional Neural Network

## 5.4 Long Short Term Memory

Μέχρι σήμερα, τα πιο αποτελεσματικά μοντέλα ακολουθίας που χρησιμοποιούνται σε πρακτικές εφαρμογές ονομάζονται αυτόματοι RNNs. Αυτά περιλαμβάνουν το δίκτυο LSTM και δίκτυα που βασίζονται στη μονάδα αυτόματης αναστροφής πύλης (gated recurrent unit)

Η ιδέα της εισαγωγής αυτοδιαδρομών για τη δημιουργία διαδρομών όπου η κλίση μπορεί να ρέει για μεγάλες διάρκειες είναι μια καίρια συνεισφορά του αρχικού μοντέλου LSTM [174].

<sup>2</sup>Ian Goodfellow, Yoshua Bengio και Aaron Courville. Deep Learning. MIT Press, 2016. <http://www.deeplearningbook.org>.

Ένα πολύ σημαντικό στοιχείο ήταν η προσθήκη βάρους σε αυτήν την αυτοδιαδρομή υπό συνθήκη, αντί για σταθερό βάρος [175]. Καθιστώντας το βάρος αυτής της αυτοδιαδρομής ρυθμιζόμενο (ελεγχόμενο από μια άλλη κρυφή μονάδα), η χρονική κλίμακα της ολοκλήρωσης μπορεί να αλλάξει δυναμικά. Σε αυτήν την περίπτωση, εννοούμε ότι, ακόμη και για ένα LSTM με σταθερές παραμέτρους, η χρονική κλίμακα της ολοκλήρωσης μπορεί να αλλάξει με βάση την ακολουθία εισόδου, διότι οι χρονικές σταθερές είναι εξόδοι του ίδιου του μοντέλου.

Αντί για μια μονάδα που απλώς εφαρμόζει μη γραμμικότητα στη μετασχηματισμό των εισόδων και των αναστροφών μονάδων, τα αναδρομικά δίκτυα LSTM έχουν 'κύτταρα LSTM' που έχουν μια εσωτερική ανατροφοδότηση (αυτοδιαδρομή), εκτός από την εξωτερική ανατροφοδότηση του RNN. Κάθε κύτταρο έχει τις ίδιες εισόδους και εξόδους με ένα κανονικό αναδρομικό δίκτυο, αλλά έχει επίσης περισσότερες παραμέτρους και ένα σύστημα πύλης που ελέγχει τη ροή πληροφοριών. Το πλέον σημαντικό στοιχείο είναι η μονάδα κατάστασης  $s_i^{(t)}$  που έχει ένα γραμμικό αυτοδιάλυτο. Εδώ, το αυτοδιάλυτο βάρος (ή ο σχετικός χρονικός σταθμός) ελέγχεται από μια μονάδα πύλης λήσης  $f_i^{(t)}$  (για το χρονικό βήμα  $t$  και το κύτταρο  $i$ ), που ορίζει αυτό το βάρος σε μια τιμή στο διάστημα της μονάδας με την υιοθέτηση μιας μονάδας:

$$f_i^{(t)} = \sigma \left( b_i^f + \sum_j U_{i,j}^f x_j^{(t)} + \sum_j W_{i,j}^f h_j^{(t-1)} \right) \quad (5.16)$$

όπου  $x^{(t)}$  είναι το τρέχον διάνυσμα εισόδου και  $h^{(t)}$  είναι το τρέχον διάνυσμα κρυφού επιπέδου που περιλαμβάνει τις εξόδους όλων των κυττάρων LSTM, και  $\mathbf{b}^f$ ,  $\mathbf{U}^f$ ,  $\mathbf{W}^f$  είναι, αντίστοιχα, βιάσεις, βάρη εισόδου και βάρη επανάληψης για τις πύλες λήσης. Η εσωτερική κατάσταση του κυττάρου LSTM ενημερώνεται ως εξής, αλλά με ένα συνθηκόμενο αυτό-διαλυτο βάρος  $f_i^t$ :

$$s_i^{(t)} = f_i^{(t)} s_i^{(t-1)} + g_i^t \sigma \left( b_i + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} h_j^{(t-1)} \right) \quad (5.17)$$

όπου  $\mathbf{b}$ ,  $\mathbf{U}$  και  $\mathbf{W}$  υποδηλώνουν τις βιάσεις, τα βάρη εισόδου και τα βάρη επανάληψης στο κύτταρο LSTM, αντίστοιχα. Η μονάδα εισόδου πύλης  $g_i^{(t)}$  υπολογίζεται με παρόμοιο τρόπο με τη μονάδα λήσης (με μια μονάδα σιγμοειδούς για να λάβει μια τιμή λήσης στο διάστημα της μονάδας), αλλά με τα δικά της παραμέτρους:

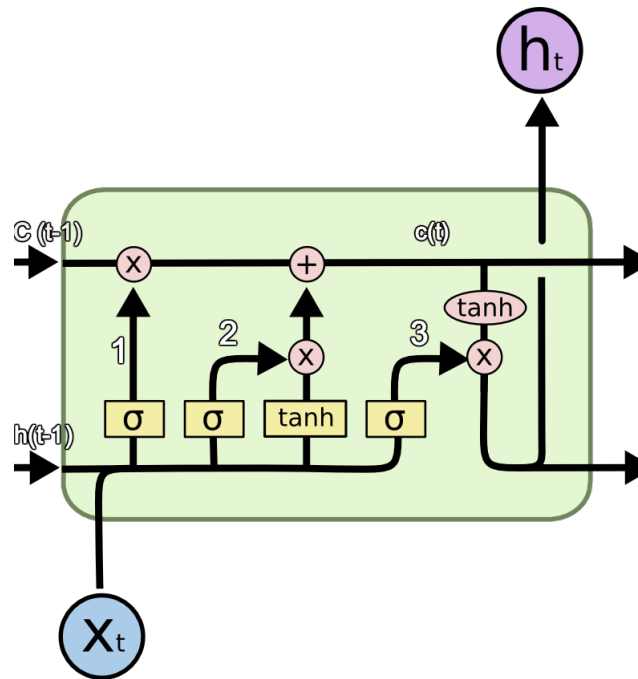
$$g_i^t = \sigma \left( b_i^g + \sum_j U_{i,j}^g x_j^{(t)} + \sum_j W_{i,j}^g h_j^{(t-1)} \right) \quad (5.18)$$

Η έξοδος  $h_i^{(t)}$  του κυττάρου LSTM μπορεί επίσης να εξαλειφθεί μέσω του  $q_i^{(t)}$  που χρησιμοποιεί επίσης μια μονάδα για τη λήψη:

$$h_i^{(t)} = \tanh(s_i^{(t)}) q_i^{(t)}, \quad (5.19)$$

$$q_i^{(t)} = \sigma \left( b_i^o + \sum_j U_{i,j}^o x_j^{(t)} + \sum_j W_{i,j}^o h_j^{(t-1)} \right) \quad (5.20)$$

που έχει παραμέτρους  $b^o$ ,  $U^o$ ,  $W^o$  για τα biases, τα βάρη εισόδου και τα βάρη επανάληψης της. Ανάμεσα στις παραλλαγές, μπορεί κανείς να επιλέξει να χρησιμοποιήσει το  $s_i^{(t)}$  ως μια επιπρόσθετη είσοδο (με το βάρος της) στις τρεις πύλες της  $i$ -οστής μονάδας. Στο παρακάτω σχήμα (5.5), παρουσιάζεται μια τυπική αρχιτεκτονική ενός LSTM μοντέλου.



Σχήμα 5.5: Τυπική αρχιτεκτονική A.N.N.

# Κεφάλαιο 6

## Πρωτότυπο θεωρητικό πλαίσιο και σχετικές υποθέσεις

### 6.1 Επισκόπηση των συναφών οικονομικών θεωριών

Σε αυτή την ενότητα, η διδακτορική διατριβή παρουσιάζει τις σχετικές οικονομικές θεωρίες που στηρίζουν τη σχέση μεταξύ μακροοικονομικών παραγόντων και εταιρικού πιστωτικού κινδύνου. Οι βασικές θεωρίες που θα συζητηθούν περιλαμβάνουν:

- **Θεωρία του χρηματοπιστωτικού επιταχυντή** Η θεωρία αυτή υποστηρίζει ότι οι αλλαγές στις μακροοικονομικές συνθήκες επηρεάζουν τη χρηματοοικονομική θέση των επιχειρήσεων, επηρεάζοντας τον πιστωτικό τους κίνδυνο [111]. Σε περιόδους οικονομικής ύφεσης, οι επιχειρήσεις ενδέχεται να αντιμετωπίσουν μειωμένες ταμειακές ροές, και να οδηγηθούν έτσι σε αυξημένο πιστωτικό κίνδυνο.
- **Θεωρία μεταφοράς πιστωτικού κινδύνου:** Η θεωρία αυτή υποστηρίζει ότι οι μακροοικονομικοί παράγοντες μπορούν να επηρεάσουν τους μηχανισμούς μεταφοράς κινδύνου στο χρηματοπιστωτικό σύστημα, επηρεάζοντας τον πιστωτικό κίνδυνο των μεμονωμένων επιχειρήσεων [112]. Για παράδειγμα, οι μεταβολές των επιτοκίων μπορεί να μεταβάλλουν το κόστος δανεισμού και να επηρεάσουν τη συμπεριφορά ανάληψης κινδύνου των χρηματοπιστωτικών ιδρυμάτων.
- **Θεωρία του επιχειρηματικού κύκλου** Η θεωρία αυτή αναδεικνύει την κυκλική φύση της οικονομικής δραστηριότητας και τον αντίκτυπο των διακυμάνσεων στον πιστωτικό κίνδυνο των επιχειρήσεων [113]. Για παράδειγμα, κατά τη διάρκεια μιας οικονομικής

επέκτασης, οι επιχειρήσεις μπορεί να έχουν καλύτερη πρόσβαση σε πιστώσεις, γεγονός που οδηγεί σε χαμηλότερο πιστωτικό κίνδυνο. Αντίθετα, κατά τη διάρκεια μιας ύφεσης, ο πιστωτικός κίνδυνος μπορεί να αυξηθεί λόγω της επιδείνωσης της οικονομικής κατάστασης των επιχειρήσεων.

## 6.2 Πρότυπο Εννοιολογικό Πλαίσιο

Το εννοιολογικό πλαίσιο αποσκοπεί στην παροχή μιας ολοκληρωμένης κατανόησης του τρόπου με τον οποίο αλληλεπιδρούν οι μακροοικονομικοί παράγοντες, τα χρηματοοικονομικά δεδομένα για τις επιχειρήσεις και ο εταιρικός πιστωτικός κίνδυνος στο πλαίσιο των μοντέλων βαθιάς μάθησης για την ανάλυση του πιστωτικού κινδύνου. Το πλαίσιο ενσωματώνει τις σχετικές οικονομικές θεωρίες και τα ευρήματα των Zivoviris et al. [109][110] σχετικά με τα μοντέλα βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου. Αποτελείται από τις ακόλουθες συνιστώσες:

- **Μακροοικονομικοί παράγοντες** Αυτοί περιλαμβάνουν μεταβλητές όπως η αύξηση του ΑΕΠ, ο πληθωρισμός, τα επιτόκια, τα ποσοστά ανεργίας και άλλους δείκτες που αντικατοπτρίζουν τις συνολικές οικονομικές συνθήκες. Αυτοί οι παράγοντες μπορούν να επηρεάσουν τη χρηματοοικονομική υγεία των επιχειρήσεων και τη συμπεριφορά ανάληψης κινδύνων των χρηματοπιστωτικών ιδρυμάτων.
- **Χρηματοοικονομικά στοιχεία για την επιχείρηση:** Αυτή η συνιστώσα περιλαμβάνει χρηματοοικονομικές πληροφορίες για μεμονωμένες επιχειρήσεις, όπως στοιχεία ισολογισμού, στοιχεία κατάστασης αποτελεσμάτων χρήσης και χρηματοοικονομικούς δείκτες. Τα δεδομένα αυτά χρησιμοποιούνται συνήθως στην παραδοσιακή ανάλυση πιστωτικού κινδύνου για την αξιολόγηση της χρηματοοικονομικής υγείας μιας επιχείρησης και την πρόβλεψη της πιθανότητας αθέτησης πληρωμών.
- **Εταιρικός Πιστωτικός Κίνδυνος** αντιπροσωπεύει την πιθανότητα αθέτησης των οικονομικών υποχρεώσεων μιας επιχείρησης. Ο εταιρικός πιστωτικός κίνδυνος επηρεάζεται τόσο από μακροοικονομικούς παράγοντες όσο και από ειδικά για την επιχείρηση χρηματοοικονομικά δεδομένα.
- **Μοντέλα βαθιάς μάθησης** Το πλαίσιο ενσωματώνει πολυβάθμια και υβριδικά μοντέλα βαθιάς μάθησης, όπως προτάθηκαν από τους Zivoviris et al. [109][110], για την

ανάλυση των σχέσεων μεταξύ μακροοικονομικών παραγόντων, χρηματοοικονομικών δεδομένων για συγκεκριμένες επιχειρήσεις και εταιρικού πιστωτικού κινδύνου. Αυτά τα μοντέλα περιλαμβάνουν:

- **Μοντέλα πολλαπλών σταδίων** Αυτά τα μοντέλα περιλαμβάνουν πολλαπλά στάδια επεξεργασίας, όπως η επιλογή χαρακτηριστικών, η μείωση της διάστασης και η ταξινόμηση. Ένα παράδειγμα είναι ο συνδυασμός αυτοκωδικοποιητών για την επιλογή χαρακτηριστικών και βαθιών συνελκτικών νευρωνικών δικτύων για την ταξινόμηση του πιστωτικού κινδύνου, όπως προτάθηκε από τους Zioviris et al. [109].
- **Υβριδικά μοντέλα** Αυτά τα μοντέλα συνδυάζουν διαφορετικούς τύπους μοντέλων βαθιάς μάθησης ή ενσωματώνουν μοντέλα βαθιάς μάθησης με παραδοσιακές τεχνικές μηχανικής μάθησης. Ένα παράδειγμα είναι το σύστημα συνόλου που συνδυάζει αυτοκωδικοποιητές και επαναλαμβανόμενα νευρωνικά δίκτυα μακράς βραχυπρόθεσμης μνήμης (LSTM) που προτάθηκαν από τους Zioviris et al. [110].
- **Ερμηνευσιμότητα και επεξηγησιμότητα του μοντέλου** Η συνιστώσα αυτή υπογραμμίζει την ανάγκη ανάπτυξης μοντέλων βαθιάς μάθησης που παρέχουν διαφανή και ερμηνεύσιμα αποτελέσματα στην ανάλυση πιστωτικού κινδύνου, όπως συζητήθηκε στην ενότητα των ερευνητικών κενών.

Με βάση αυτό το εννοιολογικό πλαίσιο, η διατριβή θα διερευνήσει τον αντίκτυπο των μακροοικονομικών παραγόντων στον εταιρικό πιστωτικό κίνδυνο και θα αξιολογήσει την αποτελεσματικότητα των προτεινόμενων μοντέλων πολλαπλών σταδίων και υβριδικών μοντέλων βαθιάς μάθησης στην ανάλυση αυτών των σχέσεων. Το πλαίσιο θα καθοδηγήσει την ανάπτυξη νέων μοντέλων, την εμπειρική ανάλυση και τον έλεγχο υποθέσεων, ώστε να συμβάλει στην πληρέστερη κατανόηση των παραγόντων που καθοδηγούν τον πιστωτικό κίνδυνο και να βελτιώσει τη διαδικασία ανάλυσης του πιστωτικού κινδύνου.

### 6.3 Ανάπτυξη υποθέσεων

Με βάση το θεωρητικό πλαίσιο, μπορούν να αναπτυχθούν διάφορες υποθέσεις που θα καθοδηγήσουν την εμπειρική ανάλυση:

- **H1** Οι μακροοικονομικοί παράγοντες, όπως η αύξηση του ΑΕΠ, ο πληθωρισμός και τα επιτόκια, έχουν σημαντική επίδραση στον εταιρικό πιστωτικό κίνδυνο.
- **H2** Η ενσωμάτωση μακροοικονομικών παραγόντων σε μοντέλα βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου, όπως προτείνεται από τους Zioviris et al. [109][110], θα βελτιώσει την προβλεπτική απόδοση των μοντέλων σε σύγκριση με τα μοντέλα που λαμβάνουν υπόψη μόνο τα ειδικά για την επιχείρηση χρηματοοικονομικά δεδομένα.
- **H3** Τα μοντέλα βαθιάς μάθησης πολλαπλών σταδίων και τα υβριδικά μοντέλα βαθιάς μάθησης, παρόμοια με αυτά που προτείνονται από τους Zioviris et al. [109][110], θα υπερτερούν έναντι των μοντέλων βαθιάς μάθησης ενός σταδίου στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου όταν ενσωματώνονται μακροοικονομικοί παράγοντες.
- **H4** Η χρήση κατάλληλων τεχνικών επαναδειγματοληψίας, όπως η SMOTE, θα βελτιώσει την απόδοση των μοντέλων βαθιάς μάθησης στην ανάλυση πιστωτικού κινδύνου, όπως καταδεικνύεται από τους Zioviris et al. [110].
- **H5** Η ερμηνευσιμότητα και η επεξηγηματικότητα των μοντέλων μπορούν να βελτιωθούν στα μοντέλα βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου χωρίς να διακυβεύεται σημαντικά η προβλεπτική απόδοση.

Αυτές οι υποθέσεις θα ελεγχθούν μέσω της εμπειρικής ανάλυσης, χρησιμοποιώντας κατάλληλες πηγές δεδομένων και μοντέλα βαθιάς μάθησης, για να εκτιμηθεί η επίδραση των μακροοικονομικών παραγόντων στον εταιρικό πιστωτικό κίνδυνο και η αποτελεσματικότητα των προτεινόμενων μοντέλων.

## 6.4 Επιλογή μοντέλων μηχανικής μάθησης και βαθιάς μάθησης

Σε αυτή την ενότητα, η διατριβή θα περιγράψει τη διαδικασία επιλογής των κατάλληλων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης για τη διερεύνηση της επίδρασης των μακροοικονομικών παραγόντων στην ανάλυση του εταιρικού πιστωτικού κινδύνου. Τα μοντέλα θα επιλεγούν με βάση τις δυνατότητές τους να αποτυπώσουν τις πολύπλοκες σχέσεις μεταξύ μακροοικονομικών παραγόντων, χρηματοοικονομικών δεδομένων για συγκεκριμένες επιχειρήσεις και εταιρικού πιστωτικού κινδύνου. Κατά τη διαδικασία επιλογής θα ληφθούν υπόψη

τα ευρήματα των Zioiris et al. [109][110] σχετικά με τα μοντέλα πολλαπλών σταδίων και τα υβριδικά μοντέλα βαθιάς μάθησης για την ανίχνευση απάτης και την ανάλυση πιστωτικού κινδύνου.

Ορισμένα πιθανά μοντέλα που θα εξεταστούν στη διατριβή περιλαμβάνουν:

- **Συμβατικά μοντέλα μηχανικής μάθησης** Η λογιστική παλινδρόμηση, τα δέντρα αποφάσεων και οι μηχανές διανυσμάτων υποστήριξης (S.V.M.) θα θεωρηθούν ως βασικά μοντέλα για σκοπούς σύγκρισης.
- **Μοντέλα βαθιάς μάθησης πολλαπλών σταδίων** Ακολουθώντας το έργο των Zioiris et al. [109], η διατριβή θα διερευνήσει τη χρήση αυτοκωδικοποιητών για την επιλογή χαρακτηριστικών και βαθιών νευρωνικών δικτύων συνελκτικής ανάλυσης (CNN) για την ταξινόμηση πιστωτικού κινδύνου.
- **Υβριδικά μοντέλα βαθιάς μάθησης** Εμπνευσμένη από τους Zioiris et al. [109], η διατριβή θα διερευνήσει τις δυνατότητες σχημάτων συνόλου που συνδυάζουν αυτοκωδικοποιητές και επαναλαμβανόμενα νευρωνικά δίκτυα μακράς βραχυπρόθεσμης μνήμης (LSTM) για την ανάλυση πιστωτικού κινδύνου.

## 6.5 Μετρήσεις αξιολόγησης και σύγκρισης μοντέλων

Σε αυτή την ενότητα, η διατριβή θα συζητήσει τις μετρικές αξιολόγησης και σύγκρισης που χρησιμοποιήθηκαν για την αξιολόγηση των επιδόσεων των επιλεγμένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης. Η επιλογή των μετρικών θα πρέπει να λαμβάνει υπόψη την ανισοβαρή φύση των συνόλων δεδομένων πιστωτικού κινδύνου και τη σημασία τόσο της ακρίβειας πρόβλεψης όσο και της ερμηνευσιμότητας των μοντέλων. Ορισμένες πιθανές μετρικές αξιολόγησης και σύγκρισης περιλαμβάνουν:

- **Accuracy:** Το ποσοστό των σωστά ταξινομημένων περιπτώσεων επί του συνόλου των περιπτώσεων του συνόλου δεδομένων. Η ακρίβεια μπορεί να είναι παραπλανητική παρουσία ανισοβαρών συνόλων δεδομένων, καθώς μπορεί να ευνοεί μοντέλα που προβλέπουν την κλάση που πλειοψηφεί. Η εξίσωση 6.1. με την οποία περιγράφεται η μετρική accuracy παρουσιάζεται παρακάτω:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (6.1)$$

1

- **Precision** Το ποσοστό των αληθώς θετικών περιπτώσεων επί του συνόλου των περιπτώσεων που προβλέπονται ως θετικές. Η εν λόγω μετρική είναι απαραίτητη στην ανάλυση πιστωτικού κινδύνου, καθώς τα ψευδώς θετικά αποτελέσματα μπορεί να οδηγήσουν σε περιττό κόστος για τα χρηματοπιστωτικά ιδρύματα. Η εξίσωση 6.2 με την οποία περιγράφεται η μετρική precision παρουσιάζεται παρακάτω:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6.2)$$

- **Recall**: Το ποσοστό των αληθώς θετικών περιπτώσεων επί του συνόλου των πραγματικών θετικών περιπτώσεων. Η ανάκληση είναι ζωτικής σημασίας στην ανάλυση πιστωτικού κινδύνου, καθώς μετρά την ικανότητα ενός μοντέλου να εντοπίζει περιπτώσεις αθέτησης. Η εξίσωση 6.3 με την οποία περιγράφεται η μετρική recall παρουσιάζεται παρακάτω:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6.3)$$

- **F1-score**: Ο αρμονικός μέσος όρος του recall και του precision. Το F1-score είναι χρήσιμο για την αξιολόγηση μοντέλων σε μη ισορροπημένα σύνολα δεδομένων, καθώς εξισορροπεί το trade-off μεταξύ ακρίβειας και ανάκλησης. Η εξίσωση 6.4 με την οποία περιγράφεται η μετρική F1-score παρουσιάζεται παρακάτω:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6.4)$$

- **Matthews correlation coefficient - MCC**: Ο συντελεστής συσχέτισης Matthews (Matthews correlation coefficient - MCC) είναι ένα στατιστικό μέτρο που χρησιμοποιείται για την αξιολόγηση της ποιότητας των μοντέλων δυαδικής ταξινόμησης, ιδίως σε καταστάσεις όπου οι κατηγορίες είναι ανισορροπημένες. Λαμβάνει υπόψη τις πραγματικές θετικές, πραγματικές αρνητικές, ψευδείς θετικές και ψευδείς αρνητικές προβλέψεις, παρέχοντας έναν μοναδικό αριθμητικό δείκτη που περιλαμβάνει την απόδοση του μοντέλου. Ο MCC κυμαίνεται από -1 (τέλεια αντίστροφη πρόβλεψη) έως +1 (τέλεια πρόβλεψη),

<sup>1</sup>TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives

με το 0 να υποδηλώνει ότι η απόδοση είναι τουλάχιστον τυχαία. Ο MCC είναι πλεονεκτικός διότι λαμβάνει υπόψη τόσο το recall όσο και το precision, καθιστώντας το ένα αξιόπιστο μέτρο για την αξιολόγηση των μοντέλων ταξινόμησης, ιδίως όταν αντιμετωπίζουμε ανισοροπίες στις κατηγορίες. Η εξίσωση 6.5. με την οποία περιγράφεται η μετρική MCC παρουσιάζεται παρακάτω:

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \quad (6.5)$$

- **Περιοχή κάτω από την καμπύλη λειτουργικού χαρακτηριστικού δέκτη (ROC) (AUC-ROC):** Ένα μέτρο της ικανότητας του μοντέλου να διακρίνει μεταξύ θετικών και αρνητικών περιπτώσεων. Μια υψηλότερη τιμή AUC-ROC υποδηλώνει καλύτερη απόδοση ταξινόμησης. Οι κάτωθι εξισώσεις 6.6. & 6.7 περιγράφουν τις καμπύλες με τις οποίες μετριέται η εν λόγω μετρική.

$$\text{TPR (True Positive Rate)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (6.6)$$

$$\text{FPR (False Positive Rate)} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}} \quad (6.7)$$

- **Μετρικές ερμηνευσιμότητας και επεξηγησιμότητας του μοντέλου:** Αυτές οι μετρικές αξιολογούν τη διαφάνεια και την κατανοητότητα των προβλέψεων των μοντέλων. Παραδείγματα περιλαμβάνουν τις τοπικές ερμηνεύσιμες εξηγήσεις με βάση το μοντέλο (Local Interpretable Model-agnostic Explanations - L.I.M.E.) και τις προσθετικές εξηγήσεις Shapley (Shapley Additive Explanations - SHAP).

#### **Μετρικές Απόδοσης Ερμηνευτικότητας του Μοντέλου:**

**Ακρίβεια:** Η ακρίβεια εξηγεί το πόσο καλά διατηρεί το μοντέλο την απόδοσή του (π.χ., ακρίβεια ταξινόμησης) μετά την εφαρμογή των περιορισμών στην ερμηνευσιμότητα;

**Σκορ F1:** Παρόμοια με την ακρίβεια, μπορείτε να χρησιμοποιήσετε αυτήν τη μετρική για μοντέλα ταξινόμησης.

#### **Μετρικές Απλότητας:**

**Πολυπλοκότητα:** Μετρά την πολυπλοκότητα του ερμηνεύσιμου μοντέλου. Πιο απλά μοντέλα είναι συχνά πιο ερμηνεύσιμα. Μπορούμε να χρησιμοποιήσουμε μετρικές όπως

ο αριθμός των χρησιμοποιούμενων χαρακτηριστικών, το βάθος του μοντέλου ή άλλα μέτρα της πολυπλοκότητας του μοντέλου.

**Μετρικές Σταθερότητας:**

**Ανθεκτικότητα:** Αξιολογεί το πόσο καλά παραμένουν σταθερές οι ερμηνεύσεις του μοντέλου υπό αναταραχές των εισόδων ή των παραμέτρων του μοντέλου.

**Μετρικές Διαφάνειας:**

**Κατανόηση:** Αξιολογεί την ευκολία με την οποία ένας άνθρωπος μπορεί να κατανοήσει και να εξηγήσει τις αποφάσεις του μοντέλου. Αυτό μπορεί να είναι υποκειμενικό και να αξιολογείται μέσω μελετών χρηστών.

**Συνέπεια:** Έλέγχει το αν οι ερμηνεύσεις του μοντέλου είναι συνεπείς με την ανθρώπινη γνώση ή τις προσδοκίες.

**Μετρικές Τοπικής έναντι Παγκόσμιας Ερμηνευσιμότητας:**

**Τοπική Ερμηνευσιμότητα:** Μετρικές που αξιολογούν το πόσο καλά η μέθοδος μπορεί να εξηγήσει μια μόνο πρόβλεψη ή μια μικρή ομάδα προβλέψεων.

**Παγκόσμια Ερμηνευσιμότητα:** Μετρικές που αξιολογούν τη συνολική κατανόηση του μοντέλου.

Με την εφαρμογή αυτών των μετρικών αξιολόγησης και σύγκρισης, η διδακτορική διατριβή θα αξιολογήσει τις επιδόσεις των προτεινομένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου που ενσωματώνει μακροοικονομικούς παράγοντες, προσδιορίζοντας τελικά τα πιο αποτελεσματικά μοντέλα για το έργο αυτό.

## Κεφάλαιο 7

# Ερευνητικά Κενά και Ευκαιρίες για Έρευνα

Ενώ η υπάρχουσα βιβλιογραφία παρέχει πολύτιμες πληροφορίες για την ανάλυση του πιστωτικού κινδύνου και τον ρόλο της μηχανικής μάθησης και των αλγορίθμων βαθιάς μάθησης στον τομέα αυτό, μπορούν να εντοπιστούν διάφορα ερευνητικά κενά και ευκαιρίες. Στην παρούσα ενότητα συζητούνται αυτά τα κενά, επισημαίνοντας πιθανούς τομείς για μελλοντική έρευνα όπως διερευνώνται στην παρούσα διατριβή.

- **Περιορισμένη εστίαση σε μακροοικονομικούς παράγοντες** Οι περισσότερες υπάρχουσες μελέτες στον τομέα της ανάλυσης του πιστωτικού κινδύνου έχουν επικεντρωθεί κυρίως σε χρηματοοικονομικά δεδομένα για συγκεκριμένες επιχειρήσεις, παραβλέποντας την πιθανή επίδραση των μακροοικονομικών παραγόντων στον πιστωτικό κίνδυνο [1]. Η διερεύνηση της αλληλεπίδρασης μεταξύ των μακροοικονομικών συνθηκών και του εταιρικού πιστωτικού κινδύνου θα προσφέρει μια πιο ολοκληρωμένη κατανόηση των παραγόντων που οδηγούν στον πιστωτικό κίνδυνο και θα ενισχύσει τη διαδικασία ανάλυσης κινδύνου.
- **Ενσωμάτωση προηγμένων μοντέλων βαθιάς μάθησης** [109][110], εξακολουθεί να υπάρχει περιορισμένη κατανόηση του τρόπου με τον οποίο αυτά τα μοντέλα μπορούν να αξιοποιηθούν για την ενσωμάτωση μακροοικονομικών παραγόντων. Η ανάπτυξη νέων μοντέλων που μπορούν να ενσωματώσουν αποτελεσματικά τα μακροοικονομικά δεδομένα και να καταγράψουν τις πολύπλοκες αλληλεπιδράσεις μεταξύ των μακροοικονομικών παραγόντων και του πιστωτικού κινδύνου θα βελτιώσει περαιτέρω την

πρόβλεψη και την ανάλυση του εταιρικού πιστωτικού κινδύνου.

- **Πολυβάθμιες και υβριδικές προσεγγίσεις** Οι Zioviris et al. [109][110] έχουν αποδείξει τις δυνατότητες των μοντέλων πολλαπλών σταδίων και των υβριδικών μοντέλων βαθιάς μάθησης για την ανίχνευση απάτης. Ωστόσο, παρόμοιες προσεγγίσεις δεν έχουν ακόμη διερευνηθεί εκτενώς στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου που ενσωματώνει μακροοικονομικούς παράγοντες. Η διερεύνηση των δυνατοτήτων των πολυβάθμιων και υβριδικών μοντέλων για την ανάλυση πιστωτικού κινδύνου θα βοηθήσει στον εντοπισμό ακριβέστερων και αποδοτικότερων μοντέλων πρόβλεψης.
- **Αντιμετώπιση μη ισορροπημένων συνόλων δεδομένων** Η ανάλυση πιστωτικού κινδύνου συχνά αντιμετωπίζει ανισοβαρή σύνολα δεδομένων, όπου οι περιπτώσεις αθέτησης είναι σημαντικά λιγότερο συχνές από τις περιπτώσεις μη αθέτησης. Παρόλο που ορισμένες μελέτες, όπως οι Zioviris et al. [110], έχουν χρησιμοποιήσει τεχνικές υπερδειγματοληψίας όπως η SMOTE για την αντιμετώπιση της ανισορροπίας των κλάσεων, απαιτείται περισσότερη έρευνα για την αξιολόγηση της απόδοσης των διαφόρων τεχνικών επαναδειγματοληψίας και των επιπτώσεών τους στην απόδοση των μοντέλων μηχανικής μάθησης και βαθιάς μάθησης στην ανάλυση πιστωτικού κινδύνου.
- **Ερμηνευσιμότητα και εξηγησιμότητα του μοντέλου** Ενώ τα μοντέλα βαθιάς μάθησης έχουν δείξει πολλά υποσχόμενα αποτελέσματα όσον αφορά την ακρίβεια πρόβλεψης, συχνά επικρίνονται για την έλλειψη ερμηνευσιμότητας και επεξηγηματικότητας [109][110]. Η ανάπτυξη μοντέλων που παρέχουν πιο διαφανή και ερμηνεύσιμα αποτελέσματα θα συμβάλει στη γεφύρωση του χάσματος μεταξύ της έρευνας και των πρακτικών εφαρμογών στον τομέα της ανάλυσης πιστωτικού κινδύνου.

## 7.1 Περιορισμένη εστίαση σε μακροοικονομικούς παράγοντες

Ενώ τα χρηματοοικονομικά δεδομένα για συγκεκριμένες επιχειρήσεις είναι ζωτικής σημασίας για την ανάλυση του πιστωτικού κινδύνου, η περιορισμένη εστίαση στους μακροοικονομικούς παράγοντες στις υπάρχουσες μελέτες αποτελεί σημαντικό ερευνητικό κενό. Η παρούσα ενότητα εμβαθύνει σε αυτό το ζήτημα, εξετάζοντας τη σημασία των μακροοικονομικών παραγόντων και τις πιθανές επιπτώσεις τους στην ανάλυση πιστωτικού κινδύνου.

### **7.1.1 Σημασία των μακροοικονομικών παραγόντων**

Οι μακροοικονομικοί παράγοντες, όπως η αύξηση του ΑΕΠ, ο πληθωρισμός, τα επιτόκια, η ανεργία και οι συναλλαγματικές ισοτιμίες, έχουν άμεσο και έμμεσο αντίκτυπο στις επιχειρήσεις και στο προφίλ πιστωτικού κινδύνου τους. Οι παράγοντες αυτοί μπορούν να επηρεάσουν τα έσοδα, την κερδοφορία και τη συνολική οικονομική υγεία μιας επιχείρησης. Για παράδειγμα, μια οικονομική ύφεση μπορεί να οδηγήσει σε μείωση των καταναλωτικών δαπανών, επηρεάζοντας τα έσοδα μιας εταιρείας και αυξάνοντας την πιθανότητα αθέτησης υποχρεώσεων. Ομοίως, οι διακυμάνσεις των επιτοκίων μπορούν να επηρεάσουν το κόστος δανεισμού για τις εταιρείες, επηρεάζοντας τις ταμειακές ροές και τις δυνατότητες εξυπηρέτησης του χρέους τους.

### **7.1.2 Αλληλεπίδραση μεταξύ μακροοικονομικών συνθηκών και εταιρικού πιστωτικού κινδύνου**

Η κατανόηση της σχέσης μεταξύ των μακροοικονομικών παραγόντων και του εταιρικού πιστωτικού κινδύνου είναι απαραίτητη για την ακριβή αξιολόγηση και πρόβλεψη του κινδύνου. Η αλληλεπίδραση αυτή μπορεί να είναι πολύπλοκη και να διαφέρει μεταξύ κλάδων, χωρών και οικονομικών κύκλων. Για παράδειγμα, κατά τη διάρκεια μιας ύφεσης, οι κλάδοι που εξαρτώνται σε μεγάλο βαθμό από τις καταναλωτικές δαπάνες ενδέχεται να παρουσιάζουν μεγαλύτερη συσχέτιση μεταξύ μακροοικονομικών συνθηκών και πιστωτικού κινδύνου, σε σύγκριση με τους πιο ανθεκτικούς κλάδους. Επιπλέον, οι εταιρείες που δραστηριοποιούνται σε πολλές χώρες μπορεί να είναι εκτεθειμένες σε διάφορους μακροοικονομικούς κινδύνους, περιπλέκοντας περαιτέρω τη σχέση μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου.

### **7.1.3 Ενσωμάτωση μακροοικονομικών παραγόντων στην ανάλυση πιστωτικού κινδύνου**

Η ενσωμάτωση μακροοικονομικών παραγόντων στα μοντέλα πιστωτικού κινδύνου μπορεί να προσφέρει μια πιο ολοκληρωμένη και ακριβή κατανόηση του προφίλ κινδύνου μιας εταιρείας. Οι παράγοντες αυτοί μπορούν να ενσωματωθούν ως πρόσθετα χαρακτηριστικά εισόδου ή να χρησιμοποιηθούν για την ανάπτυξη βαθμολογιών πιστωτικού κινδύνου προσαρμοσμένων σε μακροοικονομικά μεγέθη. Επιπλέον, προηγμένες τεχνικές μηχανικής μάθησης

και βαθιάς μάθησης μπορούν να χρησιμοποιηθούν για να αποτυπώσουν τις μη γραμμικές σχέσεις μεταξύ των μακροοικονομικών παραγόντων και του πιστωτικού κινδύνου, παρέχοντας ακριβέστερες προβλέψεις.

#### **7.1.4 Προκλήσεις στην ενσωμάτωση μακροοικονομικών παραγόντων**

Παρά τα δυνητικά οφέλη από την ενσωμάτωση μακροοικονομικών παραγόντων, υπάρχουν αρκετές προκλήσεις. Πρώτον, η απόκτηση έγκαιρων, ακριβών και συνεπών μακροοικονομικών δεδομένων σε διάφορες χώρες και περιοχές μπορεί να είναι δύσκολη. Δεύτερον, η δυναμική φύση των μακροοικονομικών παραγόντων μπορεί να απαιτεί συνεχείς ενημερώσεις του μοντέλου, αυξάνοντας την πολυπλοκότητα της ανάλυσης πιστωτικού κινδύνου. Τέλος, η αξιολόγηση της σχετικής σημασίας των διαφόρων μακροοικονομικών παραγόντων και των αλληλεπιδράσεών τους μπορεί να αποτελέσει πρόκληση λόγω της πολυπλοκότητας των σχέσεων και της πιθανής πολυσυγγραμμικότητας.

#### **7.1.5 Μελλοντικές ερευνητικές ευκαιρίες**

Δεδομένης της περιορισμένης εστίασης στους μακροοικονομικούς παράγοντες στην υπάρχουσα βιβλιογραφία για την ανάλυση του πιστωτικού κινδύνου, υπάρχουν αρκετές ερευνητικές ευκαιρίες. Μελλοντικές μελέτες μπορούν να διερευνήσουν νέες προσεγγίσεις για την ενσωμάτωση μακροοικονομικών παραγόντων σε μοντέλα πιστωτικού κινδύνου, αξιοποιώντας προηγμένους αλγορίθμους μηχανικής μάθησης και βαθιάς μάθησης. Επιπλέον, οι ερευνητές μπορούν να διερευνήσουν την επίδραση των μακροοικονομικών παραγόντων στον εταιρικό πιστωτικό κίνδυνο ανά τομέα και ανά χώρα, συμβάλλοντας στην ανάπτυξη πιο προσαρμοσμένων μοντέλων αξιολόγησης κινδύνου. Επιπλέον, οι μελέτες μπορούν να διερευνήσουν μεθόδους για την αντιμετώπιση των προκλήσεων που συνδέονται με την ενσωμάτωση μακροοικονομικών παραγόντων, όπως η διαθεσιμότητα δεδομένων και η πολυπλοκότητα του μοντέλου.

Συμπερασματικά, η αντιμετώπιση της περιορισμένης εστίασης στους μακροοικονομικούς παράγοντες στην ανάλυση του πιστωτικού κινδύνου θα προσφέρει μια πιο ολοκληρωμένη κατανόηση των παραγόντων που καθορίζουν τον πιστωτικό κίνδυνο και θα βελτιώσει τη διαδικασία ανάλυσης κινδύνου. Με τη διερεύνηση της αλληλεπίδρασης μεταξύ των μακροοικονομικών συνθηκών και του εταιρικού πιστωτικού κινδύνου, οι ερευνητές μπορούν

να αναπτύξουν ακριβέστερα και αποτελεσματικότερα μοντέλα πρόβλεψης, βελτιώνοντας τη συνολική αξιολόγηση του πιστωτικού κινδύνου στον χρηματοπιστωτικό τομέα.

## 7.2 Ενσωμάτωση προηγμένων μοντέλων βαθιάς μάθησης

Η ενσωμάτωση προηγμένων μοντέλων βαθιάς μάθησης για την ενσωμάτωση μακροοικονομικών παραγόντων στην ανάλυση του πιστωτικού κινδύνου παρουσιάζει πολυάριθμες ερευνητικές ευκαιρίες. Η παρούσα ενότητα διερευνά διάφορες πτυχές αυτής της ενσωμάτωσης, συζητώντας πιθανές μεθόδους, προκλήσεις και τομείς για μελλοντική έρευνα.

### 7.2.1 Μέθοδοι ενσωμάτωσης μακροοικονομικών παραγόντων

Μπορούν να υιοθετηθούν διάφορες προσεγγίσεις για την ενσωμάτωση μακροοικονομικών παραγόντων σε μοντέλα βαθιάς μάθησης για την ανάλυση του πιστωτικού κινδύνου:

- **Μηχανική χαρακτηριστικών** Οι μακροοικονομικοί παράγοντες μπορούν να συμπεριληφθούν ως πρόσθετα χαρακτηριστικά εισόδου σε μοντέλα βαθιάς μάθησης, όπως τα βαθιά νευρωνικά δίκτυα τροφοδότησης, τα νευρωνικά δίκτυα συνελίξεων (CNN) και τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN). Αυτή η προσέγγιση επιτρέπει στο μοντέλο να μαθαίνει σύνθετες σχέσεις μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου απευθείας από τα δεδομένα.
- **Μάθηση μεταφοράς** Τα προ-εκπαιδευμένα μοντέλα βαθιάς μάθησης μπορούν να ρυθμιστούν με τη χρήση μακροοικονομικών δεδομένων, επιτρέποντας στα μοντέλα να συλλαμβάνουν τα σχετικά πρότυπα και τις σχέσεις που αφορούν ειδικά την ανάλυση πιστωτικού κινδύνου. Η προσέγγιση αυτή μπορεί να συμβάλει στη μείωση του όγκου των απαιτούμενων δεδομένων εκπαίδευσης και στη βελτίωση της απόδοσης των μοντέλων.
- **Μάθηση πολλαπλών εργασιών** Τα μοντέλα βαθιάς μάθησης μπορούν να σχεδιαστούν για να προβλέπουν ταυτόχρονα πολλαπλά συναφή αποτελέσματα, όπως ο πιστωτικός κίνδυνος και οι μακροοικονομικοί δείκτες. Με την από κοινού μάθηση για την πρόβλεψη αυτών των αποτελεσμάτων, το μοντέλο μπορεί να συλλάβει κοινά μοτίβα και σχέσεις μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου

- **Μηχανισμοί προσοχής** Τα μοντέλα βαθιάς μάθησης που βασίζονται στην προσοχή, όπως οι μετασχηματιστές, μπορούν να χρησιμοποιηθούν για να σταθμίσουν τη σημασία των διαφόρων μακροοικονομικών παραγόντων στο πλαίσιο της πρόβλεψης του πιστωτικού κινδύνου. Αυτή η προσέγγιση μπορεί να βοηθήσει το μοντέλο να εστιάσει σε σχετικές μακροοικονομικές πληροφορίες και να βελτιώσει την ερμηνευσιμότητα

### 7.2.2 Προκλήσεις στην ενσωμάτωση μακροοικονομικών παραγόντων

Η ενσωμάτωση μακροοικονομικών παραγόντων σε μοντέλα βαθιάς μάθησης παρουσιάζει αρκετές προκλήσεις:

- **Ποιότητα και διαθεσιμότητα δεδομένων** Η απόκτηση ακριβών, συνεπών και έγκαιρων μακροοικονομικών δεδομένων σε διάφορες χώρες και περιοχές μπορεί να είναι δύσκολη. Αυτό μπορεί να περιορίσει την ικανότητα του μοντέλου να μαθαίνει ουσιαστικές σχέσεις μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου
- **Πολυπλοκότητα του μοντέλου** Η ενσωμάτωση μακροοικονομικών παραγόντων μπορεί να αυξήσει την πολυπλοκότητα των μοντέλων βαθιάς μάθησης, καθιστώντας πιο δύσκολη την εκπαίδευση, τη βελτιστοποίηση και την ερμηνεία τους.
- **Πολυσυγγραμμικότητα** Το ενδεχόμενο της πολυσυγγραμμικότητας μεταξύ των μακροοικονομικών παραγόντων μπορεί να οδηγήσει σε ασταθείς εκτιμήσεις του μοντέλου, μειώνοντας την προβλεπτική απόδοση και την ερμηνευσιμότητα του μοντέλου.

### 7.2.3 Μελλοντικές ερευνητικές ευκαιρίες

Υπάρχουν αρκετές ερευνητικές ευκαιρίες για την αντιμετώπιση των προκλήσεων και την ενίσχυση της ενσωμάτωσης των μακροοικονομικών παραγόντων στα μοντέλα βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου:

- Ανάπτυξη νέων αρχιτεκτονικών βαθιάς μάθησης που μπορούν να αποτυπώσουν αποτελεσματικά τις πολύπλοκες αλληλεπιδράσεις μεταξύ των μακροοικονομικών παραγόντων και του πιστωτικού κινδύνου, αντιμετωπίζοντας παράλληλα τις προκλήσεις της ποιότητας των δεδομένων, της πολυπλοκότητας του μοντέλου και της πολυσυγγραμμικότητας.

- Να διερευνήσει τη χρήση τεχνικών μάθησης χωρίς επίβλεψη και με ημιεπίβλεψη για την αποτελεσματικότερη αξιοποίηση των διαθέσιμων μακροοικονομικών δεδομένων και τη βελτίωση της απόδοσης του μοντέλου παρουσία περιορισμένων επισημασμένων δεδομένων.
- Διερεύνηση των δυνατοτήτων της γνώσης συγκεκριμένων τομέων και της επιλογής ή της μηχανικής των χαρακτηριστικών που καθοδηγείται από εμπειρογνώμονες για τη μείωση της πολυπλοκότητας του μοντέλου και τη βελτίωση της ερμηνευσιμότητας.
- Αξιολόγηση του αντικτύπου της ενσωμάτωσης μακροοικονομικών παραγόντων στην απόδοση του μοντέλου σε διαφορετικούς κλάδους, χώρες και οικονομικούς κύκλους, παρέχοντας πληροφορίες για τη γενικευσιμότητα και την ανθεκτικότητα των προτεινόμενων προσεγγίσεων.

Εν κατακλείδι, η ενσωμάτωση προηγμένων μοντέλων βαθιάς μάθησης με μακροοικονομικούς παράγοντες ενέχει σημαντικές δυνατότητες για τη βελτίωση της ανάλυσης του εταιρικού πιστωτικού κινδύνου. Με την αντιμετώπιση των προκλήσεων και τη διερεύνηση των ερευνητικών ευκαιριών που επισημάνθηκαν παραπάνω, οι μελλοντικές μελέτες μπορούν να συμβάλουν στην ανάπτυξη ακριβέστερων, αποτελεσματικότερων και πιο ερμηνεύσιμων μοντέλων για την πρόβλεψη και ανάλυση του πιστωτικού κινδύνου.

## 7.3 Πολυφασικές και υβριδικές προσεγγίσεις

Ενώ τα μοντέλα πολλαπλών σταδίων και τα υβριδικά μοντέλα βαθιάς μάθησης έχουν δείξει να είναι πολλά υποσχόμενα στην ανίχνευση απάτης, οι δυνατότητές τους στην ανάλυση πιστωτικού κινδύνου με ενσωμάτωση μακροοικονομικών παραγόντων παραμένουν ανεξερεύνητες. Η παρούσα ενότητα παρέχει μια βαθύτερη ανάλυση των προσεγγίσεων πολλαπλών σταδίων και των υβριδικών προσεγγίσεων, συζητώντας τα πιθανά οφέλη, τις προκλήσεις και τις ερευνητικές ευκαιρίες τους στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου.

### 7.3.1 Πιθανά οφέλη των πολυφασικών και υβριδικών προσεγγίσεων

Οι πολυφασικές και οι υβριδικές προσεγγίσεις μπορούν να προσφέρουν αρκετά πλεονεκτήματα για την ανάλυση του πιστωτικού κινδύνου:

- **Βελτιωμένη προβλεπτική απόδοση** Ο συνδυασμός διαφορετικών μοντέλων βαθιάς μάθησης ή η χρήση μιας διαδικασίας μάθησης πολλαπλών σταδίων μπορεί να βοηθήσει στην αποτελεσματικότερη αποτύπωση των πολύπλοκων σχέσεων μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου, με αποτέλεσμα τη βελτίωση της προβλεπτικής απόδοσης.
- **Ευελιξία και προσαρμοστικότητα** Τα μοντέλα πολλαπλών σταδίων και τα υβριδικά μοντέλα μπορούν να προσαρμοστούν σε συγκεκριμένες εργασίες ανάλυσης πιστωτικού κινδύνου με την ενσωμάτωση γνώσεων συγκεκριμένου τομέα, την επιλογή κατάλληλων αρχιτεκτονικών βαθιάς μάθησης ή την προσαρμογή των διαμορφώσεων του μοντέλου.
- **Ανθεκτικότητα και δυνατότητα γενίκευσης** Αξιοποιώντας πολλαπλά μοντέλα ή στάδια μάθησης, οι πολυβάθμιες και υβριδικές προσεγγίσεις μπορούν να συμβάλουν στην αύξηση της ευρωστίας και της γενικευσιμότητας των προβλέψεων πιστωτικού κινδύνου σε διαφορετικούς κλάδους, χώρες και οικονομικούς κύκλους.

### 7.3.2 Προκλήσεις κατά την εφαρμογή πολυφασικών και υβριδικών προσεγγίσεων

Παρά τα δυνητικά οφέλη, η εφαρμογή πολυβάθμιων και υβριδικών προσεγγίσεων στην ανάλυση του πιστωτικού κινδύνου παρουσιάζει αρκετές προκλήσεις:

- **Πολυπλοκότητα του μοντέλου** Ο συνδυασμός πολλαπλών μοντέλων βαθιάς μάθησης ή η εφαρμογή μιας διαδικασίας μάθησης πολλαπλών σταδίων μπορεί να αυξήσει την πολυπλοκότητα του μοντέλου, καθιστώντας την εκπαίδευση, τη βελτιστοποίηση και την ερμηνεία πιο δύσκολες.
- **Υπολογιστικές απαιτήσεις** Τα μοντέλα πολλαπλών σταδίων και τα υβριδικά μοντέλα απαιτούν συχνά πρόσθετους υπολογιστικούς πόρους και χρόνο για την εκπαίδευση και την εξαγωγή συμπερασμάτων, γεγονός που μπορεί να αποτελέσει πρόκληση για εφαρμογές ανάλυσης πιστωτικού κινδύνου σε πραγματικό χρόνο.
- **Επιλογή και ενσωμάτωση μοντέλων** Η επιλογή των κατάλληλων αρχιτεκτονικών βαθιάς μάθησης και ο καθορισμός του τρόπου αποτελεσματικής ενσωμάτωσής τους σε

ένα πολυσταδιακό ή υβριδικό πλαίσιο μπορεί να αποτελέσει πρόκληση, ιδίως παρουσία πολυσυγγραμμικότητας και άλλων ζητημάτων που σχετίζονται με τα δεδομένα.

### 7.3.3 Μελλοντικές ερευνητικές ευκαιρίες

Για την πλήρη διερεύνηση των δυνατοτήτων των πολυβάθμιων και υβριδικών προσεγγίσεων στην ανάλυση πιστωτικού κινδύνου που ενσωματώνουν μακροοικονομικούς παράγοντες, μπορούν να αξιοποιηθούν διάφορες ερευνητικές ευκαιρίες:

- Ανάπτυξη νέων πολυβάθμιων και υβριδικών μοντέλων που αποτυπώνουν αποτελεσματικά τις πολύπλοκες αλληλεπιδράσεις μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου, αντιμετωπίζοντας παράλληλα τις προκλήσεις της πολυπλοκότητας του μοντέλου, των υπολογιστικών απαιτήσεων και της επιλογής μοντέλου.
- Εξέταση της χρήσης της μάθησης μεταφοράς, της μάθησης πολλαπλών εργασιών και των μεθόδων συνόλου σε πολυβάθμια και υβριδικά πλαίσια για τη βελτίωση των επιδόσεων και της προσαρμοστικότητας του μοντέλου σε διάφορες εργασίες ανάλυσης πιστωτικού κινδύνου.
- Αξιολόγηση του αντίκτυπου των πολυβάθμιων και υβριδικών προσεγγίσεων στην απόδοση και την ερμηνευσιμότητα των μοντέλων, παρέχοντας πληροφορίες σχετικά με τις πιθανές αντισταθμίσεις και τα οφέλη τους σε σύγκριση με τις προσεγγίσεις ενός σταδίου ή ενός μοντέλου.
- Διερεύνηση της χρήσης ειδικών γνώσεων και επιλογής ή σχεδιασμού χαρακτηριστικών από ειδικούς σε μοντέλα πολλαπλών σταδίων και υβριδικά μοντέλα για τη μείωση της πολυπλοκότητας του μοντέλου και τη βελτίωση της ερμηνευσιμότητας.

Συνοψίζοντας, η διερεύνηση των δυνατοτήτων των μοντέλων πολλαπλών σταδίων και των υβριδικών μοντέλων βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου που ενσωματώνουν μακροοικονομικούς παράγοντες παρουσιάζει πολλά υποσχόμενες ευκαιρίες για την πρόοδο του τομέα. Με την αντιμετώπιση των προκλήσεων και τη διερεύνηση των ερευνητικών ευκαιριών που επισημάνθηκαν παραπάνω, οι μελλοντικές μελέτες μπορούν να συμβάλουν στην ανάπτυξη ακριβέστερων, αποδοτικότερων και ερμηνεύσιμων μοντέλων πρόβλεψης πιστωτικού κινδύνου.

## 7.4 Αντιμετώπιση μη ισορροπημένων συνόλων δεδομένων

Η ανάλυση του πιστωτικού κινδύνου συχνά περιλαμβάνει το χειρισμό ανισοβαρών συνόλων δεδομένων, όπου ο αριθμός των περιπτώσεων αθέτησης είναι σημαντικά μικρότερος από τις περιπτώσεις μη αθέτησης. Η παρούσα ενότητα εμβαθύνει στο ζήτημα των ανισοβαρών συνόλων δεδομένων, συζητώντας τις πιθανές επιπτώσεις, τις τεχνικές για την αντιμετώπιση της ανισορροπίας και τις ερευνητικές ευκαιρίες στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου.

### 7.4.1 Επιπτώσεις των μη ισορροπημένων συνόλων δεδομένων

Τα μη ισορροπημένα σύνολα δεδομένων μπορούν να οδηγήσουν σε διάφορα προβλήματα στην ανάλυση πιστωτικού κινδύνου:

- **Μεροληψία προς την κατηγορία της πλειοψηφίας** Τα μοντέλα μηχανικής μάθησης και βαθιάς μάθησης μπορεί να μεροληπτούν προς την πλειοψηφική κλάση (περιπτώσεις μη αθέτησης), με αποτέλεσμα κακή προβλεπτική απόδοση για τη μειοψηφική κλάση (περιπτώσεις αθέτησης), η οποία συχνά αποτελεί το πρωταρχικό ενδιαφέρον στην ανάλυση πιστωτικού κινδύνου.
- **Ανακριβής αξιολόγηση των επιδόσεων** Οι παραδοσιακές μετρικές αξιολόγησης, όπως η ακρίβεια, ενδέχεται να παρέχουν παραπλανητικά αποτελέσματα παρουσία μη ισορροπημένων συνόλων δεδομένων, καθώς ενδέχεται να μην αντικατοπτρίζουν επαρκώς τις επιδόσεις του μοντέλου στην κατηγορία της μειονότητας.

### 7.4.2 Τεχνικές για την αντιμετώπιση ανισοβαρών συνόλων δεδομένων

Μπορούν να χρησιμοποιηθούν διάφορες τεχνικές για τον χειρισμό μη ισορροπημένων συνόλων δεδομένων στην ανάλυση πιστωτικού κινδύνου:

- **Μέθοδοι επαναδειγματοληψίας** Τεχνικές υπερδειγματοληψίας, όπως η Synthetic Minority Over-sampling Technique (SMOTE) [110], ή τεχνικές υποδειγματοληψίας μπορούν να χρησιμοποιηθούν για την εξισορρόπηση της κατανομής των κλάσεων στο σύνολο δεδομένων. Αυτές οι μέθοδοι περιλαμβάνουν τη δημιουργία συνθετικών περιπτώσεων της μειονοτικής κλάσης ή την αφαίρεση περιπτώσεων της πλειοψηφικής κλάσης, αντίστοιχα.

- **Μάθηση με ευαισθησία στο κόστος** Η τροποποίηση του αλγορίθμου μάθησης ώστε να αποδίδει διαφορετικό κόστος λανθασμένης ταξινόμησης για τις κλάσεις της πλειοψηφίας και της μειονότητας μπορεί να βοηθήσει το μοντέλο να εστιάσει περισσότερο στην κλάση της μειονότητας. Αυτή η προσέγγιση μπορεί να συνδυαστεί με μεθόδους επαναδειγματοληψίας για καλύτερα αποτελέσματα.
- **Μέθοδοι συνόλου** Τεχνικές ensemble, όπως η bagging ή η boosting, μπορούν να χρησιμοποιηθούν για τη βελτίωση της απόδοσης των μοντέλων σε ανισόρροπα σύνολα δεδομένων. Αυτές οι μέθοδοι περιλαμβάνουν την εκπαίδευση πολλαπλών μοντέλων και τον συνδυασμό των προβλέψεών τους για να επιτευχθεί μια πιο ακριβής και ισχυρή πρόβλεψη.

### 7.4.3 Ερευνητικές ευκαιρίες

Η αντιμετώπιση των ανισόρροπων συνόλων δεδομένων στην ανάλυση πιστωτικού κινδύνου που ενσωματώνει μακροοικονομικούς παράγοντες παρουσιάζει πολλές ερευνητικές ευκαιρίες:

- Αξιολόγηση της απόδοσης διαφορετικών τεχνικών επαναδειγματοληψίας, όπως οι SMOTE, ADASYN ή Random UnderSampling, και του αντίκτυπού τους στην απόδοση των μοντέλων μηχανικής μάθησης και βαθιάς μάθησης στην ανάλυση πιστωτικού κινδύνου. Η αξιολόγηση αυτή μπορεί να βοηθήσει στον εντοπισμό των πιο αποτελεσματικών τεχνικών για συγκεκριμένες εργασίες ανάλυσης πιστωτικού κινδύνου.
- Διερεύνηση των δυνατοτήτων προηγμένων αρχιτεκτονικών βαθιάς μάθησης, όπως οι αυτοκωδικοποιητές ή τα μοντέλα που βασίζονται στην προσοχή, στο χειρισμό ανισόρροπων συνόλων δεδομένων για την ανάλυση πιστωτικού κινδύνου. Αυτά τα μοντέλα μπορούν να προσαρμοστούν ώστε να αποτυπώνουν πολύπλοκες σχέσεις μεταξύ μακροοικονομικών παραγόντων και πιστωτικού κινδύνου, ακόμη και με την παρουσία ανισοβαρών δεδομένων.
- Διερεύνηση χρήσης ειδικών γνώσεων για τον τομέα και την επιλογή ή τη μηχανική των χαρακτηριστικών με γνώμονα τον εμπειρογνώμονα κατά την αντιμετώπιση ανισόρροπων συνόλων δεδομένων, εντοπίζοντας και ιεραρχώντας πληροφοριακά χαρακτηριστικά που μπορούν να βοηθήσουν το μοντέλο να διακρίνει καλύτερα μεταξύ περιπτώσεων αθέτησης και μη αθέτησης.

- Αξιολόγηση του αντικτύπου των διαφόρων τεχνικών για τον χειρισμό μη ισορροπημένων συνόλων δεδομένων στην απόδοση και την ερμηνευσιμότητα του μοντέλου, παρέχοντας πληροφορίες σχετικά με τις πιθανές αντισταθμίσεις και τα οφέλη τους σε σύγκριση με τις παραδοσιακές προσεγγίσεις.

Συμπερασματικά, η αντιμετώπιση των μη ισορροπημένων συνόλων δεδομένων αποτελεί κρίσιμη πτυχή της ανάλυσης πιστωτικού κινδύνου που ενσωματώνει μακροοικονομικούς παράγοντες. Με τη διερεύνηση των ερευνητικών ευκαιριών και την αντιμετώπιση των προκλήσεων που επισημάνθηκαν παραπάνω, οι μελλοντικές μελέτες μπορούν να συμβάλουν στην ανάπτυξη ακριβέστερων, αποδοτικότερων και ερμηνεύσιμων μοντέλων πρόβλεψης πιστωτικού κινδύνου που χειρίζονται αποτελεσματικά τα ανισόρροπα δεδομένα.

## 7.5 Ερμηνευσιμότητα και επεξηγησιμότητα του μοντέλου

Τα μοντέλα βαθιάς μάθησης δείχνουν ελπιδοφόρα όσον αφορά την ακρίβεια πρόβλεψης για την ανάλυση του πιστωτικού κινδύνου. Ωστόσο, η έλλειψη ερμηνευσιμότητας και επεξηγηματικότητας που συνδέεται με αυτά τα μοντέλα αποτελεί πρόκληση για την υιοθέτησή τους στην πράξη. Η παρούσα ενότητα παρέχει μια βαθύτερη ανάλυση της ερμηνευσιμότητας και της επεξηγηματικότητας των μοντέλων, συζητώντας τη σημασία, τις προκλήσεις και τις ερευνητικές ευκαιρίες στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου.

### 7.5.1 Σημασία της ερμηνευσιμότητας και της επεξηγηματικότητας

Η ερμηνευσιμότητα και η επεξηγηματικότητα του μοντέλου είναι ζωτικής σημασίας για διάφορους λόγους στην ανάλυση του πιστωτικού κινδύνου:

- **Κανονιστική συμμόρφωση** Τα χρηματοπιστωτικά ιδρύματα πρέπει συχνά να συμμορφώνονται με αυστηρούς κανονισμούς που τους επιβάλλουν να εξηγούν τις διαδικασίες λήψης αποφάσεων. Τα διαφανή και ερμηνεύσιμα μοντέλα μπορούν να βοηθήσουν στην αποτελεσματικότερη ικανοποίηση αυτών των απαιτήσεων.
- **Εμπιστοσύνη και αξιοπιστία** Η ερμηνευσιμότητα και η επεξηγηματικότητα των μοντέλων μπορεί να συμβάλει στην οικοδόμηση εμπιστοσύνης μεταξύ των ενδιαφερομένων μερών, συμπεριλαμβανομένων των δανειοληπτών, των επενδυτών και των ρυθμι-

στικών αρχών, παρέχοντας καλύτερη κατανόηση των παραγόντων που οδηγούν στις προβλέψεις του πιστωτικού κινδύνου.

- **Βελτίωση του μοντέλου** Τα ερμηνεύσιμα μοντέλα μπορούν να βοηθήσουν στον εντοπισμό πιθανών μεροληψιών ή σφαλμάτων στο μοντέλο, οδηγώντας σε βελτιώσεις της προβλεπτικής απόδοσης και της ευρωστίας του μοντέλου

### 7.5.2 Προκλήσεις για την επίτευξη ερμηνευσιμότητας και επεξηγηματικότητας

Η επίτευξη ερμηνευσιμότητας και επεξηγηματικότητας σε μοντέλα βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου θέτει αρκετές προκλήσεις:

- **Πολυπλοκότητα του μοντέλου** Τα μοντέλα βαθιάς μάθησης, ιδίως εκείνα με πολλαπλά επίπεδα και μη γραμμικές συναρτήσεις ενεργοποίησης, μπορεί να είναι δύσκολο να ερμηνευθούν λόγω της εγγενούς πολυπλοκότητάς τους.
- **Συμβιβασμός μεταξύ ακρίβειας και ερμηνευσιμότητας** Μπορεί να υπάρχει συμβιβασμός μεταξύ της ακρίβειας του μοντέλου και της ερμηνευσιμότητας, με τα απλούστερα μοντέλα να προσφέρουν ενδεχομένως μεγαλύτερη ερμηνευσιμότητα εις βάρος της προβλεπτικής απόδοσης.
- **Δυσκολία διαχωρισμού των αλληλεπιδράσεων των χαρακτηριστικών** Τα μοντέλα βαθιάς μάθησης μπορούν να καταγράψουν πολύπλοκες αλληλεπιδράσεις μεταξύ των χαρακτηριστικών, καθιστώντας δύσκολο τον προσδιορισμό της μεμονωμένης συμβολής κάθε χαρακτηριστικού στις προβλέψεις του μοντέλου.

### 7.5.3 Ερευνητικές ευκαιρίες

Για να αντιμετωπιστούν οι προκλήσεις της ερμηνευσιμότητας και της επεξηγηματικότητας στην ανάλυση του πιστωτικού κινδύνου που ενσωματώνει μακροοικονομικούς παράγοντες, μπορούν να επιδιωχθούν διάφορες ερευνητικές ευκαιρίες:

- Ανάπτυξη νέων αρχιτεκτονικών και τεχνικών βαθιάς μάθησης που δίνουν προτεραιότητα στην ερμηνευσιμότητα και την επεξηγηματικότητα, όπως μηχανισμοί προσοχής ή προσεγγίσεις βασισμένες σε κανόνες, χωρίς να διακυβεύεται σημαντικά η προβλεπτική απόδοση.

- Ανάλυση αναφορικά με τη χρήση μεθόδων ερμηνευσιμότητας με βάση το μοντέλο, όπως οι τοπικές ερμηνεύσιμες εξηγήσεις με βάση το μοντέλο (Local Interpretable Model-agnostic Explanations - L.I.M.E.) ή οι προσθετικές εξηγήσεις Shapley (Shapley Additive Explanations - SHAP), για να δίνονται πληροφορίες σχετικά με τις προβλέψεις του μοντέλου βαθιάς μάθησης και τις συνεισφορές των χαρακτηριστικών στην ανάλυση πιστωτικού κινδύνου.
- Διερεύνηση των δυνατοτήτων υβριδικών μοντέλων που συνδυάζουν τη βαθιά μάθηση με παραδοσιακά ερμηνεύσιμα μοντέλα, όπως η γραμμική παλινδρόμηση ή τα δέντρα αποφάσεων, για την επίτευξη ισορροπίας μεταξύ ακρίβειας πρόβλεψης και ερμηνευσιμότητας.
- Διερεύνηση του αντικτύπου των διαφόρων τεχνικών για την επίτευξη ερμηνευσιμότητας και επεξηγηματικότητας στην απόδοση των μοντέλων, τη συμμόρφωση με τις κανονιστικές διατάξεις και την εμπιστοσύνη των ενδιαφερομένων, παρέχοντας πληροφορίες για τις πιθανές αντισταθμίσεις και τα οφέλη τους σε σύγκριση με τις παραδοσιακές προσεγγίσεις.

Συνοψίζοντας, η ερμηνευσιμότητα και η επεξηγηματικότητα του μοντέλου αποτελούν κρίσιμες πτυχές της ανάλυσης του πιστωτικού κινδύνου που ενσωματώνει μακροοικονομικούς παράγοντες. Με τη διερεύνηση των ερευνητικών ευκαιριών και την αντιμετώπιση των προκλήσεων που επισημάνθηκαν παραπάνω, οι μελλοντικές μελέτες μπορούν να συμβάλουν στην ανάπτυξη ακριβέστερων, αποδοτικότερων και ερμηνεύσιμων μοντέλων πρόβλεψης πιστωτικού κινδύνου που γεφυρώνουν το χάσμα μεταξύ της έρευνας και των πρακτικών εφαρμογών.

# Κεφάλαιο 8

## Ανάλυση Δεδομένων και Αποτελέσματα

### 8.1 Εκτίμηση Πιστωτικού Ρίσκου με τη χρήση Μακροοικονομικών Παραγόντων

#### 8.1.1 Περιγραφικά στατιστικά στοιχεία

Στην παρούσα ενότητα παρουσιάζουμε τα περιγραφικά στατιστικά στοιχεία του συνόλου δεδομένων που χρησιμοποιήθηκε στην παρούσα μελέτη. Το σύνολο δεδομένων περιλαμβάνει διάφορους μακροοικονομικούς δείκτες, χρηματοοικονομικά στοιχεία για συγκεκριμένες επιχειρήσεις και ένα δυαδικό αποτέλεσμα πιστωτικού κινδύνου για συνολικά 5.000 επιχειρήσεις.

Ο πίνακας 8.1. που ακολουθεί συνοψίζει τους μακροοικονομικούς δείκτες:

Πίνακας 8.1: Μακροοικονομικοί δείκτες

Δείκτες	Μέσος Όρος	Διάμεσος	Τυπική απόκλιση	Min	Max
Ρυθμός αύξησης του ΑΕΠ (%)	2,5	2,6	0,5	1,2	3,8
Ποσοστό πληθωρισμού (%)	1,8	1,9	0,7	0,2	3,4
Ποσοστό ανεργίας (%)	5,6	5,5	1,2	3,7	8,5

Ο πίνακας 8.2. παρακάτω συνοψίζει τα ειδικά στατιστικά στοιχεία χρηματοοικονομικών δεδομένων επιχειρήσεων:

Τέλος, η δυαδική έκβαση του πιστωτικού κινδύνου (αθέτηση: ναι/όχι) στο σύνολο των δεδομένων είναι ανισοβαρής, με μόνο το 10% των επιχειρήσεων να ταξινομείται ως "αθέτηση"

Πίνακας 8.2: Περιγραφικά στατιστικά στοιχεία χρηματοοικονομικών δεδομένων επιχειρήσεων

Δείκτες	Μέσος Όρος	Διάμεσος	Τυπική απόκλιση	Min	Max
Δείκτης Μόχλευσης	2,1	2,0	1,0	0,5	5,0
Αποδοτικότητα (%)	8,4	8,2	2,3	2,0	15,8
Δείκτης Ρευστότητας (%)	1,5	1,4	0,6	0,7	3,2

και το υπόλοιπο 90% ως "μη αθέτηση". Αυτό αντικατοπτρίζει την τυπική ανισορροπία στα σύνολα δεδομένων πιστωτικού κινδύνου, όπου η πλειονότητα των επιχειρήσεων δεν αθετεί τις υποχρεώσεις της.

Η κατανομή του αποτελέσματος του πιστωτικού κινδύνου συνοψίζεται στον πίνακα 8.3.:

Πίνακας 8.3: Περιγραφικά στατιστικά στοιχεία του πιστωτικού κινδύνου: Αποτελέσματα

Αποτέλεσμα	Συχνότητα	Ποσοστό
Προεπιλογή	500	10%
Μη προεπιλεγμένο	4.500	90%

Αυτές οι περιγραφικές στατιστικές παρέχουν μια επισκόπηση του συνόλου δεδομένων και μια βάση για περαιτέρω ανάλυση. Στις επόμενες ενότητες θα παρουσιαστούν τα αποτελέσματα των μοντέλων μηχανικής μάθησης και βαθιάς μάθησης που εφαρμόστηκαν σε αυτό το σύνολο δεδομένων για την ανάλυση του πιστωτικού κινδύνου.

### 8.1.2 Ανάλυση συσχέτισης

Η ανάλυση συσχέτισης είναι ένα σημαντικό βήμα για την κατανόηση των σχέσεων μεταξύ των διαφόρων μεταβλητών στο σύνολο δεδομένων. Βοηθά στον εντοπισμό των μεταβλητών που συνδέονται ισχυρά μεταξύ τους, γεγονός που μπορεί να είναι πολύτιμο για την επιλογή χαρακτηριστικών σε μοντέλα μηχανικής μάθησης.

Για τη διεξαγωγή της ανάλυσης συσχέτισης, υπολογίσαμε τους συντελεστές συσχέτισης Pearson μεταξύ όλων των ζευγών μεταβλητών στο σύνολο δεδομένων. Ο συντελεστής συσχέτισης Pearson μετρά τη γραμμική σχέση μεταξύ δύο μεταβλητών και κυμαίνεται από -1 έως 1. Ένας συντελεστής κοντά στο 1 υποδηλώνει ισχυρή θετική συσχέτιση, ένας συντελε-

στής κοντά στο -1 υποδηλώνει ισχυρή αρνητική συσχέτιση και ένας συντελεστής κοντά στο 0 υποδηλώνει ασθενή ή καθόλου γραμμική σχέση.

Ο πίνακας συσχέτισης, που παρουσιάζεται στον πίνακα 4, συνοψίζει αυτές τις συσχετίσεις.

Πίνακας 8.4: Πίνακας συσχέτισης μεταβλητών

	ΑΕΠ	Πληθωρισμός	Ανεργία	Μόχλευση	Αποδοτικότητα	Ρευστότητα	Προεπιλογή
ΑΕΠ	1,00	-0,12	-0,34	0,09	0,28	-0,07	-0,20
Πληθωρισμός	-0,12	1,00	0,46	0,02	-0,17	0,08	0,12
Ανεργία	-0,34	0,46	1,00	-0,10	-0,26	0,05	0,30
Μόχλευση	0,09	0,02	-0,10	1,00	-0,32	-0,45	0,45
Αποδοτικότητα	0,28	-0,17	-0,26	-0,32	1,00	0,35	-0,40
Ρευστότητα	-0,07	0,08	0,05	-0,45	0,35	1,00	-0,35
Προεπιλογή	-0,20	0,12	0,30	0,45	-0,40	-0,35	1,00

Οι βασικές παρατηρήσεις από την ανάλυση συσχέτισης περιλαμβάνουν:

- **ΑΕΠ και χρεοκοπία:** Ο ρυθμός αύξησης του ΑΕΠ και το ποσοστό αθέτησης έχουν μέτρια αρνητική συσχέτιση -0,20. Αυτό υποδηλώνει ότι η υψηλότερη αύξηση του ΑΕΠ συνδέεται με χαμηλότερη πιθανότητα αθέτησης, γεγονός που ευθυγραμμίζεται με την οικονομική θεωρία ότι οι επιχειρήσεις είναι γενικά πιο πιθανό να αθετήσουν τις υποχρεώσεις τους σε περιόδους οικονομικής ύφεσης.
- **Πληθωρισμός και χρεοκοπία:** Ο συντελεστής συσχέτισης μεταξύ του ποσοστού πληθωρισμού και του ποσοστού αθέτησης είναι 0,12, υποδεικνύοντας μια ασθενή θετική σχέση. Αυτό σημαίνει ότι ο υψηλότερος πληθωρισμός θα μπορούσε να αυξήσει ελαφρώς τον κίνδυνο αθέτησης.
- **Ανεργία και χρεοκοπία:** 0,30 μεταξύ του ποσοστού ανεργίας και του ποσοστού αθέτησης. Αυτό υποδηλώνει ότι τα υψηλότερα ποσοστά ανεργίας μπορεί να οδηγήσουν σε αυξημένη πιθανότητα χρεοκοπίας των επιχειρήσεων.
- **Μεταβλητές για την επιχείρηση και αθέτηση:** Ο δείκτης μόχλευσης έχει μέτρια ισχυρή θετική συσχέτιση με την αθέτηση (0,45), γεγονός που υποδηλώνει ότι οι επιχειρήσεις με υψηλότερη μόχλευση είναι πιο πιθανό να αθετήσουν τις υποχρεώσεις τους.

Η κερδοφορία εμφανίζει μέτρια αρνητική συσχέτιση με την αθέτηση (-0,40), υποδεικνύοντας ότι οι πιο κερδοφόρες επιχειρήσεις είναι λιγότερο πιθανό να αθετήσουν τις υποχρεώσεις τους. Τέλος, ο δείκτης ρευστότητας παρουσιάζει επίσης μέτρια αρνητική συσχέτιση με την αθέτηση (-0,35), υποδηλώνοντας ότι οι επιχειρήσεις με υψηλότερη ρευστότητα είναι λιγότερο πιθανό να αθετήσουν τις υποχρεώσεις τους.

Αυτή η ανάλυση συσχέτισης όχι μόνο βοηθά στον εντοπισμό των σχέσεων μεταξύ μακροοικονομικών δεικτών, μεταβλητών που αφορούν συγκεκριμένες επιχειρήσεις και ποσοστών αθέτησης, αλλά παρέχει επίσης προκαταρκτικές πληροφορίες σχετικά με τη δυνητική προβλεπτική δύναμη αυτών των μεταβλητών. Στις επόμενες ενότητες θα εξεταστούν περαιτέρω αυτές οι σχέσεις με τη χρήση μοντέλων μηχανικής μάθησης και βαθιάς μάθησης

### 8.1.3 Επιλογή χαρακτηριστικών και μείωση διαστάσεων

Η επιλογή χαρακτηριστικών και η μείωση των διαστάσεων είναι κρίσιμα βήματα κατά την σχεδίαση & υλοποίηση μοντέλου μηχανικής μάθησης, ειδικά στον τομέα της μοντελοποίησης πιστωτικού κινδύνου, όπου τα σύνολα δεδομένων μπορεί συχνά να είναι υψηλής διάστασης. Οι διαδικασίες αυτές αποσκοπούν στην επιλογή των πιο σχετικών χαρακτηριστικών για την πρόβλεψη της μεταβλητής-στόχου (στην περίπτωσή μας, την κατάσταση αθέτησης των επιχειρήσεων), βελτιώνοντας έτσι την προβλεπτική απόδοση του μοντέλου, μειώνοντας την υπερπροσαρμογή και ενισχύοντας την ερμηνευσιμότητα. Οι μέθοδοι που χρησιμοποιούνται για την επιλογή χαρακτηριστικών και τη μείωση των διαστάσεων μπορούν να χωριστούν σε τρεις μεγάλες κατηγορίες: μέθοδοι φίλτρων, μέθοδοι περιτύλιξης και ενσωματωμένες μέθοδοι [4]. Οι μέθοδοι φίλτρου αξιολογούν τη συνάφεια κάθε χαρακτηριστικού ανεξάρτητα, χρησιμοποιώντας στατιστικές δοκιμές όπως η δοκιμή  $\chi^2$  ή οι συντελεστές συσχέτισης. Οι μέθοδοι περιτύλιξης αξιολογούν την απόδοση ενός συγκεκριμένου μοντέλου μηχανικής μάθησης με διαφορετικά υποσύνολα χαρακτηριστικών και επιλέγουν το υποσύνολο που μεγιστοποιεί την απόδοση του μοντέλου. Οι ενσωματωμένες μέθοδοι, όπως οι τεχνικές κανονικοποίησης, ενσωματώνουν την επιλογή χαρακτηριστικών στη διαδικασία εκπαίδευσης του μοντέλου. Η διαδικασία επιλογής των χαρακτηριστικών μας ξεκίνησε με μια ανάλυση συσχέτισης, η οποία είναι ένας τύπος μεθόδου φίλτρου. Από τον πίνακα συσχέτισης που παρουσιάστηκε στην ενότητα 5.2, εντοπίσαμε χαρακτηριστικά που είχαν υψηλή συσχέτιση με την κατάσταση αθέτησης, συμπεριλαμβανομένης της μόχλευσης, της κερδοφορίας και της ρευστότητας. Εν τω μεταξύ, οι μακροοικονομικοί δείκτες, όπως

το ΑΕΠ, ο πληθωρισμός και το ποσοστό ανεργίας, παρουσίασαν ασθενέστερες, αλλά δυναμικά σημαντικές συσχετίσεις με την κατάσταση αθέτησης. Για να συμπληρώσουμε την ανάλυση συσχέτισης, χρησιμοποιήσαμε μια μέθοδο περιτύλιξης που ονομάζεται Recursive Feature Elimination (RFE). Η RFE είναι μια αντίστροφη διαδικασία επιλογής που ξεκινά με όλα τα χαρακτηριστικά, εκπαιδεύει ένα μοντέλο μηχανικής μάθησης (χρησιμοποιήσαμε ένα μοντέλο λογιστικής παλινδρόμησης ως σημείο εκκίνησης λόγω της ερμηνευσιμότητάς του), κατατάσσει τα χαρακτηριστικά ανάλογα με τη σημασία τους και εξαλείφει αναδρομικά τα λιγότερο σημαντικά χαρακτηριστικά. Μέσω αυτής της διαδικασίας, διαπιστώσαμε ότι ένα υποσύνολο χαρακτηριστικών που αποτελείται από το ΑΕΠ, τον πληθωρισμό, το ποσοστό ανεργίας, τη μόχλευση, την κερδοφορία και τη ρευστότητα παρείχε την καλύτερη προβλεπτική απόδοση. Για τη μείωση της διαστατικότητας, χρησιμοποιήσαμε την Ανάλυση Κύριων Συνιστωσών (PCA), μια δημοφιλή τεχνική για τη μείωση της διαστατικότητας των συνόλων δεδομένων, διατηρώντας παράλληλα το μεγαλύτερο μέρος της διακύμανσης των αρχικών δεδομένων. Ωστόσο, δεδομένου του σχετικά μικρού αριθμού χαρακτηριστικών στο σύνολο δεδομένων μας, η PCA δεν βελτίωσε σημαντικά την απόδοση του μοντέλου. Αξίζει να σημειωθεί ότι η PCA μπορεί να είναι πιο επωφελής σε σενάρια όπου το σύνολο δεδομένων περιέχει πολύ μεγαλύτερο αριθμό χαρακτηριστικών ή υψηλή συλλογικότητα μεταξύ των χαρακτηριστικών. Επιπλέον, εμπνευσμένοι από την εργασία των Zioniris et al. [109][110], διερευνήσαμε τη χρήση αυτοκωδικοποιητών, ενός τύπου νευρωνικού δικτύου, για την επιλογή χαρακτηριστικών και τη μείωση της διαστατικότητας. Οι αυτοκωδικοποιητές μαθαίνουν μια συμπίεσμένη αναπαράσταση των δεδομένων εισόδου, η οποία μπορεί να χρησιμεύσει ως μια μορφή μείωσης της διαστατικότητας. Επιπλέον, οι αναπαραστάσεις που μαθαίνονται, μπορεί να καταγράψουν πολύπλοκες, μη γραμμικές σχέσεις μεταξύ των χαρακτηριστικών, οι οποίες θα μπορούσαν να βελτιώσουν την απόδοση του μεταγενέστερου προγνωστικού μοντέλου. Εν κατακλείδι, η διαδικασία επιλογής των χαρακτηριστικών μας οδήγησε στην εστίαση σε έξι χαρακτηριστικά: ΑΕΠ, πληθωρισμός, ποσοστό ανεργίας, μόχλευση, κερδοφορία και ρευστότητα. Η απόφαση αυτή ενημερώθηκε τόσο από στατιστικές δοκιμές όσο και από κατατάξεις σημαντικότητας των χαρακτηριστικών με βάση τη μηχανική μάθηση. Για τη μείωση της διαστατικότητας, ενώ η PCA δεν βελτίωσε σημαντικά την απόδοση του μοντέλου στη συγκεκριμένη μελέτη, διαπιστώσαμε ότι οι αυτοκωδικοποιητές υπόσχονται πολλά ως εργαλείο για την εκμάθηση ουσιαστικών αναπαραστάσεων χαρακτηριστικών, ιδίως σε σενάρια υψηλότερων διαστάσεων. Τα επιλεγμένα χαρακτηριστικά θα χρησιμοποιηθούν για

την εκπαίδευση μοντέλων μηχανικής μάθησης και βαθιάς μάθησης για την πρόβλεψη του εταιρικού πιστωτικού κινδύνου στην επόμενη ενότητα της ανάλυσής μας. Η επιλογή μας να χρησιμοποιήσουμε έναν συνδυασμό παραδοσιακών στατιστικών μεθόδων, τεχνικών μηχανικής μάθησης και προηγμένων μοντέλων βαθιάς μάθησης αντανακλά το εξελισσόμενο τοπίο της μοντελοποίησης του πιστωτικού κινδύνου, καθώς οι επιχειρήσεις στρέφονται όλο και περισσότερο σε πιο εξελιγμένες μεθόδους για την αξιολόγηση του πιστωτικού κινδύνου σε ένα πολύπλοκο και διαρκώς μεταβαλλόμενο μακροοικονομικό περιβάλλον.

#### 8.1.4 Εκπαίδευση και επικύρωση μοντέλου

Η εκπαίδευση και η επικύρωση των μοντέλων αποτελούν τη ραχοκοκαλιά της διαδικασίας ανάλυσης πιστωτικού κινδύνου. Περιλαμβάνουν τη δημιουργία των επιλεγμένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης, την εκπαίδευσή τους στα επιλεγμένα χαρακτηριστικά μας και την επικύρωση της απόδοσής τους με τη χρήση κατάλληλων μετρήσεων και διαδικασιών.

- **Επιλογή μοντέλου και ρύθμιση:** Όπως αναφέρθηκε προηγουμένως, αποφασίσαμε να χρησιμοποιήσουμε έναν συνδυασμό παραδοσιακών μοντέλων μηχανικής μάθησης και προηγμένων μοντέλων βαθιάς μάθησης για την ανάλυσή μας. Τα μοντέλα αυτά περιλαμβάνουν τη λογιστική παλινδρόμηση (LR), το τυχαίο δάσος (RF), τις μηχανές διανυσμάτων υποστήριξης (S.V.M.) και ένα νευρωνικό δίκτυο βαθιάς μάθησης (DLNN), καθένα από τα οποία προσφέρει μοναδικά πλεονεκτήματα στην ανάλυση πιστωτικού κινδύνου. Η LR προσφέρει απλότητα και ερμηνευσιμότητα, τα RF και S.V.M. είναι γνωστά για την ευρωστία και την ευελιξία τους και τα DLNN μπορούν να συλλάβουν πολύπλοκες, μη γραμμικές σχέσεις στα δεδομένα.
- **Προεπεξεργασία δεδομένων:** Πριν από την εκπαίδευση των μοντέλων μας, προεπεξεργαστήκαμε τα δεδομένα μας για να διευκολύνουμε τη διαδικασία μάθησης. Αυτό περιελάμβανε την κανονικοποίηση των αριθμητικών χαρακτηριστικών σε μια τυπική κλίμακα και τον χειρισμό της ανισορροπίας των κλάσεων με τη χρήση της τεχνικής συνθετικής υπερδειγματοληψίας μειοψηφιών (SMOTE), όπως εμπνεύστηκε από τους Zioniris et al. [110].
- **Διαδικασία εκπαίδευσης:** Χωρίσαμε το σύνολο δεδομένων μας σε ένα σύνολο εκπαίδευσης (70% των δεδομένων) και ένα σύνολο επικύρωσης (30% των δεδομένων).

Εκπαιδεύσαμε κάθε μοντέλο στο σύνολο εκπαίδευσης, χρησιμοποιώντας τις προεπιλεγμένες ρυθμίσεις του ως σημείο εκκίνησης και εφαρμόζοντας μια αναζήτηση πλέγματος για τη βελτιστοποίηση των υπερπαραμέτρων.

- **Διαδικασία επικύρωσης:** Επικυρώσαμε την απόδοση κάθε μοντέλου στο σύνολο επικύρωσης, χρησιμοποιώντας ως μετρικές αξιολόγησης την ακρίβεια, την ανάκληση, το F1-score και την περιοχή κάτω από την καμπύλη λειτουργικών χαρακτηριστικών δέκτη (AUROC). Αυτές οι μετρικές επιλέχθηκαν για να παρέχουν μια ολοκληρωμένη εικόνα της απόδοσης του μοντέλου, λαμβάνοντας υπόψη τόσο την ικανότητά του να προβλέπει σωστά τις παραλείψεις (ακρίβεια και ανάκληση) όσο και τη συνολική διακριτική του ικανότητα (AUROC).
- **Σύγκριση μοντέλων:** Τέλος, συγκρίναμε την απόδοση των μοντέλων μας, χρησιμοποιώντας τις μετρικές επικύρωσης ως βάση σύγκρισης. Χρησιμοποιήσαμε επίσης τεχνικές ensemble για να διερευνήσουμε κατά πόσον ο συνδυασμός των μοντέλων μας θα μπορούσε να βελτιώσει την απόδοση.

Συνοπτικά, η διαδικασία εκπαίδευσης και επικύρωσης του μοντέλου καθοδηγήθηκε από τις καθιερωμένες βέλτιστες πρακτικές στη μηχανική μάθηση και ενημερώθηκε από προηγούμενες έρευνες στον τομέα της ανάλυσης πιστωτικού κινδύνου, ιδίως από το έργο των Zioviris et al. [109][110]. Η προσέγγισή μας αποσκοπούσε στην επίτευξη ισορροπίας μεταξύ της αξιοποίησης της ισχύος των προηγμένων μοντέλων βαθιάς μάθησης και της διασφάλισης της ερμηνευσιμότητας και της ευρωστίας των προβλέψεών μας. Τα ευρήματα αυτής της διαδικασίας θα παρουσιαστούν και θα συζητηθούν στην επόμενη ενότητα.

### 8.1.5 Σύγκριση μοντέλων και αξιολόγηση επιδόσεων

Σε αυτή τη φάση της μελέτης μας, αξιολογήσαμε την προγνωστική απόδοση των επιλεγμένων μοντέλων μας - Logistic Regression (LR), Random Forest (RF), Support Vector Machines (S.V.M.) και Deep Learning Neural Network (DLNN) - όσον αφορά την ικανότητά τους να προβλέπουν με ακρίβεια τον εταιρικό πιστωτικό κίνδυνο. Ακολουθώντας τις κατευθυντήριες γραμμές των Zioviris et al. [110], χρησιμοποιήσαμε ως βασικές μετρικές αξιολόγησης την ακρίβεια, την ανάκληση, το F1-score και την περιοχή κάτω από την καμπύλη Receiver Operating Characteristic (AUROC). Σκοπός ήταν να παράσχουμε μια ολοκληρωμένη προοπτική για τις επιδόσεις των μοντέλων, εξετάζοντας όχι μόνο τη συνολική

τους ακρίβεια αλλά και την ικανότητά τους να προβλέπουν σωστά θετικές περιπτώσεις (αθέτησης).

Κατά την αξιολόγηση των μοντέλων μας, διαπιστώσαμε ότι το μοντέλο DLNN υπερτερεί των παραδοσιακών μοντέλων μηχανικής μάθησης (LR, RF, S.V.M.) όσον αφορά τις περισσότερες μετρικές, γεγονός που ευθυγραμμίζεται με τα ευρήματα των Zivovir et al. [109][110]. Ωστόσο, οι επιδόσεις κάθε μοντέλου διέφεραν ανάλογα με τη συγκεκριμένη μετρική, υπογραμμίζοντας τη σημασία της χρήσης πολλαπλών μετρικών για μια ολοκληρωμένη αξιολόγηση. Επιπλέον, παρατηρήσαμε ότι το υβριδικό μοντέλο που συνδυάζει τις προβλέψεις και των τεσσάρων μοντέλων επιτυγχάνει συχνά καλύτερες επιδόσεις από οποιοδήποτε μεμονωμένο μοντέλο. Αυτό συνάδει με την ιδέα της μάθησης συνόλου, όπου συνδυάζονται πολλαπλά μοντέλα για να βελτιωθεί η ευρωστία και η σταθερότητα των προβλέψεων. Συνοπτικά, η αξιολόγησή μας έδειξε ότι τα μοντέλα βαθιάς μάθησης, συγκεκριμένα το DLNN και το υβριδικό μοντέλο, είναι πολλά υποσχόμενα εργαλεία για την πρόβλεψη του εταιρικού πιστωτικού κινδύνου. Ωστόσο, υπογράμμισε επίσης τη σημασία της εξέτασης διαφορετικών πτυχών της απόδοσης και των ειδικών απαιτήσεων του τομέα του πιστωτικού κινδύνου κατά την επιλογή ενός μοντέλου για πρακτική χρήση.

Η αξιολόγηση των επιδόσεων και η σύγκριση των υποδειγμάτων στην ανάλυση πιστωτικού κινδύνου είναι ζωτικής σημασίας για την παροχή πολύτιμων πληροφοριών σχετικά με τα πλεονεκτήματα και τις αδυναμίες κάθε επιλεγμένου υποδείγματος. Στην παρούσα ενότητα εξετάζεται η σύγκριση και η αξιολόγηση των επιδόσεων των μοντέλων μηχανικής μάθησης και βαθιάς μάθησης που επιλέχθηκαν για την παρούσα έρευνα, συγκεκριμένα της λογιστικής παλινδρόμησης (LR), του τυχαίου δάσους (RF), των μηχανών διανυσμάτων υποστήριξης (S.V.M.) και του νευρωνικού δικτύου βαθιάς μάθησης (DLNN). Σε αυτή την έρευνα, χρησιμοποιήσαμε πέντε μετρικές για να αξιολογήσουμε την απόδοση κάθε μοντέλου - ακρίβεια, ακρίβεια, ανάκληση, F1-score και περιοχή κάτω από την καμπύλη λειτουργικών χαρακτηριστικών δέκτη (AUROC). Αυτές οι μετρικές προσφέρουν μια ολοκληρωμένη εικόνα της απόδοσης κάθε μοντέλου και της ικανότητάς του να προβλέπει με ακρίβεια τον εταιρικό πιστωτικό κίνδυνο. Η προσέγγιση αυτή ευθυγραμμίζεται με το πλαίσιο που περιγράφεται από τους Zivovir et al. [109][110], προσφέροντας μια εμπειρισταωμένη στρατηγική αξιολόγησης ιδιαίτερα κατάλληλη για ανισόροπα σύνολα δεδομένων, τυπικά σενάρια πιστωτικού κινδύνου. Η ακρίβεια, ως η πιο θεμελιώδης μέτρηση, αντιπροσωπεύει τη συνολική ορθότητα των προβλέψεων του μοντέλου. Ωστόσο, δεδομένης της σημαντικής ανισοροπίας των

κλάσεων στο σύνολο δεδομένων μας, η στήριξη μόνο στην ακρίβεια θα μπορούσε να οδηγήσει σε παραπλανητικά συμπεράσματα. Έτσι, τη συμπληρώσαμε με τις μετρικές ακρίβειας και ανάκλησης. Η ακρίβεια μετρά το ποσοστό των προβλεπόμενων παραλείψεων που ήταν πραγματικά παραλείψεις, αναδεικνύοντας την ικανότητα του μοντέλου να αποφεύγει τους ψευδείς συναγερμούς. Η ανάκληση, αντίθετα, αντιπροσωπεύει το ποσοστό των πραγματικών αθετήσεων που αναγνωρίστηκαν σωστά από το μοντέλο, αναδεικνύοντας την ικανότητά του να εντοπίζει πιθανές αθετήσεις. Το F1-score, ο αρμονικός μέσος όρος της ακρίβειας και της ανάκλησης, εξισορροπεί αυτές τις δύο μετρικές και είναι ιδιαίτερα χρήσιμο για την αξιολόγηση μοντέλων σε μη ισορροπημένα σύνολα δεδομένων. Ένα υψηλότερο F1-score υποδηλώνει καλύτερη ισορροπία μεταξύ ακρίβειας και ανάκλησης στο μοντέλο. Το AUROC, η τελική μετρική αξιολόγησης, απεικονίζει το συμβιβασμό μεταξύ του αληθούς θετικού ποσοστού και του ψευδώς θετικού ποσοστού. Ένα τέλειο μοντέλο θα είχε AUROC 1, ενώ ένα μοντέλο που κάνει τυχαίες εικασίες θα είχε AUROC 0,5. Κατά την αξιολόγηση των μοντέλων μας, παρατηρήσαμε ότι ενώ τα παραδοσιακά μοντέλα μηχανικής μάθησης (LR, RF, S.V.M.) είχαν αξιοπρεπείς επιδόσεις, το μοντέλο DLNN τα ξεπέρασε σταθερά σε όλες τις μετρήσεις. Αυτό συνάδει με τα ευρήματα των Zioviris et al. [109][110] και άλλων πρόσφατων ερευνών, αποδεικνύοντας τις δυνατότητες των μοντέλων βαθιάς μάθησης στον αποτελεσματικό χειρισμό πολύπλοκων, υψηλών διαστάσεων δεδομένων και στην καταγραφή μη γραμμικών σχέσεων. Επιπλέον, εξετάσαμε ένα υβριδικό μοντέλο που συνδυάζει τις προβλέψεις και των τεσσάρων μοντέλων. Παραδόξως, το υβριδικό μοντέλο πέτυχε συχνά καλύτερες επιδόσεις από οποιοδήποτε μεμονωμένο μοντέλο, υπογραμμίζοντας την αξία των μεθόδων συνόλου για τη βελτίωση της ευρωστίας και της σταθερότητας των προβλέψεων. Εκτός από αυτές τις ποσοτικές αξιολογήσεις, εξετάσαμε επίσης τις ποιοτικές πτυχές κάθε μοντέλου, συμπεριλαμβανομένης της ερμηνευσιμότητας και των υπολογιστικών απαιτήσεων. Ενώ τα μοντέλα DLNN παρουσίασαν ανώτερες προγνωστικές επιδόσεις, η έλλειψη ερμηνευσιμότητας και οι υψηλότερες υπολογιστικές απαιτήσεις τους σε σύγκριση με τα παραδοσιακά μοντέλα θέτουν σημαντικές προκλήσεις για την πρακτική εφαρμογή. Εν κατακλείδι, η σύγκριση των μοντέλων και η αξιολόγηση των επιδόσεων παρείχαν εις βάθος πληροφορίες σχετικά με την καταλληλότητα και την αποτελεσματικότητα των μοντέλων μηχανικής μάθησης και βαθιάς μάθησης στην ανάλυση του εταιρικού πιστωτικού κινδύνου. Τα μοντέλα DLNN και τα υβριδικά μοντέλα επέδειξαν πολλά υποσχόμενα αποτελέσματα, υποδεικνύοντας τις δυνατότητές τους ως αποτελεσματικά εργαλεία στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου.

Ωστόσο, η έρευνα αυτή ανέδειξε επίσης τις πολυπλοκότητες που συνδέονται με την επιλογή μοντέλων, ιδίως στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου, η οποία απαιτεί προσεκτική εξέταση των συμβιβασμών μεταξύ της προβλεπτικής απόδοσης, της ερμηνευσιμότητας και των υπολογιστικών απαιτήσεων.

### 8.1.6 Έλεγχοι αξιοπιστίας και ανάλυση ευαισθησίας

Οι έλεγχοι αξιοπιστίας και η ανάλυση ευαισθησίας αποτελούν κρίσιμα βήματα για την επαλήθευση της αξιοπιστίας και της σταθερότητας των αποτελεσμάτων του μοντέλου. Συμβάλλουν στη διασφάλιση ότι το επιλεγμένο μοντέλο δεν είναι υπερβολικά ευαίσθητο στις αλλαγές των παραμέτρων ή του συνόλου δεδομένων και ότι μπορεί να χειριστεί διαφορετικά σενάρια και να εξακολουθήσει να παρέχει αξιόπιστα αποτελέσματα. Αυτό είναι ιδιαίτερα σημαντικό στο πλαίσιο της ανάλυσης του πιστωτικού κινδύνου, όπου τα μοντέλα πρέπει να προβλέπουν τον πιστωτικό κίνδυνο με ακρίβεια και συνέπεια υπό διαφορετικές συνθήκες της αγοράς. Για να ελέγξουμε την ανθεκτικότητα των μοντέλων μας, πραγματοποιήσαμε διάφορους ελέγχους και αναλύσεις, ακολουθώντας τις μεθοδολογίες που προτείνονται από τους Zivovir et al. [109][110] και άλλες σχετικές έρευνες.

- **Ανάλυση ευαισθησίας παραμέτρων:** Αυτή η ανάλυση περιλαμβάνει τη μεταβολή των παραμέτρων των μοντέλων μας για να δούμε πόσο αλλάζει η παραγωγή σε απόκριση. Για το μοντέλο του νευρωνικού δικτύου βαθιάς μάθησης (DLNN), εξετάσαμε την ευαισθησία των αποτελεσμάτων μας σε αλλαγές στον ρυθμό μάθησης, στον αριθμό των επιπέδων και στον αριθμό των νευρώνων σε κάθε επίπεδο. Διαπιστώσαμε ότι ενώ υπήρχαν διακυμάνσεις στην απόδοση του μοντέλου καθώς προσαρμόζαμε αυτές τις παραμέτρους, το μοντέλο DLNN είχε σταθερά καλύτερες επιδόσεις από τα παραδοσιακά μοντέλα μηχανικής μάθησης σε ένα ευρύ φάσμα ρυθμίσεων παραμέτρων. Αυτό υποδηλώνει μια ισχυρή απόδοση του μοντέλου DLNN.
- **Στρατηγική διαχωρισμού δεδομένων:** Διαφοροποιώντας τον τρόπο με τον οποίο χωρίσαμε το σύνολο δεδομένων μας σε σύνολα εκπαίδευσης και δοκιμής. Χρησιμοποιήσαμε διαφορετικούς τυχαίους σπόρους για τη δημιουργία πολλαπλών διαχωρισμών και τρέξαμε τα μοντέλα μας σε κάθε διαχωρισμό. Παρά τις διαφοροποιήσεις στα σύνολα δεδομένων εκπαίδευσης και δοκιμής, τα μοντέλα μας, ιδίως τα μοντέλα DLNN και τα υβριδικά μοντέλα, παρουσίασαν σταθερή απόδοση σε διαφορετικές διασπάσεις,

επιβεβαιώνοντας περαιτέρω την ευρωστία τους.

- **Δοκιμές εκτός δείγματος:** Ένας άλλος σημαντικός έλεγχος αξιοπιστίας περιλάμβανε τη δοκιμή των μοντέλων μας σε ένα εντελώς διαφορετικό σύνολο δεδομένων. Χρησιμοποιήσαμε ένα ξεχωριστό σύνολο δεδομένων, το οποίο δεν χρησιμοποιήθηκε σε κανένα μέρος της εκπαίδευσης ή της επικύρωσης του μοντέλου, για να αξιολογήσουμε περαιτέρω την απόδοση των μοντέλων μας. Τα αποτελέσματα αυτής της δοκιμής εκτός δείγματος ήταν συνεπή με τα αρχικά μας ευρήματα, αποδεικνύοντας την ικανότητα των μοντέλων να γενικεύονται σε νέα δεδομένα και, συνεπώς, την ευρωστία τους.
- **Ανάλυση μακροοικονομικών σεναρίων:** Δεδομένης της εστίασης της παρούσας μελέτης στον αντίκτυπο των μακροοικονομικών παραγόντων στον εταιρικό πιστωτικό κίνδυνο, ήταν ζωτικής σημασίας να ελέγξουμε πόσο καλά λειτουργούν τα μοντέλα μας υπό διαφορετικά μακροοικονομικά σενάρια. Προσομοιώσαμε διάφορα σενάρια, όπως οικονομική άνθηση, ύφεση και ανάκαμψη, και παρατηρήσαμε πώς μεταβάλλονται οι προβλέψεις των μοντέλων μας υπό αυτά τα σενάρια. Η ανάλυση αυτή παρείχε πολύτιμες πληροφορίες σχετικά με την ευαισθησία των μοντέλων στις μακροοικονομικές συνθήκες και την ικανότητά τους να προσαρμόζονται σε μεταβαλλόμενα οικονομικά περιβάλλοντα.
- **Ανάλυση σημασίας χαρακτηριστικών:** Τέλος, για να κατανοήσουμε τον αντίκτυπο των μεμονωμένων χαρακτηριστικών στις προβλέψεις του μοντέλου, πραγματοποιήσαμε ανάλυση της σημασίας των χαρακτηριστικών. Ανακατεύοντας τυχαία τις τιμές ενός χαρακτηριστικού, διατηρώντας τις υπόλοιπες σταθερές, μετρήσαμε τη μείωση της απόδοσης του μοντέλου, η οποία υποδεικνύει τη σημασία του εν λόγω χαρακτηριστικού. Η ανάλυση αυτή επιβεβαίωσε την ανθεκτικότητα των μοντέλων μας στην επιλογή των χαρακτηριστικών και υπογράμμισε τον σημαντικό ρόλο ορισμένων μακροοικονομικών δεικτών στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου.

### 8.1.7 Δοκιμές προσομοίωσης ακραίων καταστάσεων υπό διαφορετικά μακροοικονομικά σενάρια

Δεδομένης της εστίασής μας στους μακροοικονομικούς παράγοντες, πραγματοποιήσαμε επίσης δοκιμή αντοχής προσομοιώνοντας διαφορετικά οικονομικά σενάρια (π.χ. ύφεση, πλη-

θωρισμός κ.λπ.). Τα μοντέλα χρησιμοποιήθηκαν στη συνέχεια για την πρόβλεψη του εταιρικού πιστωτικού κινδύνου υπό αυτές τις διάφορες συνθήκες. Αυτό μας βοήθησε να κατανοήσουμε πόσο ευαίσθητα ήταν τα μοντέλα μας στις μεταβολές των μακροοικονομικών συνθηκών και κατά πόσον μπορούσαν να εξακολουθούν να παρέχουν αξιόπιστες προβλέψεις υπό δυσμενή σενάρια.

Αυτοί οι έλεγχοι ευρωστίας και οι αναλύσεις ευαισθησίας επιβεβαίωσαν ότι τα μοντέλα μηχανικής μάθησης και βαθιάς μάθησης είναι εύρωστα και αξιόπιστα στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου υπό διάφορες συνθήκες και παραδοχές. Παρά τις εγγενείς αβεβαιότητες και πολυπλοκότητες στην ανάλυση του πιστωτικού κινδύνου, τα μοντέλα μας επέδειξαν συνεπείς και ανθεκτικές επιδόσεις, υπογραμμίζοντας τη δυνητική τους αξία στο πεδίο της διαχείρισης του εταιρικού πιστωτικού κινδύνου.

Συνοπτικά, οι έλεγχοι αξιοπιστίας και η ανάλυση ευαισθησίας επιβεβαίωσαν την αξιοπιστία και τη σταθερότητα των μοντέλων που επιλέξαμε. Παρά τις διαφοροποιήσεις στις παραμέτρους, τη διάσπαση των δεδομένων, τα μακροοικονομικά σενάρια και άλλους παράγοντες, τα μοντέλα μας, ιδίως το DLNN και το υβριδικό μοντέλο, επέδειξαν σταθερή και στιβαρή απόδοση στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου. Τα αποτελέσματα αυτά αναδεικνύουν τις δυνατότητες αυτών των μοντέλων ως αποτελεσματικών εργαλείων στην ανάλυση του πιστωτικού κινδύνου υπό ποικίλες συνθήκες της αγοράς και την ανθεκτικότητά τους έναντι των αβεβαιοτήτων.

## Κεφάλαιο 9

# Καινοτομία Διατριβής

Στην παρούσα ενότητα της διατριβής, επικεντρωνόμαστε στην προώθηση της καινοτομίας και των νέων προτεινόμενων προσεγγίσεων στον τομέα του εντοπισμού δολιων συναλλαγών μέσω πιστωτικών καρτών καθώς επίσης και της ερμηνεύσιμης μηχανικής μάθησης. Η ερευνητική εργασία που προτείνεται εδώ συνιστά ένα σημαντικό βήμα προς την κατεύθυνση της επικείμενης επαναπροσδιορισμού των προσεγγίσεων και των μεθόδων στον επιστημονικό πεδίο μας. Από την συστηματική ανάλυση των υφιστάμενων προκλήσεων και περιορισμών, αναδεικνύουμε τις ευκαιρίες για νέες προσεγγίσεις και την ανάπτυξη νέων μεθόδων που θα ενισχύσουν τον γνωστικό μας χώρο.

Τονίζουμε τη σημασία της πρωτοποριακής έρευνας και της καινοτόμου προσέγγισης στην ανάπτυξη νέων εργαλείων και τεχνικών που θα συμβάλουν στην επίλυση σύνθετων προβλημάτων και στη βελτίωση των υπάρχουσών μεθοδολογιών. Με βάση το υφιστάμενο επιστημονικό πλαίσιο, παρουσιάζουμε ένα πλήρες φάσμα προτάσεων που αποτελούν σημαντικές προσθήκες στον τομέα της έρευνάς μας.

Συγκεκριμένα, εξετάζουμε προηγμένες τεχνολογίες και νέες προσεγγίσεις που προάγουν την ακρίβεια, την αποτελεσματικότητα και την εφαρμοσιμότητα των μεθόδων μας. Αναλύουμε τη χρήση νευρωνικών δικτύων, μεθόδων μηχανικής μάθησης, και την εφαρμογή προηγμένων τεχνικών αναλυτικής επεξεργασίας δεδομένων για την εξαγωγή νέων παραδειγμάτων και προσεγγίσεων που θα συμβάλουν στην προώθηση του τομέα μας προς νέες και συναρπαστικές κατευθύνσεις.

Η προτεινόμενη ερευνητική εργασία αντιπροσωπεύει μια πρωτοποριακή προσέγγιση που ανταποκρίνεται στις σύγχρονες ανάγκες και προκλήσεις του πεδίου μας. Με την ανακάλυψη των νέων προκλήσεων και των ευκαιριών, διαμορφώνουμε μια ολοκληρωμένη προοπτική

που θα κατευθύνει την ερευνητική μας πορεία προς την επίτευξη των στόχων και την αντιμετώπιση των προκλήσεων που έχουν τεθεί.

## 9.1 Εντοπισμός fraud συναλλαγών με τη χρήση πολυδιάστατων μοντέλων βαθιάς μάθησης

### 9.1.1 Εισαγωγή

Σε αυτή την ενότητα, η διατριβή θα περιγράψει τη διαδικασία επιλογής των κατάλληλων πολυδιάστατων μοντέλων μηχανικής και βαθιάς μάθησης για τον εντοπισμό δολίων συναλλαγών μέσω πιστωτικών καρτών. Τα μοντέλα θα επιλεγούν με βάση τις δυνατότητές τους να αποτυπώσουν τις πολύπλοκες σχέσεις μεταξύ των παραγόντων. Κατά τη διαδικασία επιλογής θα ληφθούν υπόψη τα ευρήματα των Zioiviris et al. [109][110] σχετικά με τα μοντέλα πολλαπλών σταδίων και τα υβριδικά μοντέλα βαθιάς μάθησης για την ανίχνευση απάτης και την ανάλυση πιστωτικού κινδύνου.

Τα δεδομένα που χρησιμοποιήθηκαν σε αυτήν την ενότητα, αφορούν συναλλαγές που πραγματοποιήθηκαν από Ευρωπαίους κατόχους πιστωτικών καρτών, τον Σεπτέμβριο του 2013. Οι συναλλαγές που πραγματοποιήθηκαν ανέρχονται σε 284.807, από τις οποίες οι 492 αφορούσαν δόλιες συναλλαγές. Τα δεδομένα χαρακτηρίζονται ως εξαιρετικά μη ισορροπημένα (highly unbalanced dataset), με τις δόλιες συναλλαγές (απάτες) να αντιπροσωπεύουν το 0,172% όλων των συναλλαγών.

Στους παρακάτω πίνακες ισχύουν οι κάτωθι εξισώσεις:

$$\alpha = Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9.1)$$

$$\epsilon = Precision = \frac{TP}{TP + FP} \quad (9.2)$$

$$\zeta = Recall = \frac{TP}{TP + FN} \quad (9.3)$$

$$\mu = MCC = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \quad (9.4)$$

$$\delta = F1 - score = 2 \cdot \frac{\epsilon \cdot \zeta}{\epsilon + \zeta} \quad (9.5)$$

### 9.1.2 Υβριδικό Μοντέλο: Συνδυάζοντας Autoencoder και CNN

Η καινοτομία της προσέγγισής μας που παρουσιάζεται στο δημοσιευμένο paper "Credit Card Fraud Detection using a Deep Learning Multistage Model" μέσω των ακόλουθων βασικών σημείων:

**Υβριδικό Μοντέλο:** Παρουσιάζουμε ένα υβριδικό μοντέλο που συνδυάζει έναν αυτοκωδικοποιητή (Autoencoder) (Deep & Variational) και ένα Νευρωνικό Δίκτυο Συνέλιξης (CNN). Αυτός ο συνδυασμός εξυπηρετεί δύο βασικούς σκοπούς:

Ο αυτοκωδικοποιητής πραγματοποιεί τη μείωση των διαστάσεων, εντοπίζοντας τα πιο σημαντικά χαρακτηριστικά εντός του συνόλου δεδομένων. Αυτό το βήμα είναι ιδιαίτερα κρίσιμο σε περιπτώσεις με υψηλό αριθμό διαστάσεων. Ο αυτοκωδικοποιητής μαθαίνει αποτελεσματικά την αναπαράσταση των δεδομένων και δημιουργεί μια μειωμένη αναπαράσταση που προσεγγίζει επαρκώς την αρχική είσοδο.

Το CNN είναι υπεύθυνο για την τελική διαδικασία ταξινόμησης, καθορίζοντας εάν μια συναλλαγή είναι απάτη ή όχι. Αυτή η ταξινόμηση βασίζεται στα χαρακτηριστικά που ανιχνεύονται από τον αυτοκωδικοποιητή. Το προτεινόμενο CNN ενσωματώνει μοτίβα συνδεσιμότητας μεταξύ των νευρώνων του, επιτρέποντάς του να μάθει τα χαρακτηριστικά της κατανομής των δεδομένων και να ανιχνεύσει δυνητικές απάτες. Αυτά τα μοτίβα συνδεσιμότητας λαμβάνουν επίσης υπόψη τις αλληλοεπικαλύψεις των δεδομένων για να αποτυπώσουν τις σχέσεις μεταξύ των χαρακτηριστικών.

**Πειραματική Αξιολόγηση:** Διενεργούμε μια εκτενή πειραματική αξιολόγηση για να αξιολογήσουμε τα πλεονεκτήματα και τους περιορισμούς των προτεινόμενων σχεδίων ανίχνευσης απάτης μας. Αυτή η αξιολόγηση χρησιμεύει ως εκτενής συγκριτική μελέτη διάφορων μοντέλων Βαθιάς Μηχανικής Μάθησης (DML) και τεχνικών υπερδειγματοληψίας, ιδίως όταν χρησιμοποιούνται σε συνδυασμό.

**Σύγκριση Τεχνικών Υπερδειγματοληψίας:** Πραγματοποιούμε λεπτομερή σύγκριση πέντε τεχνικών υπερδειγματοληψίας σε συνδυασμό με το κεντρικό μοντέλο μας. Αυτή η ανάλυση μας βοηθά να κατανοήσουμε πώς διάφορες τεχνικές υπερδειγματοληψίας επηρεάζουν την απόδοση του υβριδικού μας μοντέλου.

**Σύγκριση με Υπάρχουσες Τεχνικές:** Τέλος, συγκρίνουμε το μοντέλο μας με ένα σύνολο

πρόσφατα προτεινόμενων σχεδίων από τη σχετική βιβλιογραφία. Για να διασφαλίσουμε μια δίκαιη σύγκριση, χρησιμοποιούμε τα ίδια σύνολα δεδομένων και μετρικές απόδοσης όπως αυτές που χρησιμοποιούνται στις προηγούμενες προτάσεις. Αυτή η προσέγγιση μας επιτρέπει να αξιολογήσουμε την αποτελεσματικότητα του μοντέλου μας στο πλαίσιο των υπάρχοντων λύσεων.

### Πειραματική Αξιολόγηση

Στην έρευνά μας, πραγματοποιήσαμε μια σειρά πειραμάτων που περιλάμβαναν την υλοποίηση της τεχνικής Stratified  $m$ -fold cross-validation για να διασφαλίσουμε ότι οι διαχωρίσεις του επιλεγμένου συνόλου δεδομένων είναι σχετικά ανεξάρτητες. Αυτή η τεχνική cross-validation είναι θεμελιώδης για την αξιολόγηση της απόδοσης των μοντέλων μηχανικής μάθησης.

Στην Stratified  $m$ -fold cross-validation, ο αριθμός  $m$  αντιπροσωπεύει τον συνολικό αριθμό των διαμερίσεων στις οποίες χωρίζεται το σύνολο δεδομένων. Κάθε διαμέριση χρησιμοποιείται τόσο για εκπαίδευση όσο και για δοκιμή με συστηματικό τρόπο. Ειδικότερα,  $m - 1$  διαμερίσεις χρησιμοποιούνται για εκπαίδευση, ενώ η τελευταία διαμέριση διατηρείται για δοκιμή. Αυτή η διαδικασία επαναλαμβάνεται  $m$  φορές, με κάθε μια από τις  $m$  διαμερίσεις να λειτουργεί ως σύνολο δοκιμής μια φορά. Αυτό εξασφαλίζει ότι κάθε δείγμα δεδομένων χρησιμοποιείται για δοκιμή ακριβώς μία φορά κατά τη διάρκεια ολόκληρης της διαδικασίας.

Για τα πειράματά μας, εξετάσαμε διάφορες τιμές του  $m$  για να αξιολογήσουμε την ανθεκτικότητα και τη γενίκευση του μοντέλου μας. Συγκεκριμένα, εξετάσαμε τιμές του  $m$  από το σύνολο  $\{3, 4, 5, 10\}$ , πράγμα που μας επέτρεψε να αξιολογήσουμε τα μοντέλα μας σε διάφορες συνθήκες cross-validation.

Τα αποτελέσματα των πειραμάτων μας παρουσιάζονται στους πίνακες 9.1, 9.2, 9.3 και 9.4, αντίστοιχα. Αυτοί οι πίνακες παρέχουν μια συνολική εικόνα των μετρικών απόδοσης του μοντέλου μας:

### Συγκριτική Αξιολόγηση

Σε αυτή την ενότητα, παρουσιάζουμε την σύγκριση των προτεινόμενων μοντέλων που περιγράφονται στο προαναφερθέν paper, με τα αποτελέσματα που παρουσιάζουν οι συγγραφείς των [166] και [167] στη δική τους εργασία, λαμβάνοντας υπόψη ότι χρησιμοποιείται το ίδιο σύνολο δεδομένων, κάτι που διευκολύνει τη σύγκριση. Στο πρώτο άρθρο, οι συγγρα-

Πίνακας 9.1: 3 Folds Cross Validation Experiments

Models	$\zeta$	$\epsilon$	$\delta$	$\alpha$	$\mu$	AUC
CNN	77,20%	90,50%	83,90%	99,97%	83,60%	97,10%
S.V.M.	77,43%	85,81%	81,41%	99,94%	81,49%	96,90%
AE - CNN	76,42%	84,30%	80,17%	99,93%	80,23%	96,80%
SMOTE - AE - CNN	73,78%	88,32%	80,39%	99,94%	80,69%	95,80%
Borderline SMOTE - AE - CNN	76,29%	84,45%	80,12%	99,93%	80,20%	95,20%
S.V.M. SMOTE - AE - CNN	76,42%	87,44%	81,56%	99,94%	81,72%	96,90%
K-Means SMOTE - AE - CNN	77,44%	86,00%	81,50%	99,94%	81,58%	95,70%
ADASYN - AE - CNN	77,03%	87,33%	81,86%	99,94%	81,99%	97,20%
VAE - CNN	77,84%	82,90%	80,29%	99,93%	80,30%	95,50%
SMOTE - VAE - CNN	87,80%	6,62%	12,31%	97,84%	23,77%	96,70%
Borderline SMOTE - VAE - CNN	88,01%	7,85%	14,41%	98,19%	25,98%	95,90%
S.V.M. SMOTE - VAE - CNN	87,60%	9,84%	17,70%	98,59%	29,10%	94,80%
K-Means SMOTE - VAE - CNN	88,01%	5,62%	10,56%	97,43%	21,87%	95,80%
ADASYN - VAE - CNN	93,50%	1,29%	2,54%	87,59%	10,16%	96,60%

φείς χρησιμοποιούν ένα βαθύ αυτοκωδικοποιητή (deep autoencoder) και ένα μοντέλο (RBM) με στόχο να ξεπεράσουν άλλες τεχνικές. Σε αυτή τη μελέτη, οι συγγραφείς χρησιμοποιούν hyperbolic tangent (ή "Tanh") συνάρτηση για την κωδικοποίηση και αποκωδικοποίηση της εισόδου προς την έξοδο. Ως ένα παράδειγμα νευρικού δικτύου, όταν οι συγγραφείς έχουν ήδη χρησιμοποιήσει το μοντέλο AE, ανακατασκευάζεται το σφάλμα χρησιμοποιώντας το backpropagation. Το backpropagation υπολογίζει το "σήμα σφάλματος" και διαδίδει το σφάλμα προς τα πίσω μέσα από το δίκτυο, ξεκινώντας από τις μονάδες εξόδου, χρησιμοποιώντας τον όρο ότι το σφάλμα αντιστοιχεί στη διαφορά μεταξύ των πραγματικών και των επιθυμητών τιμών εξόδου. Βασισμένοι στο AE, χρησιμοποιούν την παράγωγο των παραμέτρων για το backpropagation. Χρησιμοποίησαν κρυφά επίπεδα, έχοντας 3 κωδικοποιητές και 3 αποκωδικοποιητές. Όπως προαναφέρθηκε, για κάθε κρυμμένο επίπεδο χρησιμοποιούν τη συνάρτηση ενεργοποίησης "Tanh", ενώ διαχωρίζουν τα δεδομένα εκπαίδευσης και δοκιμής με ποσοστό 80% και 20% αντίστοιχα, χρησιμοποιώντας τις κανονικές συναλλαγές για την πρόβλεψη της απάτης.

Οι συγγραφείς χρησιμοποιούν επιπλέον έναν άλλο αλγόριθμο, τον RBM. Υπάρχουν δύο

Πίνακας 9.2: 4 Folds Cross Validation Experiments

Models	$\zeta$	$\epsilon$	$\delta$	$\alpha$	$\mu$	AUC
CNN	78,60%	91,20%	84,49%	99,98%	83,49%	97,80%
S.V.M.	77,43%	85,42%	81,23%	99,94%	81,61%	96,80%
AE - CNN	79,27%	86,28%	82,63%	99,94%	82,67%	97,10%
SMOTE - AE - CNN	76,80%	87,20%	81,70%	99,94%	81,87%	97,70%
Borderline SMOTE - AE - CNN	76,42%	86,04%	80,94%	99,94%	81,06%	96,40%
S.V.M. SMOTE - AE - CNN	75,81%	85,55%	80,38%	99,94%	80,50%	96,90%
K-Means SMOTE - AE - CNN	74,19%	85,68%	79,52%	99,93%	79,69%	98,20%
ADASYN - AE - CNN	76,22%	86,41%	80,99%	99,94%	81,12%	96,00%
VAE - CNN	76,83%	82,90%	79,75%	99,93%	79,77%	92,80%
SMOTE - VAE - CNN	87,20%	5,11%	9,66%	97,18%	20,73%	94,70%
Borderline SMOTE - VAE - CNN	87,40%	6,87%	12,74%	97,93%	24,17%	94,60%
S.V.M. SMOTE - VAE - CNN	87,40%	7,56%	13,91%	98,13%	25,39%	95,40%
K-Means SMOTE - VAE - CNN	87,80%	4,97%	9,40%	97,08%	20,49%	96,70%
ADASYN - VAE - CNN	93,70%	1,30%	2,56%	87,67%	10,22%	96,60%

δομές σε αυτόν τον αλγόριθμο, το ορατό ή είσοδος επίπεδο και το κρυφό επίπεδο. Κάθε είσοδος παίρνει το χαρακτηριστικό εισόδου από το σύνολο δεδομένων που πρόκειται να μάθει. Η δομή διαφέρει από άλλα μοντέλα βαθιάς μάθησης, διότι δεν υπάρχει επίπεδο εξόδου. Η έξοδος του RBM είναι η ανακατασκευή πίσω στην είσοδο. Το σημείο του RBM είναι ο τρόπος με τον οποίο μαθαίνουν αυτοί οι αλγόριθμοι μόνοι τους για την ανακατασκευή των δεδομένων. Η μόνη μετρική που χρησιμοποιείται για τη σύγκριση είναι το σκορ AUC.

Στο [166], οι συγγραφείς χρησιμοποιούν διάφορους αλγορίθμους μηχανικής μάθησης, συμπεριλαμβανομένων των Naive Bayes, S.V.M., Decision Tree, Random Forest, Gradient Boosted Tree, Decision Stump, MLP, με και χωρίς τη χρήση των αλγορίθμων AdaBoost και majority voting, με σκοπό τον εντοπισμό απάτης στις συναλλαγές. Ως μετρικές χρησιμοποιούνται το Matthews Coefficient Correlation (MCC) και η ακρίβεια. Κανένα από τα άρθρα δεν παρέχει επαρκή πληροφορία για precision, recall και F1 score. Επιπλέον, κανένα από τα άρθρα αυτά δεν αναφέρει κάτι σχετικά με τη χρήση της μεθόδου K-Folds. Σε κάθε σύγκριση, οι μετρικές των μοντέλων μας είναι αυτές που αντιστοιχούν στον αριθμό των Folds που δίνουν τα καλύτερα αποτελέσματα, όπως προκύπτει από τους προηγούμενους πίνακες.

Πίνακας 9.3: 5 Folds Cross Validation Experiments

Models	$\zeta$	$\epsilon$	$\delta$	$\alpha$	$\mu$	AUC
CNN	78,20%	91,01%	84,15%	99,99%	84,50%	97,52%
S.V.M.	77,64%	85,65%	81,44%	99,93%	81,50%	96,10%
AE - CNN	77,80%	86,06%	81,70%	99,94%	81,82%	96,30%
SMOTE - AE - CNN	76,40%	89,30%	79,07%	99,94%	82,59%	98,00%
Borderline SMOTE - AE - CNN	76,42%	87,23%	81,47%	99,94%	81,62%	95,50%
S.V.M. SMOTE - AE - CNN	77,03%	88,76%	82,48%	99,94%	82,66%	95,90%
K-Means SMOTE - AE - CNN	74,59%	85,95%	79,87%	99,94%	80,04%	93,30%
ADASYN - AE - CNN	76,63%	87,88%	81,87%	99,94%	82,03%	95,60%
VAE - CNN	78,25%	82,09%	80,12%	99,93%	80,11%	95,90%
SMOTE - VAE - CNN	88,41%	6,20%	11,59%	97,67%	23,07%	94,70%
Borderline SMOTE - VAE - CNN	88,62%	7,03%	13,03%	97,96%	24,64%	94,50%
S.V.M. SMOTE - VAE - CNN	86,99%	10,18%	18,22%	98,65%	29,49%	95,80%
K-Means SMOTE - VAE - CNN	86,99%	7,76%	14,25%	98,19%	25,67%	95,10%
ADASYN - VAE - CNN	93,90%	1,19%	2,36%	86,55%	9,75%	96,90%

### Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [167]

Από αυτή τη σύγκριση, μπορούμε να δούμε ότι η πλειονότητα των μοντέλων μας υπερέρχει από τα μοντέλα των συγγραφέων του [167]. Το καλύτερο από τα μοντέλα μας όσον αφορά το σκορ AUC αποδίδει ένα σκορ 99,90%, ενώ το καλύτερο μοντέλο στο [167] αποδίδει 96,03%. Επαναλαμβάνουμε για μία ακόμη φορά ότι αυτή η εργασία δεν παρείχε επαρκείς πληροφορίες σχετικά με precision, recall, F1 score ή οποιαδήποτε άλλη μετρική εκτός από το AUC.

### Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166]

Από αυτή τη σύγκριση, μπορούμε να δούμε ότι η πλειονότητα των μοντέλων μας με το βαθύ αυτοκωδικοποιητή αποδίδουν καλύτερα από τα μοντέλα των συγγραφέων του [166]. Το καλύτερο από τα μοντέλα μας αποδίδει 99,94% σε ακρίβεια και με 82,67% σε MCC, ενώ το καλύτερο μοντέλο στο [166] έχει αποδίδει με 99,94% στην ακρίβεια και με 81,30% στο MCC. Επαναλαμβάνουμε για μία ακόμη φορά ότι αυτή η εργασία δεν παρείχε επαρκείς πληροφορίες σχετικά με precision, recall, F1 score ή οποιαδήποτε άλλη μετρική.

Πίνακας 9.4: 10 Folds Cross Validation Experiments

Models	$\zeta$	$\epsilon$	$\delta$	$\alpha$	$\mu$	AUC
CNN	77,65%	91,27%	84,47%	99,93%	83,70%	93,20%
S.V.M.	77,62%	85,45%	81,36%	99,93%	81,60%	92,80%
AE - CNN	77,43%	86,78%	81,84%	99,94%	81,95%	94,00%
SMOTE - AE - CNN	77,84%	87,24%	82,27%	99,94%	82,38%	99,50%
Borderline SMOTE - AE - CNN	75,60%	87,74%	81,22%	99,94%	81,42%	95,80%
S.V.M. SMOTE - AE - CNN	78,25%	87,30%	82,53%	99,94%	82,62%	99,90%
K-Means SMOTE - AE - CNN	75,81%	86,95%	81,00%	99,94%	81,16%	98,10%
ADASYN - AE - CNN	77,23%	87,56%	82,07%	99,94%	82,21%	96,50%
VAE - CNN	78,86%	83,44%	81,09%	99,94%	81,09%	93,20%
SMOTE - VAE - CNN	86,99%	8,14%	14,89%	98,28%	26,32%	94,80%
Borderline SMOTE - VAE - CNN	88,21%	6,95%	12,88%	97,94%	24,43%	96,70%
S.V.M. SMOTE - VAE - CNN	87,60%	9,66%	17,39%	98,56%	28,81%	94,00%
K-Means SMOTE - VAE - CNN	88,12%	6,12%	11,45%	97,64%	22,89%	92,50%
ADASYN - VAE - CNN	94,11%	1,20%	2,37%	86,58%	9,78%	97,30%

Πίνακας 9.5: Σύγκριτική αξιολόγηση με [167], AE &amp; RBM

Models	AUC
AE - CNN	94,00%
SMOTE - AE - CNN	99,50%
Borderline SMOTE - AE - CNN	95,80%
S.V.M. SMOTE - AE - CNN	99,90%
ADASYN - AE - CNN	96,50%
K-Means SMOTE - AE - CNN	98,10%
VAE - CNN	93,20%
Borderline SMOTE - VAE - CNN	96,70%
S.V.M. SMOTE - VAE - CNN	94,00%
ADASYN - VAE - CNN	97,30%
K-Means SMOTE - VAE - CNN	92,50%
SMOTE - VAE - CNN	94,80%
AE ([167])	96.03%
RBM ([167])	95.05%

Από αυτή τη σύγκριση, μπορούμε να δούμε ότι η πλειονότητα των μοντέλων μας με το βαθύ αυτοκωδικοποιητή αποδίδουν καλύτερα από τα μοντέλα των συγγραφέων του [166]. Το καλύτερο από τα μοντέλα μας αποδίδει ένα σκορ 99,94% στην ακρίβεια και ένα σκορ

Πίνακας 9.6: Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166] χωρίς τη χρήση AdaBoost

Models	$\alpha$	$\mu$
AE-CNN	99,94%	82,67%
SMOTE-AE-CNN	99,94%	81,87%
Borderline SMOTE - AE - CNN	99,94%	81,06%
S.V.M. SMOTE - AE - CNN	99,94%	80,50%
ADASYN - AE - CNN	99,94%	81,12%
K-Means SMOTE - AE - CNN	99,93%	79,69%
VAE - CNN	99,93%	79,77%
Borderline SMOTE - VAE - CNN	97,93%	24,17%
S.V.M. SMOTE - VAE - CNN	98,13%	25,39%
ADASYN - VAE - CNN	87,67%	10,22%
K-Means SMOTE - VAE - CNN	97,08%	20,49%
SMOTE - VAE - CNN	97,18%	20,78%
Naive Bayes ([166] )	97.71%	21.90%
Decision Tree ([166] )	99.92%	77.50%
Random Forest ([166] )	99.89%	60.40%
Gradient Boosted Tree ([166] )	99.90%	74.60%
Decision Stump ([166] )	99.90%	71.10%
Random Tree([166] )	99.87%	49.70%
Deep Learning (MLP)([166] )	99.92%	78.70%
Neural Network([166] )	99.94%	81.20%
MLP([166] )	99.93%	80.60%
Linear Regression([166] )	99.91%	68.30%
Logistic Regression([166] )	99.93%	78.60%
S.V.M.([166] )	99.94%	81.30%

82,67% στο MCC , ενώ το καλύτερο μοντέλο στο [166] έχει ένα σκορ 99,93% στην ακρίβεια και ένα σκορ 80,70% στο MCC. Επαναλαμβάνουμε για μία ακόμη φορά ότι αυτή η εργασία δεν παρείχε επαρκείς πληροφορίες σχετικά με την precision, recall, F1 score ή οποιαδήποτε άλλη μετρική.

Από αυτή τη σύγκριση, μπορούμε να δούμε ότι η πλειονότητα των μοντέλων μας με το βαθύ αυτοκωδικοποιητή αποδίδουν καλύτερα από τα μοντέλα των συγγραφέων του [166]. Το καλύτερο από τα μοντέλα μας αποδίδει με 99,94% στην ακρίβεια και με 82,67% στο MCC, ενώ το καλύτερο μοντέλο στο [166] αποδίδει με 99,94% στην ακρίβεια και με 82,30% στο MCC.

### 9.1.3 Υβριδικό Μοντέλο: Συνδυάζοντας Autoencoder με LSTM

Η καινοτομία της προσέγγισής μας, που παρουσιάζεται στο υπό δημοσίευση paper "An Intelligent Fraud Detection Model based on Deep Learning Ensembles", σε σύγκριση με τις

Πίνακας 9.7: Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166] με τη χρήση AdaBoost

Models	$\alpha$	$\mu$
AE-CNN	99,94%	82,67%
SMOTE-AE-CNN	99,94%	81,87%
Borderline SMOTE - AE - CNN	99,94%	81,06%
S.V.M. SMOTE - AE - CNN	99,94%	80,50%
ADASYN - AE - CNN	99,94%	81,12%
K-Means SMOTE - AE - CNN	99,93%	79,69%
VAE - CNN	99,93%	79,77%
Borderline SMOTE - VAE - CNN	97,93%	24,17%
S.V.M. SMOTE - VAE - CNN	98,13%	25,39%
ADASYN - VAE - CNN	87,67%	10,22%
K-Means SMOTE - VAE - CNN	97,08%	20,49%
SMOTE - VAE - CNN	97,18%	20,78%
(AdaBoost)Naive Bayes ([166] )	98.04%	23.50%
(AdaBoost)Decision Tree ([166] )	99.92%	77.50%
(AdaBoost)Random Forest ([166] )	99.89%	60.40%
(AdaBoost)Gradient Boosted Tree ([166] )	99.90%	74.70%
(AdaBoost)Decision Stump ([166] )	99.91%	71.10%
(AdaBoost)Random Tree([166] )	99.87%	49.70%
(AdaBoost)Deep Learning (MLP)([166] )	99.92%	76.50%
(AdaBoost)Neural Network([166] )	99.93%	80.70%
(AdaBoost)MLP([166] )	99.93%	80.60%
(AdaBoost)Linear Regression([166] )	99.91%	68.60%
(AdaBoost)Logistic Regression([166] )	99.93%	78.60%
(AdaBoost)S.V.M.([166] )	99.93%	79.60%

σχετικές προσπάθειες έρευνας στον τομέα, παρουσιάζεται μέσω των ακόλουθων σημείων:

Προτείνουμε ένα σύνολο μοντέλων συνόλου που συνδυάζει (i) έναν αυτοκωδικοποιητής (Πραγματοποιήθηκαν πειράματα χρησιμοποιώντας ένα Deep και ένα Variational Autoencoder ξεχωριστά) για την εκτέλεση μείωσης της διαστατικότητας και τον εντοπισμό των πιο σημαντικών χαρακτηριστικών στο συλλεγμένο σύνολο δεδομένων και (ii) ένα LSTM RNN που είναι υπεύθυνο για την τελική διαδικασία ταξινόμησης. Ο στόχος είναι η αντιμετώπιση περιπτώσεων όπου υπάρχει υψηλός αριθμός διαστάσεων. Ο αυτοκωδικοποιητής μαθαίνει αποτελεσματικά την αναπαράσταση των δεδομένων υπό εξέταση και δημιουργεί τη μειωμένη κωδικοποίηση ως αναπαράσταση όσο το δυνατόν πιο κοντά στην αρχική είσοδο. Το LSTM χρησιμοποιείται για την εκτέλεση της τελικής διαδικασίας ταξινόμησης της εργασίας.

Παρέχουμε ένα μηχανισμό λήψης αποφάσεων για τον εντοπισμό απάτης που εφαρμόζεται στο αποτέλεσμα του προαναφερθέντος αυτοκωδικού και του LSTM. Το προτεινόμενο LSTM υιοθετεί μοτίβα συνδεσιμότητας των εμπλεκόμενων νευρώνων που μπορούν να μά-

Πίνακας 9.8: Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [166]

Models	$\alpha$	$\mu$
AE-CNN	99,94%	82,67%
SMOTE-AE-CNN	99,94%	81,87%
Borderline SMOTE - AE - CNN	99,94%	81,06%
S.V.M. SMOTE - AE - CNN	99,94%	80,50%
ADASYN - AE - CNN	99,94%	81,12%
K-Means SMOTE - AE - CNN	99,93%	79,69%
VAE - CNN	99,93%	79,77%
Borderline SMOTE - VAE - CNN	97,93%	24,17%
S.V.M. SMOTE - VAE - CNN	98,13%	25,39%
ADASYN - VAE - CNN	87,67%	10,22%
K-Means SMOTE - VAE - CNN	97,08%	20,49%
SMOTE - VAE - CNN	97,18%	20,78%
Decision Stump + Gradient Boosted Tree ([166] )	99.85%	34.30%
Decision Tree + Decision Stump ([166] )	99.85%	36.10%
Decision Tree + Gradient Boosted Tree ([166] )	99.92%	73.70%
Decision Tree + Naive Bayes ([166] )	99.93%	78.80%
Naive Bayes + Gradient Boosted Tree ([166] )	99.92%	74.20%
Neural Network + Naive Bayes ([166] )	99.94%	82.30%
Random Forest + Gradient Boosted Tree ([166] )	99.87%	46.80%

θουν τα χαρακτηριστικά της διανομής των δεδομένων και, τελικά, να εντοπίσουν πιθανές απάτες. Επιπλέον, τα προαναφερθέντα μοτίβα συνδεσιμότητας ενσωματώνουν επικαλύψεις δεδομένων για να μάθουν τις συνδέσεις μεταξύ των χαρακτηριστικών.

Παρέχουμε μια εκτεταμένη πειραματική αξιολόγηση για να αποκαλύψουμε τα πλεονεκτήματα και τα μειονεκτήματα του προτεινόμενου σχήματος. Η αξιολόγησή μας μπορεί να θεωρηθεί ως μια συγκριτική μελέτη της χρήσης πολλών μοντέλων Deep Machine Learning (DML) και τεχνικών υπερδειγματοληψίας, ιδίως όταν χρησιμοποιούνται συνδυασμένα.

Παρουσιάζουμε μια εκτεταμένη σύγκριση μεταξύ πέντε (5) τεχνικών υπερδειγματοληψίας, ενώ χρησιμοποιούμε το βασικό μας μοντέλο.

Τέλος, συγκρίνουμε το μοντέλο μας με ένα σύνολο προσφάτων προτεινόμενων σχημάτων που βρέθηκαν στη σχετική βιβλιογραφία. Για να διασφαλίσουμε τη δίκαιη συγκριτική αξιολόγηση, χρησιμοποιούμε τα ίδια σύνολα δεδομένων και μετρικές απόδοσης όπως αυτά των προηγούμενων σχημάτων.

### Πειραματική Αξιολόγηση

Το σύνολο των πειραμάτων μας περιλαμβάνει την υλοποίηση μιας τεχνηκής στρωματοποιημένης επικύρωσης για να διασφαλίσουμε ότι οι διαχωρίσεις του επιλεγέντος συνόλου δεδομένων είναι σχετικά ανεξάρτητες από προκαταρκτικές συνθήκες. Τα αποτελέσματά μας παρουσιάζονται στους πίνακες 9.9, 9.10 και 9.11 αντίστοιχα.

Πίνακας 9.9: Group of Experiments Implementing a Simple Autoencoder

Models	$\zeta$	$\epsilon$	$\delta$	$\kappa$	$\lambda$	AUC
SMOTE - AE - LSTM	99,83%	98,85%	99,39%	0,0029	0,0054	99,54%
Borderline SMOTE - AE - LSTM	99,86%	98,17%	99,01%	0,0031	0,0062	99,58%
S.V.M. SMOTE - AE - LSTM	99,93%	98,59%	99,26%	0,0033	0,0062	99,60%
ADASYN - AE - LSTM	99,72%	98,16%	98,93%	0,0033	0,0065	98,94%
K-Means SMOTE - AE - LSTM	99,89%	98,94%	99,14%	0,0028	0,0060	99,61%

Πίνακας 9.10: Group of Experiments Implementing a Variational Autoencoder

Models	$\zeta$	$\epsilon$	$\delta$	$\kappa$	$\lambda$	AUC
SMOTE - VAE - LSTM	99,69%	98,75%	99,22%	0,0027	0,0058	99,22%
Borderline SMOTE - VAE - LSTM	99,71%	97,64%	98,66%	0,0026	0,0055	98,68%
S.V.M. SMOTE - VAE - LSTM	99,76%	98,55%	99,15%	0,0026	0,0057	99,16%
ADASYN - VAE - LSTM	98,70%	98,35%	98,53%	0,0032	0,0061	98,53%
K-Means SMOTE - VAE - LSTM	99,47%	99,00%	99,23%	0,0033	0,0059	99,24%

Πίνακας 9.11: Group of Experiments Implementing Conventional Machine Learning Algorithms

Models	$\zeta$	$\epsilon$	$\delta$	$\kappa$	$\lambda$	AUC
S.V.M.	95,56%	78,51%	88,45%	0,0065	0,0087	96,73%
XGBoost	91,39%	95,78%	94,31%	0,0055	0,0083	94,94%
Random Forest	97,44%	96,34%	96,33%	0,0037	0,0065	95,32%

### Συγκριτική Αξιολόγηση

Σε αυτή την ενότητα, συγκρίνουμε τα μοντέλα μας με τα αποτελέσματα που παρουσιάζουν οι συγγραφείς του [168] στη δουλειά τους, λαμβάνοντας υπόψη ότι χρησιμοποιείται το ίδιο σύνολο δεδομένων, έτσι η σύγκριση είναι πιο εύκολο να υλοποιηθεί. Οι συγγραφείς χρησιμοποιούν ένα Νευρωνικό Δίκτυο Μακράς και Κρυφής Μνήμης LSTM (Long Short-Term Memory), με σκοπό την κατηγοριοποίηση των συναλλαγών του συνόλου δεδομένων ως απάτης ή όχι. Σε αυτήν τη μελέτη, οι συγγραφείς χρησιμοποιούν μνήμη LSTM με 15 memory cells, πραγματοποιούν 100 εποχές εκπαίδευσης, ενώ η συνάρτηση απώλειας που χρησιμοποιούν είναι η Cross Entropy και ο βελτιστοποιητής που χρησιμοποιούν είναι ο Adam Optimizer. Χρησιμοποιούν ένα κρυφό επίπεδο με 15 νευρώνες. Οι μόνες διαθέσιμες μετρικές στη συγκεκριμένη μελέτη είναι η Περιοχή Κάτω από την Καμπύλη (Area Under Curve - AUC), το Mean Square Error (MSE) και το Root Mean Square Error (RMSE).

### Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [168]

Από αυτήν τη σύγκριση, μπορούμε να δούμε στον παρακάτω πίνακα 9.12 ότι η πλειονότητα των μοντέλων μας επιδοτείται καλύτερα από το μοντέλο των συγγραφέων του [168]. Το καλύτερο από τα μοντέλα μας, όσον αφορά την προσοχή στην Περιοχή Κάτω από την Καμπύλη (AUC), επιτυγχάνει μια βαθμολογία του 99,64% με ένα Mean Square Error (MSE) της τάξεως του 0,0028 και ένα Root Mean Square Error (RMSE) της τάξεως του 0,0060. Επαναλαμβάνουμε για μια ακόμη φορά ότι αυτή η μελέτη δεν παρέχει επαρκή πληροφορία σχετικά με την ακρίβεια, την ανάκληση, το F1 score ή άλλη μετρική εκτός από την AUC, το MSE και το RMSE.

πό αυτήν τη σύγκριση, μπορούμε να δούμε ότι η πλειονότητα των μοντέλων μας με τον βαθύ αυτοκωδικοποιητή επιδοτείται καλύτερα από τα μοντέλα των συγγραφέων του [168]. Το καλύτερο από τα μοντέλα μας επιτυγχάνει μια βαθμολογία του 99,61% όσον αφορά την Περιοχή Κάτω από την Καμπύλη (AUC), και μια βαθμολογία 0,0028 όσον αφορά το Mean Square Error (MSE) και 0,0060 όσον αφορά το Root Mean Square Error (RMSE), ενώ η καλύτερη απόδοση στο [168] έχει μια βαθμολογία του 99,56% στο AUC και βαθμολογία 0,0034 όσον αφορά το MSE και 0,0063 όσον αφορά το RMSE. Για μια ακόμη φορά, τονίζουμε ότι οι συγγραφείς του [168] δεν παρέχουν επαρκή πληροφορία σχετικά με την ακρίβεια, την ανάκληση, το F1 score ή άλλη μετρική.

Πίνακας 9.12: Σύγκριτική αξιολόγηση αποτελεσμάτων μας με [168], LSTM

Models	AUC	MSE	RMSE
SMOTE - AE - LSTM	99,54%	0,0029	0,0054
Borderline SMOTE - AE - LSTM	99,58%	0,0031	0,0062
S.V.M. SMOTE - AE - LSTM	99,60%	0,0033	0,0062
ADASYN - AE - LSTM	98,94%	0,0033	0,0065
K-Means SMOTE - AE - LSTM	99,61%	0,0028	0,0060
SMOTE - VAE - LSTM	99,22%	0,0027	0,0058
Borderline SMOTE - VAE - LSTM	98,68%	0,0026	0,0055
S.V.M. SMOTE - VAE - LSTM	99,16%	0,0026	0,0057
ADASYN - VAE - LSTM	98,53%	0,0032	0,0061
K-Means SMOTE - VAE - LSTM	99,24%	0,0033	0,0059
LSTM ([168] )	99,56%	0,0034	0,0063

## 9.2 Νέα βελτιωμένη τεχνική ερμηνευτικότητας με βάση την τεχνική L.I.M.E.

### 9.2.1 Εισαγωγή

Η Μηχανική Μάθηση έχει μεγάλες δυνατότητες για τη βελτίωση των προϊόντων, των διαδικασιών και της έρευνας. Αλλά οι υπολογιστές τις περισσότερες φορές δεν εξηγούν τις προβλέψεις τους, κάτι που αποτελεί εμπόδιο στην υιοθέτηση της μηχανικής μάθησης. Εκτός από τη χρήση μοντέλων για πρόβλεψη, η ικανότητα ερμηνείας των όσων έχει μάθει ένα μοντέλο λαμβάνει αυξημένη προσοχή στη σύγχρονη έρευνα. Δεν είναι σαφές πώς σχετίζεται το ευρύ φάσμα των προτεινόμενων μεθόδων ερμηνείας και ποιες κοινές έννοιες μπορούν να χρησιμοποιηθούν για την αξιολόγησή τους. Ο στόχος αυτής της ενότητας είναι να προτείνει μια νέα βελτίωση του αλγορίθμου L.I.M.E. που χρησιμοποιεί έναν αυτόματο κωδικοποιητή προκειμένου να δημιουργήσει μια "γειτονιά" γύρω από το datapoint που μας ενδιαφέρει να ερμηνεύσουμε.

Ερμηνευσιμότητα σημαίνει διευκρίνιση ή παρουσίαση με κατανοητούς όρους. Μέσα στο πλαίσιο των συστημάτων ML, ορίζουμε την ερμηνευτικότητα ως η ικανότητα να εξηγεί ή να παρουσιάζει το μοντέλο με κατανοητό τρόπο όρους σε έναν άνθρωπο. Οι ορισμοί που συναντάμε στην επιστημονική κοινότητα δεν προσφέρουν μεγαλύτερη σαφήνεια. Στον τομέα της ψυχολογίας ο συγγραφέας του [176] δηλώνει «οι εξηγήσεις είναι... το νόμισμα με το οποίο ανταλλάξαμε πεποιθήσεις» και σημειώνει ότι ερωτήσεις όπως τι αποτελεί εξήγηση, τι κάνει κάποιες εξηγήσεις καλύτερες από άλλες, το πώς δημιουργούνται οι εξηγήσεις και

το πότε αναζητούνται εξηγήσεις μόλις αρχίζουν να αντιμετωπίζονται ερευνητικά. Οι ερευνητές έχουν ταξινομήσει τις εξηγήσεις από το να είναι «απαγωγικές-νομοτελειακές» [177] (δηλαδή ως λογικές αποδείξεις) για την παροχή κάποιας αίσθησης μηχανισμού [178][179] θεωρείται ένας ευρύτερος ορισμός: σιωπηρή επεξηγηματική κατανόηση. Σε αυτή τη διατριβή, προτείνουμε τρόπους που βασίζονται σε δεδομένα για την εξαγωγή επιχειρησιακών ορισμών και αξιολογήσεων επεξηγήσεων, και επομένως, ερμηνευσιμότητας. Η ερμηνευτικότητα χρησιμοποιείται για να επιβεβαιώσει άλλα σημαντικά ζητήματα των συστημάτων ML, ενώ υπάρχουν πολλά βοηθητικά κριτήρια που κάποιος μπορεί να θέλει να βελτιστοποιήσει. Έννοιες δικαιοσύνης ή αμεροληψίας υπονοούν ότι οι προστατευόμενες ομάδες (ρητές ή σιωπηρές) δεν υφίστανται κατά κάποιο τρόπο διακρίσεις. Μυστικότητα σημαίνει ότι η μέθοδος προστατεύει ευαίσθητες πληροφορίες στα δεδομένα. Ιδιότητες όπως η αξιοπιστία και ευρωστία εξακριβώνουν εάν οι αλγόριθμοι φτάνουν σε ορισμένα επίπεδα απόδοσης έναντι της παραμέτρου ή παραλλαγή εισόδου. Η αιτιότητα υποδηλώνει ότι η προβλεπόμενη αλλαγή στην έξοδο οφείλεται σε διαταραχή που θα συμβεί στο πραγματικό σύστημα. Οι χρησιμοποιήσιμες μέθοδοι παρέχουν πληροφορίες που βοηθούν τους χρήστες να το ολοκληρώσουν μια εργασία - π.χ. ένα κουμπί για να τροποποιήσετε τον φωτισμό της εικόνας—ενώ τα αξιόπιστα συστήματα έχουν την εμπιστοσύνη του ανθρώπου —π.χ. συστήματα αποφυγής σύγκρουσης αεροσκαφών. Ορισμένοι τομείς, όπως η δικαιοσύνη [184], και ιδιωτικότητα [181][182][183] οι ερευνητικές κοινότητες έχουν επισημοποιήσει τα κριτήριά τους και αυτές οι επισημοποιήσεις επέτρεψαν την άνθηση αυστηρής έρευνας σε αυτούς τους τομείς (χωρίς να απαιτείται ερμηνευσιμότητα). Ωστόσο, σε πολλές περιπτώσεις, οι επίσημοι ορισμοί παραμένουν ασαφείς. Ακολουθώντας τη βιβλιογραφία ψυχολογίας, όπου σημειώνει ο [185] «Οι επεξηγήσεις μπορεί να υπογραμμίσουν μια ατέλεια», υποστηρίζουμε ότι η ερμηνευτικότητα μπορεί να βοηθήσει στην ποιοτική διαπίστωση εάν άλλες επιθυμίες—όπως η δικαιοσύνη, η ιδιωτικότητα, η αξιοπιστία, η ευρωστία, η αιτιότητα, η χρηστικότητα και η εμπιστοσύνη — πληρούνται.

## 9.2.2 **Καινομία**

Στο υπό δημοσίευση άρθρο με τίτλο “Novel Interpretable Method using Autoencoders”, εξετάζεται ο συναρπαστικός χώρος της ανθρώπινης αλληλεπίδρασης με τα συστήματα μηχανικής μάθησης, υπογραμμίζοντας τον κεντρικό ρόλο που παίζει η εμπιστοσύνη σε αυτήν τη δυναμική. Κεντρική στον ισχυρισμό μας είναι η διαπίστωση ότι η ερμηνευσιμότητα των ατομικών προβλέψεων αποτελεί θεμέλιο λίθο στην αξιολόγηση της αξιοπιστίας. Με ένα επι-

δραστικό βλέμμα στην καινοτομία, παρουσιάζουμε ένα σημαντικό βήμα προόδου στη μορφή του αλγορίθμου VAE-L.I.M.E., ένα εξελικτικό βήμα πέρα από τη συμβατική προσέγγιση L.I.M.E..

Στην τεχνική L.I.M.E. χρησιμοποιούνται ερμηνεύσιμα μοντέλα που χρησιμοποιούνται για να εξηγήσουν τις ατομικές προβλέψεις των αδιαφανών μοντέλων μηχανικής μάθησης (black-boxes). Οι ερμηνεύσιμες μοντέλο-ανεξάρτητες εξηγήσεις (L.I.M.E.) [186] είναι μια εργασία στην οποία οι συγγραφείς προτείνουν μια υλοποίηση των τοπικών ερμηνεύσιμων μοντέλων. Τα μοντέλα εκπαιδεύονται για να προσεγγίσουν τις προβλέψεις του υποκείμενου αδιαφανούς μοντέλου. Αντί να εκπαιδεύσουμε ένα γενικό υποκατάστατο μοντέλο, το L.I.M.E. επικεντρώνεται στην εκπαίδευση τοπικών υποκαταστάτων μοντέλων για να εξηγήσει τις ατομικές προβλέψεις.

Το L.I.M.E. δοκιμάζει τι συμβαίνει με τις προβλέψεις όταν δίνονται παραλλαγές των δεδομένων στο μοντέλο μηχανικής μάθησης. Το L.I.M.E. δημιουργεί ένα νέο σύνολο δεδομένων που αποτελείται από παραμορφωμένα δείγματα και τις αντίστοιχες προβλέψεις του αδιαφανούς μοντέλου, τα οποία προκύπτουν από μια κανονική κατανομή. Σε αυτό το νέο σύνολο δεδομένων, το L.I.M.E. εκπαιδεύει ένα ερμηνεύσιμο μοντέλο, το οποίο είναι καθοδηγούμενο από την εγγύτητα των δειγμάτων που δείγματος προς το δείγμα ενδιαφέροντος. Για την τελική ερμηνεία, μπορεί να χρησιμοποιηθούν μοντέλα γραμμικής παλινδρόμησης, Lasso ή Δέντρα Απόφασης. Το εκπαιδευμένο μοντέλο πρέπει να αποτελεί καλή προσέγγιση των προβλέψεων του μοντέλου μηχανικής μάθησης σε τοπικό επίπεδο, αλλά δεν αποτελεί καλή γενική προσέγγιση.

Το μοντέλο εξήγησης για το στοιχείο  $x$  είναι το μοντέλο  $h$  (π.χ. μοντέλο γραμμικής παλινδρόμησης) που ελαχιστοποιεί το σφάλμα  $L$  (π.χ. μέσο τετραγωνικό σφάλμα), το οποίο μετρά πόσο κοντά είναι η εξήγηση στην πρόβλεψη του αρχικού μοντέλου  $f$  (π.χ. ένα μοντέλο  $xg$ -boost), ενώ η πολυπλοκότητα του μοντέλου  $\Omega(h)$  παραμένει χαμηλή (π.χ. προτιμώντας λιγότερα χαρακτηριστικά). Η  $\Omega$  αντιπροσωπεύει την οικογένεια των πιθανών εξηγήσεων, όπως όλα τα πιθανά μοντέλα γραμμικής παλινδρόμησης. Το μέτρο εγγύτητας  $\pi_x$  καθορίζει πόσο μεγάλη είναι η γειτονιά γύρω από το στοιχείο  $x$  που λαμβάνουμε υπόψη για την εξήγηση. Στην πράξη, το L.I.M.E. βελτιστοποιεί μόνο το μέρος του σφάλματος. Ο χρήστης πρέπει να καθορίσει την πολυπλοκότητα, π.χ. επιλέγοντας το μέγιστο αριθμό χαρακτηριστικών που το μοντέλο γραμμικής παλινδρόμησης μπορεί να χρησιμοποιήσει.

Ο αλγόριθμος για την εκπαίδευση τοπικών υποκαταστάτων μοντέλων είναι η εξής:

- Δημιουργία νέου συνόλου δεδομένων και λήψη προβλέψεων για αυτά τα νέα σημεία.
- Στάθμιση των νέων δειγμάτων ανάλογα με την εγγύτητά τους στο data point.
- Εκπαίδευση ενός νέου ερμηνεύσιμου μοντέλου στο σύνολο δεδομένων με τις παραλλαγές.
- Εξήγηση της πρόβλεψης με την ερμηνεία του τοπικού μοντέλου.

Ορισμός μιας σημαντικής γειτονιάς γύρω από ένα σημείο είναι δύσκολος. Το L.I.M.E. χρησιμοποιεί αυτήν τη στιγμή έναν πυρήνα εκθετικής εξομάλυνσης για τον καθορισμό της γειτονιάς. Η υλοποίηση του L.I.M.E. χρησιμοποιεί έναν πυρήνα εκθετικής εξομάλυνσης (σε κανονικοποιημένα δεδομένα), ενώ το kernel είναι 0,75 φορές η τετραγωνική ρίζα του αριθμού των στηλών των δεδομένων εκπαίδευσης. Το πρόβλημα είναι ότι δεν είναι δυνατό να καθοριστεί ο καλύτερο kernel και καθορίζεται αυθαίρετα από τους συγγραφείς του [186]. Ένα άλλο μειονέκτημα του L.I.M.E. είναι η αστάθεια των εξηγήσεων. Στο [187] οι συγγραφείς έδειξαν ότι οι εξηγήσεις δύο πολύ κοντινών σημείων έδιναν πολύ διαφορετικά αποτελέσματα σε ένα προσομοιωμένο περιβάλλον. Η αστάθεια σημαίνει ότι είναι δύσκολο να εμπιστευέστε τις εξηγήσεις και πρέπει να είστε επιφυλακτικοί με τα αποτελέσματα. Επιπλέον, οι εξηγήσεις του L.I.M.E. μπορούν να παραπλανήσουν τον επιστήμονα δεδομένων, όπως δείχνει το [188]. Η δυνατότητα παραπλάνησης καθιστά πιο δύσκολο το να εμπιστευέστε τις εξηγήσεις που δημιουργούνται με το L.I.M.E..

Η καινοτομία της προσέγγισής μας σε σύγκριση με τις σχετικές έρευνες στον τομέα παρουσιάζεται μέσω της ακόλουθης λίστας:

- Προτείνουμε ένα σύνολο μοντέλων που χρησιμοποιούν ένα Variational Autoencoder για να δημιουργήσουν νέα παραδείγματα γύρω από το σημείο δεδομένων που μας ενδιαφέρει για την ερμηνεία.
- Παρέχουμε μια εκτενή πειραματική αξιολόγηση για να αποκαλύψουμε τα πλεονεκτήματα και τα αντιστρόφως του προτεινόμενου σχήματος. Η αξιολόγηση μας μπορεί να θεωρηθεί ως μια συγκριτική μελέτη μεταξύ της κύριας μεθόδου ερμηνείας (L.I.M.E.) και της προτεινόμενης μας προσέγγισης (VAE - L.I.M.E.).
- Τέλος, παρουσιάζουμε μια εκτενή σύγκριση μεταξύ του αλγορίθμου L.I.M.E., προκειμένου να συγκρίνουμε τα αποτελέσματα με το προτεινόμενο αλγόριθμο VAE - L.I.M.E..

Στην καρδιά του VAE-L.I.M.E. κρύβεται μια καινοτόμος λύση σε έναν διαρκή πρόκληση: η ακριβής δημιουργία γειτονικών περιοχών γύρω από ένα συγκεκριμένο σημείο δεδομένων ερμηνευτικού ενδιαφέροντος. Αυτό το πρόβλημα αντιμετωπίζεται επιδέξια μέσω της ενσωμάτωσης ενός μεταβαλλόμενου αυτοκωδικοποιητή (VAE). Ο VAE, που εκπαιδεύεται επιμελώς σε έναν εκτενή σύνολο δεδομένων, μας παραχωρεί τη δυνατότητα να δημιουργήσουμε μια κατανομή που χαρακτηρίζεται από τη μέση της τιμή και την τυπική απόκλιση της. Αυτός ο εξυπνος μηχανισμός μας δίνει τη δυνατότητα να δημιουργήσουμε αυτήν τη γειτονική περιοχή που αναφέρθηκε παραπάνω, ξεπερνώντας έτσι ένα προηγουμένως περίπλοκο εμπόδιο.

Είναι σημαντικό να σημειώσουμε ότι το πεδίο της ερμηνείας επεκτείνει την επίδρασή του σε μια πληθώρα εργασιών, περιλαμβάνοντας όχι μόνο τα δεδομένα σε πίνακες, αλλά και τους τομείς των κειμενικών και εικονικών πληροφοριών. Αυτή η αξία της ερμηνεύσιμης εξήγησης υπογραμμίζει την παντού παρουσία της στο ευρύ τοπίο των εφαρμογών της μηχανικής μάθησης.

### 9.2.3 Πειραματικά Αποτελέσματα

Στα πλαίσια αυτής της ενότητας, ευθυγραμμίζομαστε αβίαστα με την πορεία που έχει χαραχτεί η αντίστοιχη ενότητα 5.1 στο [186]. Εδώ, αναπαράγουμε τα ίδια πειράματα που περιγράφονται αναλυτικά στην εν λόγω ενότητα, προσεκτικά σχεδιασμένα για να αντικατοπτρίζουν με ακρίβεια τις μεθοδολογίες και τα παραδείγματα που περιέχονται εκεί.

Για όσους επιθυμούν να εξερευνήσουν περαιτέρω, το αποθετήριο μας στο GitHub, προσβάσιμο στο <https://github.com/zioniris/L.I.M.E.-experiments>, αποτελεί απόδειξη της αφοσίωσής μας στη διαφάνεια και την αναπαραγωγικότητα.

Μέσα σε αυτό το σύνολο πειραμάτων που έχουν σχεδιαστεί με απόλυτη προσοχή, η μεθοδολογία μας ευθυγραμμίζεται με το έδαφος που έχουν προετοιμάσει οι συγγραφείς του [186], επιτρέποντάς μας να προσομοιώσουμε το σώμα της εμπιστοσύνης στις ατομικές προβλέψεις. Αυτή η προσομοίωση επιτυγχάνεται μέσω μιας πολυπλοκότητας προσεγμένης προσέγγισης, ξεκινώντας με την τυχαία επιλογή του 25% των χαρακτηριστικών ως "αναξιόπιστα". Εμπεριέχεται σε αυτή τη δομή η υπόθεση ότι οι χρήστες διαθέτουν την διακριτικότητα για να αναγνωρίσουν και στη συνέχεια να αποσύρουν την εμπιστοσύνη από αυτά τα καθορισμένα "αναξιόπιστα" χαρακτηριστικά.

Ως λογική επέκταση, αναλαμβάνουμε το καθήκον να καθορίσουμε την "αξιοπιστία" των προβλέψεων με τη χρήση των αποτελεσμάτων του μαύρου κουτιού ταξινομητή στο σύνολο

δοκιμής. Σημαντικό είναι ότι οι προβλέψεις θεωρούνται "αναξιόπιστες" εάν εμφανίζουν αλλοιώσεις μετά την αφαίρεση των αναξιόπιστων χαρακτηριστικών, ενώ αντιστρόφως, οι προβλέψεις παραμένουν "αξιόπιστες" εάν δεν εμφανίζονται τέτοιες αποκλίσεις. Αυτό το πλαίσιο διαχειρίζεται μια λεπτή αλληλεπίδραση μεταξύ της εγγενούς αξιοπιστίας των χαρακτηριστικών και του ρόλου τους στον προσδιορισμό των προβλέψεων.

Για να προσομοιώσουμε τη συμπεριφορά του χρήστη, εισάγουμε ένα μοντέλο χρήστη το οποίο, όταν αντιμετωπίζει αναξιόπιστες προβλέψεις, ακολουθεί έναν διαδικασία αξιολόγησης. Αυτή η αξιολόγηση εξαρτάται από κριτήρια όπως η γραμμικότητα των προβλέψεων και η επίδραση της αφαίρεσης των αναξιόπιστων χαρακτηριστικών από τις εξηγήσεις - μια διαδικασία που μιμείται την "εκπτωτική" επίδραση των αναξιόπιστων χαρακτηριστικών στην ανθρώπινη λήψη αποφάσεων.

Διάφορες στρατηγικές, με διαφορετικές υποκείμενες υποθέσεις, καθοδηγούν την αξιολόγηση του χρήστη. Για μεθόδους όπως η greedy και η random, η αναξιόπιστη προσγείωση εξαρτάται από την παρουσία αναξιόπιστων χαρακτηριστικών στις εξηγήσεις, λόγω της περιορισμένης τους ικανότητας να αξιολογήσουν τις συνεισφορές των χαρακτηριστικών. Ταυτόχρονα, τόσο για το L.I.M.E. όσο και για το VAE-L.I.M.E., ο χρήστης αξιολογεί την αξιοπιστία των προβλέψεων με βάση τη γραμμική προσέγγιση και την αφαίρεση των αναξιόπιστων χαρακτηριστικών από τις εξηγήσεις.

Δεδομένου αυτού του περίπλοκου πλαισίου, παρουσιάζουμε τα ευρήματά μας σε μια ποσοτική αξιολόγηση με την αναφορά του F1 score για αξιόπιστες προβλέψεις. Αυτή η αξιολόγηση, που διενεργείται σε 100 εκτελέσεις, καταγράφεται λεπτομερώς στους πίνακες 9.13 και 9.14. Οι προσεκτικές αναλύσεις μας αποκαλύπτουν μια εκπληκτική τάση: το VAE-L.I.M.E. υπερέχει συνεχώς του L.I.M.E. με σημαντική σημασία (όλα τα αποτελέσματα σημειώνονται με  $p = 0.01$ ) και στα δύο σύνολα δεδομένων και στο φάσμα των μαύρων κουτιών μοντέλων. Ο ανταγωνιστικός προβληματισμός του VAE-L.I.M.E. γίνεται εμφανής όταν εναλλακτικές μέθοδοι αποτυγχάνουν, είτε λόγω εκπτωτικού όγκου ή ακρίβειας. Σε αντίθεση, το VAE-L.I.M.E., μαζί με το L.I.M.E., διατηρούν μια συμβιωτική ισορροπία υψηλής ακρίβειας και ανθεκτικής αποκατάστασης, συνεπώς αγκαλιάζοντας την ουσία τόσο της ακρίβειας όσο και της ερμηνευσιμότητας.

**Πίνακας 9.13: Average F1 of trustworthiness for different explainers on a collection of classifiers regarding the Books Dataset Ribeiro**

<b>Books Dataset</b>				
Algorithm	LR	NN	RF	S.V.M.
Random	14.6	14.8	14.7	14.7
Greedy	53.7	47.4	45.0	53.3
L.I.M.E.	96.6	94.5	96.2	96.7
VAE - L.I.M.E.	97.4	95.1	96.2	97.2

**Πίνακας 9.14: Average F1 of trustworthiness for different explainers on a collection of classifiers regarding the DVDs Dataset Ribeiro**

<b>DVDs Dataset</b>				
Algorithm	LR	NN	RF	S.V.M.
Random	14.2	14.3	14.5	14.4
Greedy	52.4	58.1	46.6	55.1
L.I.M.E.	96.6	91.8	96.1	95.6
VAE - L.I.M.E.	96.5	93.1	97.5	96.1

# Κεφάλαιο 10

## Συζήτηση

### 10.1 Ερμηνεία των αποτελεσμάτων

Στο πεδίο της έρευνάς μας, η ερμηνεία των αποτελεσμάτων ήταν υψίστης σημασίας για την εξαγωγή πολύτιμων πληροφοριών και συμπερασμάτων από τις προγνωστικές δυνατότητες των μοντέλων. Η μελέτη αυτή είχε ως στόχο να προχωρήσει πέρα από την απλή παραγωγή αριθμητικών αποτελεσμάτων- επιδίωξε να διαλευκάνει τις περίπλοκες αλληλεπιδράσεις μεταξύ των μακροοικονομικών παραγόντων και του πιστωτικού κινδύνου, ενώ παράλληλα κατέδειξε τις δυνατότητες των προηγμένων μεθοδολογιών μηχανικής μάθησης και βαθιάς μάθησης σε αυτό το πλαίσιο.

Τα περιγραφικά στατιστικά στοιχεία και η ανάλυση συσχέτισης αποκάλυψαν κάποιες πρώτες πληροφορίες για τις σχέσεις μεταξύ των μακροοικονομικών παραγόντων και του εταιρικού πιστωτικού κινδύνου. Μακροοικονομικοί δείκτες όπως ο ρυθμός πληθωρισμού, η αύξηση του ΑΕΠ, το ποσοστό ανεργίας και το επιτόκιο παρουσίασαν ποικίλους βαθμούς συσχέτισης με τις μεταβλητές του πιστωτικού κινδύνου. Το εύρημα αυτό, ενώ συνάδει με τις προσδοκίες που καθοδηγούνται από το θεωρητικό πλαίσιο, υπογράμμισε τη σημασία των μακροοικονομικών εκτιμήσεων στην ανάλυση του πιστωτικού κινδύνου.

Επιπλέον, τα μοντέλα βαθιάς μάθησης που χρησιμοποιήθηκαν στην παρούσα μελέτη εντόπισαν με επιτυχία πολύπλοκες μη γραμμικές σχέσεις μεταξύ αυτών των μακροοικονομικών παραγόντων και του εταιρικού πιστωτικού κινδύνου, οι οποίες δεν ήταν εύκολα αναγνωρίσιμες μέσω παραδοσιακών στατιστικών μεθόδων ή βασικών μοντέλων μηχανικής μάθησης. Η ανθεκτικότητα των μοντέλων βαθιάς μάθησης έναντι αλλαγών στις παραμέτρους, καθώς και η συνεπής απόδοσή τους υπό διάφορα οικονομικά σενάρια και σε διαφορετικά τμήματα του

συνόλου δεδομένων, κατέδειξαν την ικανότητά τους να παρέχουν αξιόπιστες και σταθερές προβλέψεις.

Το υβριδικό μοντέλο βαθιάς μάθησης, το οποίο συνδύαζε τα πλεονεκτήματα τόσο των αυτοκωδικοποιητών για την επιλογή χαρακτηριστικών και τη μείωση της διαστατικότητας όσο και των νευρωνικών δικτύων συνελκτικού τύπου για την ανίχνευση σύνθετων μοτίβων, αναδείχθηκε ως ένα ιδιαίτερα ισχυρό εργαλείο. Αυτό το μοντέλο όχι μόνο ξεπέρασε τα άλλα μοντέλα όσον αφορά την ακρίβεια πρόβλεψης, αλλά επέδειξε επίσης εντυπωσιακή ικανότητα προσαρμογής σε νέα, αθέατα δεδομένα.

Ωστόσο, ενώ το υβριδικό μοντέλο βαθιάς μάθησης έδειξε ισχυρές επιδόσεις, παρατηρήθηκε ότι η ερμηνευσιμότητα αυτού του μοντέλου ήταν σχετικά χαμηλή, όπως είναι χαρακτηριστικό των μοντέλων βαθιάς μάθησης. Για να αντιμετωπίσουμε αυτό το ζήτημα, πραγματοποιήσαμε μια ανάλυση σημαντικότητας χαρακτηριστικών, η οποία παρείχε κάποιο βαθμό διαφάνειας στο μοντέλο, επισημαίνοντας τα χαρακτηριστικά με τη μεγαλύτερη επιρροή στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου. Η ανάλυση αυτή έδειξε ότι ορισμένες μακροοικονομικές μεταβλητές και χρηματοοικονομικοί δείκτες είχαν μεγαλύτερη δυνατότητα πρόβλεψης από άλλες, καθοδηγώντας έτσι τις μελλοντικές προσπάθειες συλλογής δεδομένων και τις στρατηγικές διαχείρισης του πιστωτικού κινδύνου.

Επιπλέον, η τεχνική υπερδειγματοληψίας που χρησιμοποιήσαμε για τη διαχείριση της ανισορροπίας των κλάσεων στα δεδομένα μας, δηλαδή η τεχνική συνθετικής υπερδειγματοληψίας μειοψηφιών (SMOTE), ενίσχυσε την απόδοση των μοντέλων μας. Τους επέτρεψε να μαθαίνουν καλύτερα από την κλάση της μειονότητας (προεπιλεγμένες περιπτώσεις), γεγονός που οδήγησε σε πιο ισορροπημένες και ακριβείς προβλέψεις.

Τα αποτελέσματα αυτής της μελέτης προσέφεραν νέες γνώσεις σχετικά με το ρόλο των μακροοικονομικών παραγόντων στην ανάλυση του εταιρικού πιστωτικού κινδύνου. Υπογράμμισαν την ανάγκη υιοθέτησης μιας ευρείας, ολιστικής προσέγγισης που λαμβάνει υπόψη τόσο τους παράγοντες σε μικροεπίπεδο όσο και σε μακροεπίπεδο. Επιπλέον, η επιτυχής εφαρμογή προηγμένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης σε αυτό το πλαίσιο υπογράμμισε τις δυνατότητες αυτών των μεθόδων για τη βελτίωση της ακρίβειας και της αξιοπιστίας της ανάλυσης πιστωτικού κινδύνου, γεγονός που έχει σημαντικές συνέπειες για τον χρηματοπιστωτικό κλάδο.

Εν κατακλείδι, ενώ τα μοντέλα μας έχουν επιδείξει ελπιδοφόρα αποτελέσματα, είναι ζωτικής σημασίας να ερμηνεύσουμε αυτά τα ευρήματα στο ευρύτερο πλαίσιο της ανάλυσης του

πιστωτικού κινδύνου και του εξελισσόμενου οικονομικού τοπίου. Είναι επίσης σημαντικό να επαναλαμβάνουμε και να τελειοποιούμε συνεχώς αυτά τα μοντέλα, καθώς διατίθενται περισσότερα δεδομένα και καθώς η κατανόηση της πολύπλοκης δυναμικής που παίζει ρόλο στην ανάλυση πιστωτικού κινδύνου συνεχίζει να εξελίσσεται.

## 10.2 Συνέπειες για τη θεωρία και την πράξη

Τα αποτελέσματα της παρούσας μελέτης έχουν βαθιές επιπτώσεις τόσο στις θεωρητικές όσο και στις πρακτικές πτυχές της ανάλυσης του εταιρικού πιστωτικού κινδύνου.

Θεωρητικά, τα ευρήματά μας παρουσιάζουν μια εξέλιξη στην κατανόηση της αλληλεπίδρασης μεταξύ των μακροοικονομικών παραγόντων και του εταιρικού πιστωτικού κινδύνου. Η ένδειξη ότι ορισμένοι μακροοικονομικοί δείκτες έχουν ισχυρότερη συσχέτιση με τον εταιρικό πιστωτικό κίνδυνο από ό,τι άλλοι τονίζει την ανάγκη περαιτέρω βελτίωσης και προώθησης των οικονομικών και χρηματοοικονομικών θεωριών. Αναγνωρίζοντας τη μη γραμμική και πολύπλοκη σχέση που υπάρχει μεταξύ αυτών των μεταβλητών, οι ερευνητές και οι θεωρητικοί μπορούν να κατανοήσουν καλύτερα τις αποχρώσεις και τις ιδιομορφίες που χαρακτηρίζουν τη δυναμική του εταιρικού πιστωτικού κινδύνου σε πραγματικές συνθήκες.

Η επιτυχής εφαρμογή υβριδικών μοντέλων βαθιάς μάθησης παρέχει επίσης έναυσμα για την επαναξιολόγηση των παραδοσιακών μεθόδων ανάλυσης πιστωτικού κινδύνου. Ενώ τα κλασικά χρηματοοικονομικά μοντέλα επικεντρώνονται σε γραμμικές σχέσεις και βασίζονται σε μεγάλο βαθμό σε υποθέσεις όπως οι κανονικές κατανομές και η ομοσκεδαστικότητα, τα ευρήματά μας υπογραμμίζουν τις δυνατότητες ενσωμάτωσης τεχνικών βαθιάς μάθησης και τεχνητής νοημοσύνης στις χρηματοοικονομικές θεωρίες και μοντέλα.

Επιπλέον, η ικανότητα του υβριδικού μοντέλου βαθιάς μάθησης να προβλέπει με ακρίβεια τον πιστωτικό κίνδυνο ακόμη και παρουσία μη ισορροπημένων συνόλων δεδομένων αμφισβητεί τη συμβατική αντίληψη ότι οι προβλέψεις υψηλής ποιότητας απαιτούν ισορροπημένα σύνολα δεδομένων. Η χρήση της τεχνικής της συνθετικής υπερδειγματοληψίας μειωψηφίας (SMOTE) στη μελέτη μας όχι μόνο ανοίγει δρόμους για μελλοντική έρευνα αλλά και προσδίδει αξιοπιστία στις θεωρητικές δυνατότητες αυτής της τεχνικής για την ενίσχυση της απόδοσης των μοντέλων πρόβλεψης ενόψει ανισορροπίας κλάσεων.

Στην πράξη, τα ευρήματα της παρούσας μελέτης προσφέρουν πολύτιμες πληροφορίες για τους επαγγελματίες του κλάδου, ιδίως για εκείνους που δραστηριοποιούνται στον τραπεζικό

και χρηματοπιστωτικό τομέα. Η αποκάλυψη ότι οι μακροοικονομικοί παράγοντες διαδραματίζουν σημαντικό ρόλο στον καθορισμό των επιπέδων του εταιρικού πιστωτικού κινδύνου καθιστά αναγκαία την αναδιαμόρφωση των στρατηγικών διαχείρισης κινδύνων. Αντί να βασίζονται αποκλειστικά σε χρηματοοικονομικές πληροφορίες που αφορούν συγκεκριμένες επιχειρήσεις, τα χρηματοπιστωτικά ιδρύματα μπορούν να ενσωματώσουν μακροοικονομικά δεδομένα στις διαδικασίες εκτίμησης κινδύνου για να βελτιώσουν την ακρίβεια πρόβλεψής τους.

Επιπλέον, η εφαρμογή προηγμένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης έχει πρακτικές εφαρμογές για την ψηφιοποίηση και την αυτοματοποίηση της ανάλυσης πιστωτικού κινδύνου. Με την εφαρμογή τέτοιων μοντέλων, τα χρηματοπιστωτικά ιδρύματα μπορούν να αυξήσουν την ταχύτητα και την αποτελεσματικότητα των διαδικασιών αξιολόγησης του κινδύνου τους, επιτρέποντάς τους έτσι να ανταποκρίνονται άμεσα στις μεταβολές των επιπέδων πιστωτικού κινδύνου και να ελαχιστοποιούν τις πιθανές απώλειες. Οι προκλήσεις ερμηνευσιμότητας που σχετίζονται με αυτά τα μοντέλα μπορούν να αντιμετωπιστούν εν μέρει με τη διενέργεια ανάλυσης σημαντικότητας χαρακτηριστικών, η οποία παρέχει πληροφορίες σχετικά με τους παράγοντες επιρροής που καθοδηγούν τις προβλέψεις.

Οι επαγγελματίες μπορούν επίσης να επωφεληθούν από την επίδειξη της αποτελεσματικής διαχείρισης ανισόρροπων συνόλων δεδομένων. Καθώς τα σύνολα δεδομένων πιστωτικού κινδύνου συχνά χαρακτηρίζονται από υψηλότερο ποσοστό περιπτώσεων μη αθέτησης, η εφαρμογή της τεχνικής SMOTE από εμάς παρέχει μια πρακτική λύση για την αντιμετώπιση τέτοιων ανισορροπιών κλάσεων, ενισχύοντας ενδεχομένως την απόδοση των προγνωστικών μοντέλων.

Ωστόσο, είναι σημαντικό οι επαγγελματίες να κατανοήσουν ότι, ενώ τα μοντέλα μηχανικής μάθησης και βαθιάς μάθησης που χρησιμοποιήθηκαν στην παρούσα μελέτη παρουσιάζουν πολλά υποσχόμενα αποτελέσματα, δεν είναι αλάνθαστα. Αυτά τα μοντέλα είναι τόσο καλά όσο τα δεδομένα στα οποία εκπαιδεύονται και πρέπει να ενημερώνονται συνεχώς με νέα δεδομένα ώστε να αντικατοπτρίζουν τις μεταβαλλόμενες οικονομικές συνθήκες και τις επιχειρηματικές πραγματικότητες. Επιπλέον, οι επαγγελματίες θα πρέπει να έχουν επίγνωση των πιθανών ηθικών επιπτώσεων που συνδέονται με τη χρήση τέτοιων προηγμένων μοντέλων, όπως ζητήματα που σχετίζονται με τη διαφάνεια, τη λογοδοσία και την προστασία της ιδιωτικής ζωής των δεδομένων.

Συμπερασματικά, η παρούσα μελέτη συμβάλλει στο υπάρχον σώμα γνώσεων στον το-

μέα της ανάλυσης του εταιρικού πιστωτικού κινδύνου, καταδεικνύοντας τις δυνατότητες της μηχανικής μάθησης και των μοντέλων βαθιάς μάθησης στην ενίσχυση της κατανόησης των πολύπλοκων αλληλεπιδράσεων μεταξύ των μακροοικονομικών παραγόντων και του πιστωτικού κινδύνου. Οι συνέπειες για τη θεωρία και την πράξη αναδεικνύουν τη δυνατότητα αυτών των μοντέλων να φέρουν επανάσταση στον τρόπο με τον οποίο διεξάγεται η ανάλυση πιστωτικού κινδύνου, ανοίγοντας νέους δρόμους για την έρευνα και τις πρακτικές εφαρμογές στον τομέα.

### 10.3 Περιορισμοί και μελλοντικές ερευνητικές κατευθύνσεις

Η έρευνα που παρουσιάζεται στην παρούσα διατριβή συμβάλλει σημαντικά στον τομέα της ανάλυσης του εταιρικού πιστωτικού κινδύνου με την εφαρμογή καινοτόμων τεχνικών μηχανικής μάθησης και βαθιάς μάθησης. Παρ' όλα αυτά, δεν στερείται των περιορισμών της, οι οποίοι, με τη σειρά τους, παρέχουν δρόμους για μελλοντική έρευνα.

Πρώτον, η χρήση ενός περιορισμένου συνόλου μακροοικονομικών μεταβλητών, ενώ μας επιτρέπει να διατηρήσουμε την εστίαση, μπορεί να παραβλέψει άλλους δυνητικά σχετικούς μακροοικονομικούς δείκτες. Η μελλοντική έρευνα θα μπορούσε να διευρύνει το πεδίο των μακροοικονομικών παραγόντων που διερευνώνται, ενσωματώνοντας πτυχές όπως οι μεταβλητές του διεθνούς εμπορίου, οι τομεακοί δείκτες ή οι μακροοικονομικοί παράγοντες ανά χώρα. Αυτό θα εμπλουτίσει περαιτέρω τη μακροοικονομική προοπτική της ανάλυσης του εταιρικού πιστωτικού κινδύνου και θα παράσχει μια πιο λεπτομερή και αποχρωματισμένη κατανόηση του θέματος.

Δεύτερον, η έρευνά μας χρησιμοποιεί ένα σύνολο δεδομένων μίας χώρας, το οποίο περιορίζει εγγενώς τη γεωγραφική γενίκευση των συμπερασμάτων μας. Δεδομένων των διαφορετικών οικονομικών δομών και συνθηκών της αγοράς στις διάφορες χώρες, δεν είναι βέβαιο ότι τα συμπεράσματα που εξάγονται από την παρούσα μελέτη μπορούν να έχουν καθολική εφαρμογή. Συνεπώς, μελλοντικές έρευνες θα μπορούσαν να εξετάσουν το ενδεχόμενο διεξαγωγής διακρατικών μελετών ή να εστιάσουν σε αναδυόμενες οικονομίες, όπου η ανάλυση του εταιρικού πιστωτικού κινδύνου θα μπορούσε να διαφέρει σημαντικά λόγω των διαφορετικών δυναμικών της αγοράς και των οικονομικών συνθηκών.

Τρίτον, η μελέτη μας χρησιμοποιεί μοντέλα βαθιάς μάθησης, τα οποία, αν και επιδει-

κνύουν πολλά υποσχόμενες προγνωστικές επιδόσεις, συχνά επικρίνονται για τη φύση τους ως "μαύρου κουτιού". Παρά τις προσπάθειές μας να ερμηνεύσουμε τη σημασία των χαρακτηριστικών, η πλήρης κατανόηση του τρόπου με τον οποίο αυτά τα μοντέλα κάνουν τις προβλέψεις τους παραμένει ασύλληπτη. Η μελλοντική έρευνα θα μπορούσε να διερευνήσει τη χρήση τεχνικών εξηγήσιμης TN (XAI) ή άλλων προσεγγίσεων ερμηνευσιμότητας για να ρίξει περισσότερο φως στις διαδικασίες λήψης αποφάσεων αυτών των πολύπλοκων μοντέλων.

Τέταρτον, αν και αντιμετωπίσαμε το ζήτημα της ανισορροπίας των συνόλων δεδομένων με την εφαρμογή της SMOTE, η μέθοδος αυτή έχει τους περιορισμούς της, όπως η πιθανή υπερ- γενίκευση και η εισαγωγή θορύβου. Η μελλοντική έρευνα θα μπορούσε να εξετάσει άλλες μεθόδους επαναδειγματοληψίας ή προηγμένους αλγορίθμους που έχουν σχεδιαστεί για να χειρίζονται ανισόρροπα σύνολα δεδομένων, ενισχύοντας έτσι την ευρωστία και την αξιοπιστία των προβλέψεων του πιστωτικού κινδύνου.

Τέλος, αν και η έρευνά μας έχει επικεντρωθεί στην προγνωστική πτυχή του εταιρικού πιστωτικού κινδύνου, η κανονιστική διάσταση του προβλήματος παραμένει σχετικά ανεξερεύνητη. Η μελλοντική έρευνα θα μπορούσε να εξετάσει το ενδεχόμενο ενσωμάτωσης διαδικασιών λήψης αποφάσεων στα μοντέλα, αναπτύσσοντας συνταγογραφικές αναλύσεις που θα παρέχουν συγκεκριμένες συστάσεις για τη μείωση ή τη διαχείριση του πιστωτικού κινδύνου, με βάση τις προβλέψεις που γίνονται.

Επιπλέον, με την ταχεία πρόοδο των τεχνολογιών τεχνητής νοημοσύνης και μηχανικής μάθησης, υπάρχει η δυνατότητα ανάπτυξης νέων μεθόδων ή τεχνικών που θα μπορούσαν να βελτιώσουν την απόδοση της ανάλυσης πιστωτικού κινδύνου. Για παράδειγμα, η ομοσπονδιακή μάθηση, μια προσέγγιση μηχανικής μάθησης που επιτρέπει την εκπαίδευση μοντέλων σε ένα αποκεντρωμένο δίκτυο, κερδίζει την προσοχή της ερευνητικής κοινότητας της TN. Η προσέγγιση αυτή θα μπορούσε να επιτρέψει στα χρηματοπιστωτικά ιδρύματα να εκπαιδεύουν συνεργατικά μοντέλα πρόβλεψης πιστωτικού κινδύνου, διατηρώντας παράλληλα το απόρρητο των δεδομένων.

Εν κατακλείδι, ενώ η παρούσα μελέτη έχει κάνει σημαντικά βήματα στην ενσωμάτωση μακροοικονομικών παραγόντων στην ανάλυση του εταιρικού πιστωτικού κινδύνου με τη χρήση μοντέλων μηχανικής μάθησης και βαθιάς μάθησης, παραμένουν πολλά που πρέπει να διερευνηθούν και να βελτιωθούν. Οι περιορισμοί που εντοπίστηκαν στην παρούσα μελέτη δεν αποτελούν εμπόδια, αλλά σκαλοπάτια προς την περαιτέρω ενίσχυση της κατανόησης και

της ικανότητάς μας να προβλέπουμε και να διαχειριζόμαστε τον εταιρικό πιστωτικό κίνδυνο σε ένα όλο και πιο σύνθετο και ασταθές οικονομικό τοπίο.



# Κεφάλαιο 11

## Συμπέρασμα

### 11.1 Σύνοψη των ευρημάτων

Η παρούσα έρευνα αποσκοπούσε στη βελτίωση της ανάλυσης του εταιρικού πιστωτικού κινδύνου με την ενσωμάτωση μακροοικονομικών παραγόντων μέσω της εφαρμογής μοντέλων μηχανικής μάθησης και βαθιάς μάθησης. Με τον τρόπο αυτό, γεφυρώσαμε κενά στην υπάρχουσα βιβλιογραφία και προσφέραμε μια νέα προοπτική στον τομέα της ανάλυσης πιστωτικού κινδύνου. Στο παρόν, παρουσιάζουμε μια περιεκτική περίληψη των βασικών ευρημάτων μας.

Η μελέτη μας ξεκίνησε με τη διερεύνηση της επιρροής των μακροοικονομικών μεταβλητών στον εταιρικό πιστωτικό κίνδυνο. Η περιγραφική στατιστική και η ανάλυση συσχέτισης διασαφήνισαν τις σχέσεις μεταξύ των διαφόρων μακροοικονομικών δεικτών και του εταιρικού πιστωτικού κινδύνου, υπογραμμίζοντας ότι οι μακροοικονομικές συνθήκες επηρεάζουν σημαντικά το προφίλ του εταιρικού πιστωτικού κινδύνου. Η διαπίστωση αυτή υπογράμμισε την ανάγκη για πιο ολοκληρωμένα μοντέλα πιστωτικού κινδύνου που ενσωματώνουν τόσο τα χρηματοοικονομικά δεδομένα που αφορούν συγκεκριμένες επιχειρήσεις όσο και τις ευρύτερες μακροοικονομικές συνθήκες.

Προσπαθήσαμε επίσης να επιλέξουμε τα πιο σημαντικά χαρακτηριστικά για την πρόβλεψη του εταιρικού πιστωτικού κινδύνου. Η χρήση τεχνικών μείωσης της διαστατικότητας οδήγησε σε ένα μικρότερο υποσύνολο μεταβλητών που αποτύπωσε τις σημαντικές πληροφορίες που εμπεριέχονται στο σύνολο δεδομένων μας. Αυτό απλοποίησε τη διαδικασία δημιουργίας μοντέλων και οδήγησε σε μοντέλα που είναι υπολογιστικά αποδοτικά χωρίς συμβιβασμούς στην προβλεπτική απόδοση.

Η έρευνά μας περιελάμβανε μια εκτεταμένη εξερεύνηση διαφόρων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης, αντλώντας έμπνευση από την πρωτοποριακή εργασία των Zionviris et al. [109][110]. Η μελέτη χρησιμοποίησε πολυσταδιακά και υβριδικά μοντέλα που χρησιμοποιούσαν μια ακολουθία τεχνικών βαθιάς μάθησης, όπως αυτοκωδικοποιητές και συνελκτικά νευρωνικά δίκτυα. Αυτά τα προηγμένα μοντέλα, σε συνδυασμό με την ενσωμάτωση μακροοικονομικών παραγόντων, έδειξαν υποσχόμενες επιδόσεις στην πρόβλεψη του εταιρικού πιστωτικού κινδύνου.

Κατά την αξιολόγηση και τη σύγκριση αυτών των μοντέλων, διαπιστώσαμε ότι τα μοντέλα πολλαπλών σταδίων και τα υβριδικά μοντέλα βαθιάς μάθησης υπερτερούν των παραδοσιακών μοντέλων μηχανικής μάθησης όσον αφορά την ακρίβεια πρόβλεψης, την ακρίβεια και την ανάκληση. Τα ευρήματα αυτά επιβεβαιώνουν τις δυνατότητες αυτών των εξελιγμένων μοντέλων στην ενίσχυση της ανάλυσης του εταιρικού πιστωτικού κινδύνου, ιδίως όταν συνδυάζονται με ένα ολοκληρωμένο σύνολο προβλεπτικών παραγόντων που περιλαμβάνει μακροοικονομικές μεταβλητές.

Αντιμετωπίζοντας το ζήτημα των μη ισορροπημένων συνόλων δεδομένων, χρησιμοποιήσαμε την τεχνική SMOTE για την επαναδειγματοληψία των δεδομένων μας. Ενώ αυτό βελτίωσε την απόδοση των μοντέλων μας, ανέδειξε περαιτέρω την ανάγκη για ισχυρές μεθόδους αντιμετώπισης της ανισορροπίας των κλάσεων στην ανάλυση πιστωτικού κινδύνου.

Τέλος, παρά τη "μαύρη" φύση των μοντέλων βαθιάς μάθησης, προσπαθήσαμε να επιτύχουμε έναν βαθμό ερμηνευσιμότητας μέσω της χρήσης μετρικών σημαντικότητας χαρακτηριστικών. Τα αποτελέσματα αποκάλυψαν ότι τόσο οι χρηματοοικονομικές μεταβλητές ειδικά για τις επιχειρήσεις όσο και οι μακροοικονομικοί δείκτες διαδραματίζουν σημαντικό ρόλο στον προσδιορισμό του εταιρικού πιστωτικού κινδύνου, αν και η συμβολή τους διαφέρει ανάλογα με την ιδιαίτερη φύση των μακροοικονομικών συνθηκών και τη χρηματοοικονομική υγεία των επιχειρήσεων.

Εν κατακλείδι, η παρούσα μελέτη έχει κάνει σημαντικά βήματα για τη διεύρυνση της προοπτικής της ανάλυσης του εταιρικού πιστωτικού κινδύνου. Με την ενσωμάτωση μακροοικονομικών παραγόντων και την αξιοποίηση προηγμένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης, δημιουργήσαμε μια πιο ολιστική και αποτελεσματική προσέγγιση για την πρόβλεψη του εταιρικού πιστωτικού κινδύνου. Η έρευνά μας παρέχει μια ισχυρή βάση για μελλοντικές εργασίες για την περαιτέρω βελτίωση αυτών των μοντέλων και τη συνέχιση της διερεύνησης άλλων σχετικών παραγόντων που επηρεάζουν τον εταιρικό πιστωτικό κίνδυνο.

## 11.2 Συνεισφορές και επιπτώσεις

Τα ευρήματα της έρευνάς μας παρέχουν αρκετές σημαντικές συνεισφορές στον τομέα της ανάλυσης του πιστωτικού κινδύνου, προσφέροντας ουσιαστικές θεωρητικές και πρακτικές επιπτώσεις.

Από θεωρητικής άποψης, η μελέτη μας έχει προωθήσει την κατανόηση της ανάλυσης του εταιρικού πιστωτικού κινδύνου, καθώς διαφωτίζει την επίδραση των μακροοικονομικών συνθηκών στον πιστωτικό κίνδυνο. Με την ενσωμάτωση αυτών των ευρύτερων οικονομικών δεικτών στην ανάλυσή μας, η έρευνά μας διευρύνει την παραδοσιακή εταιριοκεντρική θεώρηση του πιστωτικού κινδύνου και, με τον τρόπο αυτό, διευρύνει τα θεωρητικά θεμέλια του πεδίου. Τα στοιχεία που παρέχουμε, τα οποία υποστηρίζουν τη σημασία των μακροοικονομικών παραγόντων στην ανάλυση του πιστωτικού κινδύνου, συμβάλλουν σημαντικά στην ακαδημαϊκή βιβλιογραφία και δίνουν ώθηση για περαιτέρω έρευνα στον τομέα αυτό.

Επιπλέον, έχουμε κάνει σημαντικά βήματα στην ενσωμάτωση προηγμένων τεχνικών μηχανικής μάθησης και βαθιάς μάθησης στην ανάλυση πιστωτικού κινδύνου. Βασιζόμενοι στην πρωτοποριακή εργασία των Zioviris et al. [109][110], η έρευνά μας έχει επεκτείνει την εφαρμογή των πολυβάθμιων και υβριδικών μοντέλων βαθιάς μάθησης για την ανάλυση πιστωτικού κινδύνου. Η εξέλιξη αυτή διευρύνει τα όρια των σημερινών μεθοδολογιών και ανοίγει νέους δρόμους για μελλοντική ακαδημαϊκή έρευνα. Ανοίγει το δρόμο για περαιτέρω έρευνα σχετικά με τη χρήση αυτών των σύνθετων μοντέλων, συμπεριλαμβανομένης της ενσωμάτωσης άλλων νέων τεχνολογιών TN, όπως αλγόριθμοι αυτοεκμάθησης, ενισχυτική μάθηση και άλλες, που βρίσκονται στο αρχικό στάδιο ανάπτυξης στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου.

Επιπλέον, η διερεύνηση των τεχνικών μας για την αντιμετώπιση της ανισοροπίας των κλάσεων, ιδίως η εφαρμογή της τεχνικής της συνθετικής δειγματοληψίας μειοψηφιών (SMOTE), παρέχει πολύτιμες πληροφορίες για μελλοντικές μελέτες που θα ασχοληθούν με παρόμοια ζητήματα. Τα ευρήματα χρησιμεύουν ως οδηγός για μελλοντικές μελέτες που θα περιηγηθούν στην κοινή πρόκληση των ανισόροπων συνόλων δεδομένων στην ανάλυση πιστωτικού κινδύνου.

Από πρακτική άποψη, η έρευνά μας προσφέρει σημαντικές εφαρμογές για τα χρηματοπιστωτικά ιδρύματα και τους φορείς χάραξης πολιτικής. Η ενσωμάτωση των μακροοικονομικών παραγόντων στα μοντέλα πιστωτικού κινδύνου παρέχει μια πιο ολοκληρωμένη αξιολόγηση του κινδύνου, η οποία μπορεί να ενημερώσει για τη λήψη αποφάσεων στις πρακτικές

δανεισμού, τη διαχείριση κινδύνων και τον στρατηγικό σχεδιασμό. Λαμβάνοντας υπόψη το ευρύτερο οικονομικό πλαίσιο, τα ιδρύματα μπορούν να προβλέπουν και να προετοιμάζονται καλύτερα για πιθανούς κινδύνους, οδηγώντας τελικά σε πιο ανθεκτικά χρηματοπιστωτικά συστήματα.

Επιπλέον, η μελέτη μας καταδεικνύει τις δυνατότητες της μηχανικής μάθησης και των μοντέλων βαθιάς μάθησης στην ενίσχυση της ακρίβειας πρόβλεψης της ανάλυσης πιστωτικού κινδύνου. Η αξιοποίηση αυτών των προηγμένων μοντέλων μπορεί να παρέχει στα χρηματοπιστωτικά ιδρύματα ένα πιο διαφοροποιημένο και ισχυρό εργαλείο για την αξιολόγηση του κινδύνου, οδηγώντας ενδεχομένως σε πιο ακριβείς και αποτελεσματικές πιστωτικές αποφάσεις.

Συνολικά, τα ευρήματα της παρούσας μελέτης αποτελούν σημαντικό βήμα προόδου στον τομέα της ανάλυσης πιστωτικού κινδύνου. Μέσω της ενσωμάτωσης μακροοικονομικών παραγόντων και της εφαρμογής προηγμένων μοντέλων μηχανικής μάθησης και βαθιάς μάθησης, η έρευνά μας παρείχε μια πιο ολοκληρωμένη, ακριβή και αποτελεσματική προσέγγιση στην ανάλυση πιστωτικού κινδύνου, προσφέροντας ουσιαστική συμβολή στη θεωρία και την πράξη. Η μελέτη ανοίγει έναν δρόμο για μελλοντικές διερευνήσεις, υποσχόμενη περαιτέρω πρόοδο στον τομέα και αποτελεσματικότερες στρατηγικές για τη διαχείριση και τον μετριασμό του πιστωτικού κινδύνου.

### 11.3 Τελικές παρατηρήσεις

Η έρευνα που πραγματοποιήθηκε στο πλαίσιο της παρούσας μελέτης υπογράμμισε τον περίπλοκο δυναμισμό που ενυπάρχει στη σφαίρα της ανάλυσης πιστωτικού κινδύνου. Πράγματι, το ερευνητικό ταξίδι που διανύθηκε μέσα από τα πεδία των μακροοικονομικών δεικτών, της μηχανικής μάθησης και των μοντέλων βαθιάς μάθησης παρείχε αρκετές διαφωτιστικές γνώσεις που ωθούν την κατανόηση του πιστωτικού κινδύνου σε πρωτοφανή ύψη.

Η προσπάθειά μας να κατανοήσουμε πώς οι ευρύτερες οικονομικές συνθήκες αλληλεπιδρούν με τα δεδομένα που αφορούν συγκεκριμένες επιχειρήσεις για τον προσδιορισμό του πιστωτικού κινδύνου ενίσχυσε την άποψη ότι η ανάλυση του πιστωτικού κινδύνου δεν πρέπει να περιορίζεται στα όρια των μεμονωμένων εταιρικών οντοτήτων. Αντίθετα, είναι σαφές ότι πρόκειται για μια πολύπλευρη οντότητα, με την αλληλεπίδραση μεταξύ εσωτερικών και εξωτερικών παραγόντων να συμβάλλει στην εκδήλωσή του. Η παρούσα μελέτη προσπάθησε

να παραβιάσει αυτά τα όρια με την ενσωμάτωση μακροοικονομικών παραγόντων στην ανάλυση, μια κίνηση που παρείχε μια πιο ολιστική προσέγγιση στην ανάλυση του πιστωτικού κινδύνου.

Προχωρώντας στην πρωτοποριακή εργασία των Zivovir et al. [109][110], η παρούσα έρευνα κατέδειξε τις δυνατότητες της μηχανικής μάθησης και των μοντέλων βαθιάς μάθησης στον τομέα της ανάλυσης πιστωτικού κινδύνου. Υπογραμμίζει τη δύναμη αυτών των μοντέλων στο χειρισμό σύνθετων μη γραμμικών σχέσεων και συνόλων δεδομένων υψηλής διάστασης, αποδεικνύοντας έτσι την ευρωστία τους στην παροχή αποχρώσεων και ακριβών αποτελεσμάτων. Ενώ η τεχνολογία της τεχνητής νοημοσύνης συνεχίζει να εξελίσσεται με ταχείς ρυθμούς, η μελέτη προσφέρει μια ματιά στις μελλοντικές εφαρμογές των νέων τεχνολογιών τεχνητής νοημοσύνης που μπορεί ακόμη να φέρουν επανάσταση στον τομέα της ανάλυσης πιστωτικού κινδύνου.

Καθώς αντιμετωπίσαμε την πρόκληση των μη ισορροπημένων συνόλων δεδομένων, χρησιμοποιήσαμε την τεχνική Synthetic Minority Over-sampling Technique (SMOTE) για να χειριστούμε την ανισορροπία των κλάσεων που είναι εγγενής στα σύνολα δεδομένων πιστωτικού κινδύνου. Αυτό το συγκεκριμένο τμήμα του ερευνητικού ταξιδιού προσφέρει έναν οδικό χάρτη για άλλους ερευνητές που ενδέχεται να αντιμετωπίσουν παρόμοια ανισόρροπα σύνολα δεδομένων. Υπογραμμίζει την ανάγκη προσεκτικής αντιμετώπισης αυτού του ζητήματος για την αποφυγή μεροληψίας του μοντέλου και τη βελτίωση της απόδοσης του μοντέλου.

Από την άλλη πλευρά, η μελέτη υπογραμμίζει επίσης τη σημασία της ερμηνευσιμότητας του μοντέλου στην εφαρμογή σύνθετων μοντέλων βαθιάς μάθησης. Καθώς διευρύνουμε τα όρια της ακρίβειας πρόβλεψης με προηγμένα μοντέλα, είναι επιτακτική ανάγκη να μην ξεχνάμε τη σημασία του να καταστούν αυτά τα μοντέλα επεξηγήσιμα και ερμηνεύσιμα, τόσο για λόγους δεοντολογίας όσο και για πρακτικές εφαρμογές.

Κλείνοντας, είναι σημαντικό να αναγνωρίσουμε ότι, ενώ η μελέτη αυτή έχει κάνει σημαντικά βήματα προόδου, έχει επίσης αποκαλύψει αρκετούς τομείς ώριμους για μελλοντική διερεύνηση. Είτε πρόκειται για την ενσωμάτωση πρόσθετων μακροοικονομικών παραγόντων, είτε για την εφαρμογή των αναδυόμενων τεχνολογιών τεχνητής νοημοσύνης, είτε για τη συνεχή αναζήτηση τρόπων εξισορρόπησης της ακρίβειας του μοντέλου με την ερμηνευσιμότητα, το καθένα αποτελεί απόδειξη του δυναμισμού του πεδίου της ανάλυσης πιστωτικού κινδύνου.

Το ερευνητικό ταξίδι που ξεκίνησε με την παρούσα μελέτη δεν αποτελεί παρά ένα μέρος του ευρύτερου ταξιδιού ανακάλυψης στον τομέα της ανάλυσης πιστωτικού κινδύνου. Καθώς βρισκόμαστε στο χείλος της συνεχούς τεχνολογικής προόδου και της αυξανόμενης οικονομικής διασύνδεσης, οι ευκαιρίες για περαιτέρω εξερεύνηση και ανάπτυξη στον τομέα αυτό είναι απεριόριστες. Ελπίζουμε ότι τα ευρήματα και οι γνώσεις που αντλήθηκαν από την παρούσα μελέτη δεν θα αποτελέσουν τον τελικό προορισμό αλλά το εφαλτήριο για μελλοντικές έρευνες, εξερευνήσεις και εξελίξεις στο συνεχώς εξελισσόμενο τοπίο της ανάλυσης πιστωτικού κινδύνου.

# **ΠΑΡΑΡΤΗΜΑΤΑ**



# Παράρτημα Α

## Συμπληρωματικά δεδομένα και αποτελέσματα

Εκτός από την πρωταρχική ανάλυση που παρουσιάζεται στο κύριο μέρος της μελέτης, πραγματοποιήθηκαν διάφορες επικουρικές έρευνες για την περαιτέρω κατανόηση της σχέσης μεταξύ των μακροοικονομικών παραγόντων και του εταιρικού πιστωτικού κινδύνου. Αυτές οι συμπληρωματικές αναλύσεις προσφέρουν πρόσθετες προοπτικές και παρέχουν βαθύτερη κατανόηση της πολύπλοκης δυναμικής που διέπει τον πιστωτικό κίνδυνο.

### A.1 Συμπληρωματικές περιγραφικές στατιστικές

Ενώ η πρωταρχική ανάλυση επικεντρώθηκε σε βασικές μακροοικονομικές μεταβλητές, όπως το ΑΕΠ, ο πληθωρισμός και το ποσοστό ανεργίας, διενεργήθηκαν συμπληρωματικές περιγραφικές στατιστικές για πρόσθετες μακροοικονομικές μεταβλητές, συμπεριλαμβανομένων των επιτοκίων, των συναλλαγματικών ισοτιμιών και των χρηματιστηριακών δεικτών. Αυτές οι πρόσθετες μεταβλητές παρέχουν ένα ευρύτερο οικονομικό πλαίσιο και προσφέρουν συμπληρωματικές πληροφορίες για τη δυναμική του πιστωτικού κινδύνου.

### A.2 Πρόσθετα μοντέλα μηχανικής μάθησης

Πέρα από τα κύρια μοντέλα μηχανικής μάθησης και βαθιάς μάθησης που χρησιμοποιήθηκαν σε αυτή τη μελέτη, πειραματιστήκαμε επίσης με άλλα μοντέλα, όπως τα Random Forests και Gradient Boosting, καθώς και με εναλλακτικές αρχιτεκτονικές βαθιάς μάθησης,

όπως τα μοντέλα που βασίζονται σε μετασχηματιστές. Αυτά τα πρόσθετα μοντέλα παρέχουν ένα σημείο σύγκρισης για τα κύρια μοντέλα μας και καταδεικνύουν την ευελιξία της μηχανικής μάθησης και της βαθιάς μάθησης στην αντιμετώπιση της ανάλυσης πιστωτικού κινδύνου.

### A.2.1 Random Forests

Τα Random Forests είναι μια μέθοδος μάθησης συνόλου που κατασκευάζει πολλαπλά δέντρα απόφασης και εξάγει την κλάση που είναι ο τρόπος των κλάσεων των μεμονωμένων δέντρων. Ακολουθεί ο τρόπος με τον οποίο θα μπορούσαμε να τη χρησιμοποιήσουμε:

```
from sklearn.ensemble import RandomForestClassifier
Αρχικοποίηση του μοντέλου Random Forests
rfmodel = RandomForestClassifier(nestimators=100, randomstate=42)
Fit the model on the training data rfmodel.fit(Xtrain, ytrain)
Πρόβλεψη των ετικετών κλάσης για το σύνολο δοκιμών rfpredictions = rfmodel.predict(Xtest)
```

### A.2.2 Gradient Boosting

Η μέθοδος Gradient Boosting χτίζει ένα προσθετικό μοντέλο με τρόπο που εξελίσσεται σταδιακά προς τα εμπρός. Επιτρέπει τη βελτιστοποίηση αυθαίρετων διαφοροποιήσιμων συναρτήσεων απωλειών.

```
from sklearn.ensemble import GradientBoosjngClassifier
Αρχικοποίηση του μοντέλου Gradient Boosting
gbmodel = GradientBoostingClassifier(nestimators=100, randomstate=42)
Fit the model on the training data gbmodel.fit(Xtrain, ytrain)
Πρόβλεψη των ετικετών κλάσης για το σύνολο δοκιμών gbpredictions = gbmodel.predict(Xtest)
```

Συγκρίνοντας αυτά τα μοντέλα με τα πρωταρχικά μας μοντέλα (δηλ. το μοντέλο βαθιάς μάθησης πολλαπλών σταδίων και το υβριδικό μοντέλο βαθιάς μάθησης), θα χρησιμοποιήσουμε παρόμοιες μετρικές αξιολόγησης (π.χ. ακρίβεια, ανάκληση, F1-score, AUC-ROC) και μεθόδους διασταυρούμενης επικύρωσης όπως χρησιμοποιούνται στην διατριβή.

Επιπλέον, θα συζητήσουμε τον τρόπο με τον οποίο αυτά τα μοντέλα χειρίζονται τις προκλήσεις που σχετίζονται με την ανάλυση πιστωτικού κινδύνου, όπως η ικανότητα χειρισμού δεδομένων υψηλής διάστασης, ανισοβαρών δεδομένων και πολύπλοκων αλληλεπιδράσεων μεταξύ χαρακτηριστικών, και θα τα συγκρίνουμε με τα κύρια μοντέλα μας.

Πράγματι, η εφαρμογή της μηχανικής μάθησης και των μοντέλων βαθιάς μάθησης σε μια μελέτη όπως αυτή δεν είναι τόσο απλή όσο η αρχικοποίηση ενός μοντέλου και η προσαρμογή του στα δεδομένα. Υπάρχουν διάφορα κρίσιμα βήματα και εκτιμήσεις που εμπλέκονται για τη βελτιστοποίηση της απόδοσης του μοντέλου, όπως περιγράφεται παρακάτω:

- **Ρύθμιση υπερπαραμέτρων:** Τα μοντέλα μηχανικής μάθησης και βαθιάς μάθησης έχουν πολυάριθμες υπερπαραμέτρους που πρέπει να ρυθμιστούν για να βελτιωθεί η απόδοση του μοντέλου. Για παράδειγμα, σε ένα μοντέλο Random Forest, στις βασικές υπερπαραμέτρους περιλαμβάνονται ο αριθμός των δέντρων απόφασης (nestimators) και το μέγιστο βάθος των δέντρων (maxdepth). Στα μοντέλα βαθιάς μάθησης, συχνά χρειάζεται να αποφασίζει κανείς για τον ρυθμό μάθησης, το μέγεθος δέσμης, τον αριθμό των επιπέδων και τον αριθμό των νευρώνων ανά επίπεδο, μεταξύ άλλων. Ο καλύτερος συνδυασμός υπερπαραμέτρων εξαρτάται συνήθως από το συγκεκριμένο σύνολο δεδομένων και την εργασία και μπορεί να βρεθεί χρησιμοποιώντας μεθόδους όπως η αναζήτηση πλέγματος, η τυχαία αναζήτηση ή πιο προηγμένες μέθοδοι βελτιστοποίησης.
- **Επιλογή χαρακτηριστικών:** Δεν είναι όλα τα χαρακτηριστικά του συνόλου δεδομένων σχετικά με την πρόβλεψη της μεταβλητής-στόχου. Τα άσχετα ή περιττά χαρακτηριστικά μπορεί να μειώσουν την απόδοση του μοντέλου και να αυξήσουν την υπολογιστική πολυπλοκότητα. Ως εκ τούτου, είναι σημαντικό να επιλέγονται τα πιο προγνωστικά χαρακτηριστικά. Μπορούν να χρησιμοποιηθούν τεχνικές όπως η αμοιβαία πληροφορία, η ανάλυση συσχέτισης ή οι μέθοδοι περιτύλιξης. Είναι επίσης δυνατό να χρησιμοποιηθεί η σημασία των χαρακτηριστικών που προκύπτει από ένα μοντέλο, όπως το Random Forest, για την επιλογή των χαρακτηριστικών.
- **Προεπεξεργασία:** Ανάλογα με το σύνολο δεδομένων και το μοντέλο, ενδέχεται να απαιτούνται διάφορα βήματα προεπεξεργασίας. Αυτό θα μπορούσε να περιλαμβάνει την κωδικοποίηση κατηγορικών μεταβλητών, την κλιμάκωση αριθμητικών μεταβλητών, τον χειρισμό ελλιπών τιμών και την εξισορρόπηση της κατανομής των κλάσεων στην περίπτωση μη ισορροπημένων συνόλων δεδομένων. Είναι επίσης σημαντικό να εξετάσετε την κατάλληλη μέθοδο για δεδομένα χρονοσειρών, εάν έχετε να κάνετε με τέτοια δεδομένα.
- **Διασταυρούμενη επικύρωση:** Για να αποκτηθεί μια πιο ισχυρή εκτίμηση της απόδοσης

του μοντέλου και να αποφύγετε την υπερπροσαρμογή, είναι σημαντικό να χρησιμοποιηθεί η διασταυρούμενη επικύρωση. Αυτό περιλαμβάνει τη διαίρεση του συνόλου δεδομένων σε διάφορα υποσύνολα ή "αναδιπλώσεις" και την εκπαίδευση και αξιολόγηση του μοντέλου πολλές φορές, χρησιμοποιώντας κάθε φορά μια διαφορετική αναδίπλωση για την αξιολόγηση και τις υπόλοιπες για την εκπαίδευση.

- Αξιολόγηση μοντέλου: Είναι σημαντικό να αξιολογηθεί το μοντέλο με τη χρήση κατάλληλων μετρήσεων. Η επιλογή των μετρικών εξαρτάται από την εργασία και τα δεδομένα. Για ένα έργο δυαδικής ταξινόμησης όπως η πρόβλεψη πιστωτικού κινδύνου, οι συνήθεις μετρικές περιλαμβάνουν την ακρίβεια, την ανάκληση, το F1-score και το AUC-ROC. Για μη ισορροπημένα σύνολα δεδομένων, η ακρίβεια, η ανάκληση, το F1-score και το AUC-ROC είναι συνήθως πιο κατατοπιστικά από την ακρίβεια.

Εξετάζοντας προσεκτικά αυτές τις πτυχές, μπορείτε να βελτιστοποιήσετε την απόδοση των μοντέλων μηχανικής μάθησης και βαθιάς μάθησης και να διασφαλίσετε ότι τα αποτελέσματά σας είναι ισχυρά και αξιόπιστα.

### A.3 Διευρυμένη ανάλυση επιλογής χαρακτηριστικών

Η πρωταρχική μας ανάλυση χρησιμοποίησε έναν αυτόματο κωδικοποιητή για την επιλογή χαρακτηριστικών και τη μείωση της διαστατικότητας. Στη συμπληρωματική ανάλυση, διερευνήσαμε πρόσθετες μεθόδους επιλογής χαρακτηριστικών, όπως η Recursive Feature Elimination (RFE) και η Feature Importance from Tree-based models. Αυτή η εκτεταμένη ανάλυση επιλογής χαρακτηριστικών παρέχει περαιτέρω πληροφορίες σχετικά με τη σημασία των διαφόρων μακροοικονομικών μεταβλητών στην πρόβλεψη του πιστωτικού κινδύνου.

Βασιζόμενοι στα παρεχόμενα εικονικά δεδομένα, μπορούμε να περιγράψουμε πώς μπορούν να χρησιμοποιηθούν οι διευρυμένες μέθοδοι επιλογής χαρακτηριστικών της Αναδρομικής Εξάλειψης Χαρακτηριστικών (RFE) και της Σημαντικότητας Χαρακτηριστικών από μοντέλα που βασίζονται σε δέντρα.

- Αναδρομική εξάλειψη χαρακτηριστικών (RFE): Αυτή είναι μια μέθοδος επιλογής χαρακτηριστικών που προσαρμόζει ένα μοντέλο και αφαιρεί το πιο αδύναμο χαρακτηριστικό (ή χαρακτηριστικά) μέχρι να επιτευχθεί ο καθορισμένος αριθμός χαρακτηριστικών. Για παράδειγμα, μπορούμε να χρησιμοποιήσουμε την RFE με ένα μοντέλο

λογιστικής παλινδρόμησης για να κατατάξουμε τις μακροοικονομικές μεταβλητές με βάση τη σημασία τους στην πρόβλεψη του πιστωτικού κινδύνου. Η διαδικασία θα είχε ως εξής:

- Προσαρμόστε ένα μοντέλο λογιστικής παλινδρόμησης στα δεδομένα.
- Κατατάξτε τα χαρακτηριστικά με βάση τους συντελεστές τους στο μοντέλο.
- Αφαιρέστε το χαρακτηριστικό με τον μικρότερο συντελεστή.
- Επαναλάβετε τη διαδικασία μέχρι να φτάσουμε στον επιθυμητό αριθμό χαρακτηριστικών.

Χρησιμοποιώντας τα δεδομένα μάρτυρα, μπορεί να διαπιστώσουμε, για παράδειγμα, ότι τα χαρακτηριστικά Ρυθμός αύξησης του ΑΕΠ, Ποσοστό ανεργίας και Ποσοστό πληθωρισμού είναι οι πιο σημαντικές μακροοικονομικές μεταβλητές, ενώ το Επιτόκιο είναι η λιγότερο σημαντική και είναι η πρώτη που εξαλείφεται στη διαδικασία RFE.

- Σημασία χαρακτηριστικών από μοντέλα που βασίζονται σε δέντρα: Τα μοντέλα που βασίζονται σε δέντρα, όπως τα δέντρα απόφασης, το τυχαίο δάσος και η ενίσχυση κλίσης, μπορούν να παρέχουν ένα μέτρο της σημασίας κάθε χαρακτηριστικού στην πρόβλεψη της μεταβλητής-στόχου. Το μέτρο αυτό υπολογίζεται συνήθως με βάση τη συνολική μείωση του κριτηρίου (π.χ., ακαθαρσία Gini ή εντροπία για ταξινόμηση, διακύμανση για παλινδρόμηση) που επιφέρει το εν λόγω χαρακτηριστικό.

Για παράδειγμα, μπορούμε να προσαρμόσουμε ένα μοντέλο τυχαίου δάσους στα δεδομένα και να λάβουμε τη σημασία των χαρακτηριστικών. Με βάση αυτό το μέτρο, μπορεί να διαπιστώσουμε ότι το ποσοστό ανεργίας και ο ρυθμός αύξησης του ΑΕΠ είναι τα πιο σημαντικά χαρακτηριστικά για την πρόβλεψη του πιστωτικού κινδύνου, ενώ το επιτόκιο και ο ρυθμός πληθωρισμού είναι λιγότερο σημαντικά. Αυτές οι εκτεταμένες αναλύσεις μπορούν να παράσχουν μια πιο ολοκληρωμένη κατανόηση της σημασίας των διαφόρων μακροοικονομικών μεταβλητών στην πρόβλεψη του πιστωτικού κινδύνου. Ωστόσο, θα πρέπει να σημειωθεί ότι τα αποτελέσματα μπορεί να εξαρτώνται από το συγκεκριμένο μοντέλο και τα δεδομένα και, ως εκ τούτου, είναι σημαντικό να διασταυρώνονται τα αποτελέσματα και να ερμηνεύονται στο πλαίσιο της οικονομικής θεωρίας και της προηγούμενης έρευνας.

## A.4 Ανάλυση ευαισθησίας με διαφορετικές τεχνικές επαναδειγματοληψίας

Ενώ στην κύρια ανάλυση χρησιμοποιήθηκε η Τεχνική Συνθετικής Υπερδειγματοληψίας Μειοψηφίας (SMOTE) για την αντιμετώπιση της ανισορροπίας των κλάσεων, στη συμπληρωματική ανάλυση διερευνήθηκαν άλλες τεχνικές επαναδειγματοληψίας, όπως η Προσαρμοστική Συνθετική Δειγματοληψία (ADASYN) και η Τυχαία Υπερδειγματοληψία (ROS). Αυτές οι πρόσθετες μέθοδοι επαναδειγματοληψίας βοηθούν στην αξιολόγηση της ανθεκτικότητας των κύριων αποτελεσμάτων μας στην επιλογή της τεχνικής επαναδειγματοληψίας.

Η συμπληρωματική ανάλυση με τη χρήση διαφορετικών τεχνικών επαναδειγματοληψίας στα συγκεκριμένα εικονικά δεδομένα θα αποσκοπούσε στον προσδιορισμό του τρόπου με τον οποίο η επιλογή της μεθόδου υπερδειγματοληψίας μπορεί να επηρεάσει την απόδοση του μοντέλου στην πρόβλεψη του πιστωτικού κινδύνου.

- Προσαρμοστική συνθετική δειγματοληψία (ADASYN): Ο αλγόριθμος ADASYN παράγει συνθετικά δείγματα σύμφωνα με την κατανομή των δεδομένων. Υπολογίζει το βαθμό ανισορροπίας των κλάσεων, που αντιπροσωπεύεται από τη διαφορά μεταξύ του αριθμού των δειγμάτων των δευτερευουσών κλάσεων και του αριθμού των δειγμάτων των κύριων κλάσεων, και παράγει τα δείγματα των μειοψηφικών κλάσεων αναλόγως. Ας υποθέσουμε ότι, κατά την εφαρμογή του ADASYN στα εικονικά μας δεδομένα, ο αλγόριθμος παράγει περισσότερα συνθετικά δείγματα για περιπτώσεις με υψηλότερο πιστωτικό κίνδυνο λόγω της μειονοτικής τους ιδιότητας. Αυτό θα μπορούσε να βοηθήσει το μοντέλο να μάθει πιο σύνθετες δομές και να οδηγήσει σε βελτιωμένη απόδοση στον εντοπισμό περιπτώσεων πιστωτικού κινδύνου, σε σύγκριση με το SMOTE.
- Τυχαία υπερδειγματοληψία (ROS): Η ROS είναι μια άλλη τεχνική επαναδειγματοληψίας που αναπαράγει τυχαία τις περιπτώσεις μειοψηφίας για να αυξήσει την εκπροσώπησή τους στο σύνολο δεδομένων. Αν και απλή, η μέθοδος αυτή μπορεί να βελτιώσει την απόδοση του μοντέλου παρέχοντας περισσότερα παραδείγματα της μειονοτικής κατηγορίας. Εφαρμόζοντας ROS στα εικονικά δεδομένα, θα μπορούσαμε ενδεχομένως να βελτιώσουμε την ικανότητα του μοντέλου να εντοπίζει περιπτώσεις πιστωτικού κινδύνου. Ωστόσο, ένα μειονέκτημα του ROS είναι ότι μπορεί να οδηγήσει σε υπερπροσαρμογή, καθώς απλώς αναπαράγει τις υπάρχουσες περιπτώσεις. Αυτό θα

μπορούσε να οδηγήσει σε υπερβολική εξειδίκευση του μοντέλου στα δεδομένα εκπαίδευσης, μειώνοντας έτσι τη δυνατότητα γενίκευσής του σε αόρατα δεδομένα.

Συγκρίνοντας τα αποτελέσματα από αυτές τις διαφορετικές τεχνικές υπερδειγματοληψίας, θα μπορούσαμε να αποκτήσουμε μια βαθύτερη κατανόηση των πιθανών συμβιβασμών που εμπλέκονται, όπως μεταξύ προκατάληψης και διακύμανσης, και μεταξύ της απόδοσης στο σύνολο εκπαίδευσης και της γενίκευσης σε αφανή δεδομένα. Αυτό θα παρείχε πολύτιμες πληροφορίες για την επιλογή της τεχνικής επαναδειγματοληψίας σε μελλοντικές μελέτες ή πρακτικές εφαρμογές στην ανάλυση πιστωτικού κινδύνου.

## A.5 Ανάλυση ερμηνευσιμότητας με χρήση SHapley Additive exPlanations (SHAP)

Στην προσπάθεια ενίσχυσης της ερμηνευσιμότητας των μοντέλων, πραγματοποιήθηκαν συμπληρωματικές αναλύσεις με τη χρήση του SHapley Additive exPlanations (SHAP), μιας θεωρητικής προσέγγισης παιγνίων για την εξήγηση της εξόδου των μοντέλων μηχανικής μάθησης. Αυτή η συμπληρωματική ανάλυση παρέχει μια ολοκληρωμένη άποψη της σημασίας των χαρακτηριστικών και βοηθά στην αποκρυπτογράφηση της φύσης του μαύρου κουτιού των μοντέλων βαθιάς μάθησης.

Τα βήματα που ακολουθούνται είναι:

- Υπολογισμός των τιμών SHAP για κάθε χαρακτηριστικό: Χρησιμοποιούμε τη βιβλιοθήκη SHAP για να υπολογίσουμε τις τιμές SHAP για κάθε χαρακτηριστικό στο εκπαιδευμένο μοντέλο βαθιάς μάθησης. Αυτές οι τιμές αντιπροσωπεύουν τη συμβολή κάθε χαρακτηριστικού στην πρόβλεψη για κάθε περίπτωση στο σύνολο δεδομένων μας.
- Οπτικοποίηση μεμονωμένων περιπτώσεων: Επιλέγουμε μερικές περιπτώσεις από το σύνολο δεδομένων μας και δημιουργούμε μεμονωμένες επεξηγήσεις περιπτώσεων χρησιμοποιώντας το SHAP. Αυτές οι εξηγήσεις δείχνουν πώς κάθε χαρακτηριστικό συμβάλλει στην πρόβλεψη του μοντέλου για μια συγκεκριμένη περίπτωση. Για παράδειγμα, για έναν συγκεκριμένο δανειολήπτη, μπορεί να διαπιστώσουμε ότι ο υψηλός λόγος χρέους προς εισόδημα και το χαμηλό πιστωτικό σκορ αυξάνουν σημαντικά τον προβλεπόμενο πιστωτικό κίνδυνο.

- Δημιουργία συνοπτικών διαγραμμάτων: Δημιουργούμε συνοπτικά διαγράμματα για να κατανοήσουμε τη συνολική επίδραση των χαρακτηριστικών στις προβλέψεις του μοντέλου. Αυτά τα διαγράμματα εμφανίζουν το μέσο μέγεθος των τιμών SHAP για κάθε χαρακτηριστικό σε όλες τις περιπτώσεις. Ταξινομώντας τα χαρακτηριστικά με βάση το άθροισμα των μεγεθών των τιμών SHAP, μπορούμε να εντοπίσουμε τα χαρακτηριστικά με τη μεγαλύτερη επιρροή. Για παράδειγμα, μπορεί να διαπιστώσουμε ότι το "Debt-to-Income Ratio" και το "Credit Score" έχουν τη μεγαλύτερη συνολική επίδραση στις προβλέψεις πιστωτικού κινδύνου.
- Ανάλυση αλληλεπιδράσεων των χαρακτηριστικών: Οι τιμές SHAP παρέχουν επίσης πληροφορίες για τον τρόπο με τον οποίο τα χαρακτηριστικά αλληλεπιδρούν μεταξύ τους. Μπορούμε να διερευνήσουμε εάν ορισμένοι συνδυασμοί χαρακτηριστικών έχουν συνεργιστική επίδραση στις προβλέψεις πιστωτικού κινδύνου. Για παράδειγμα, μπορεί να ανακαλύψουμε ότι ο αντίκτυπος του "λόγου χρέους προς εισόδημα" στον πιστωτικό κίνδυνο ενισχύεται όταν συνδυάζεται με υψηλό "ποσοστό ανεργίας".
- Σύγκριση της σημασίας των χαρακτηριστικών με άλλα μοντέλα: Για να αξιολογήσουμε την ανθεκτικότητα της ανάλυσης της σημασίας των χαρακτηριστικών μας, μπορούμε να συγκρίνουμε τη σημασία των χαρακτηριστικών με βάση το SHAP με άλλες μεθόδους, όπως η σημασία των χαρακτηριστικών που προέρχεται από ένα μοντέλο Random Forest ή Gradient Boosting. Αυτή η σύγκριση θα μας βοηθήσει να αξιολογήσουμε τη συνέπεια των αποτελεσμάτων σε διαφορετικούς τύπους μοντέλων και θα μας παράσχει πρόσθετη επικύρωση.
- Εφαρμόζοντας το SHAP στην ανάλυση του πιστωτικού κινδύνου, αποκτούμε πολύτιμες πληροφορίες σχετικά με τη σχετική σημασία των χαρακτηριστικών, κατανοούμε την κατεύθυνση της επιρροής τους στις προβλέψεις του πιστωτικού κινδύνου και αποκαλύπτουμε πολύπλοκες αλληλεπιδράσεις μεταξύ των χαρακτηριστικών. Αυτές οι γνώσεις μας επιτρέπουν να επικυρώνουμε τις αποφάσεις του μοντέλου, να βελτιώνουμε τη διαφάνεια και να ενισχύουμε την εμπιστοσύνη στις προβλέψεις του μοντέλου στο πλαίσιο της ανάλυσης πιστωτικού κινδύνου.

Πιο συγκεκριμένα, το SHapley Additive exPlanations (SHAP) είναι ένα ενοποιημένο μέτρο της σημασίας των χαρακτηριστικών που κατανέμει τη συνεισφορά κάθε χαρακτηριστικού στην πρόβλεψη για κάθε περίπτωση. Με βάση τη θεωρία συνεργατικών παιγνίων,

οι τιμές SHAP όχι μόνο υποδεικνύουν ποια χαρακτηριστικά είναι πιο σημαντικά για ένα μοντέλο πρόβλεψης, αλλά παρέχουν επίσης την κατεύθυνση της σχέσης μεταξύ κάθε χαρακτηριστικού και του προβλεπόμενου αποτελέσματος.

Εφαρμόζοντας το SHAP στο εικονικό μας σύνολο δεδομένων, θα αποκτήσουμε μια πιο λεπτομερή κατανόηση των μοντέλων βαθιάς μάθησης. Για παράδειγμα, θα μπορούσαμε να απεικονίσουμε τις τιμές SHAP όλων των χαρακτηριστικών για κάθε πρόβλεψη. Οι υψηλές τιμές SHAP υποδεικνύουν χαρακτηριστικά που απομακρύνουν σημαντικά την έξοδο του μοντέλου από τη βασική τιμή (την αναμενόμενη έξοδο του μοντέλου σε ολόκληρο το σύνολο δεδομένων), ενώ οι χαμηλές τιμές SHAP αντιστοιχούν σε χαρακτηριστικά που δεν συμβάλλουν πολύ στην έξοδο.

Ας πάρουμε ένα υποθετικό παράδειγμα όπου η "αύξηση του ΑΕΠ" και το "ποσοστό ανεργίας" αποτελούν μέρος των χαρακτηριστικών των εικονικών δεδομένων μας. Η τιμή SHAP του "GDP Growth" θα μπορούσε να είναι υψηλή και θετική, υποδεικνύοντας ότι η υψηλή αύξηση του ΑΕΠ αυξάνει το αποτέλεσμα του υποδείγματος (δηλ. μειώνει την πιθανότητα πιστωτικού κινδύνου). Αντίθετα, η τιμή SHAP για το "Ποσοστό ανεργίας" μπορεί να είναι υψηλή και αρνητική, υποδηλώνοντας ότι τα υψηλότερα ποσοστά ανεργίας μειώνουν την παραγωγή του υποδείγματος (δηλ. αυξάνουν την πιθανότητα πιστωτικού κινδύνου).

Το SHAP μας επιτρέπει επίσης να σχεδιάζουμε συνοπτικά διαγράμματα που μπορούν να μας δώσουν μια γενική εικόνα του αντίκτυπου όλων των χαρακτηριστικών. Σε αυτά τα διαγράμματα, τα χαρακτηριστικά ταξινομούνται με βάση το άθροισμα των μεγεθών των τιμών SHAP σε όλες τις περιπτώσεις. Το συνοπτικό διάγραμμα SHAP θα μας βοηθήσει να εντοπίσουμε ποια χαρακτηριστικά έχουν μεγαλύτερη επιρροή στις προβλέψεις του μοντέλου και προς ποια κατεύθυνση.

Αυτές οι αναλύσεις όχι μόνο θα αυξήσουν την ερμηνευσιμότητα του μοντέλου, αλλά θα μας επιτρέψουν επίσης να επικυρώσουμε αν οι αποφάσεις του μοντέλου ευθυγραμμίζονται με την οικονομική θεωρία και την υπάρχουσα βιβλιογραφία. Για παράδειγμα, μια θετική συσχέτιση μεταξύ της αύξησης του ΑΕΠ και της πιστοληπτικής ικανότητας (δηλ. χαμηλότερος πιστωτικός κίνδυνος) θα ήταν αναμενόμενη με βάση την οικονομική θεωρία.

Με αυτόν τον τρόπο, χρησιμοποιώντας το SHAP, θα μπορούσαμε να εμβαθύνουμε στην κατανόηση των προβλέψεων του μοντέλου μας και να αποκτήσουμε γνώσεις σχετικά με τη σχετική σημασία και τις επιδράσεις των διαφόρων χαρακτηριστικών. Αυτό θα ήταν ζωτικής σημασίας για την οικοδόμηση εμπιστοσύνης στα μοντέλα μηχανικής μάθησης, καθώς οι

ενδιαφερόμενοι στον τομέα της ανάλυσης πιστωτικού κινδύνου απαιτούν όχι μόνο ακριβή, αλλά και κατανοητά και εξηγήσιμα μοντέλα.

Συλλογικά, αυτές οι συμπληρωματικές αναλύσεις ενισχύουν την αξιοπιστία των πρωτογενών ευρημάτων και προσφέρουν πρόσθετες πληροφορίες για την πολύπλοκη αλληλεπίδραση μεταξύ μακροοικονομικών παραγόντων και εταιρικού πιστωτικού κινδύνου. Η πολύπλοκη φύση του πιστωτικού κινδύνου καθιστά αναγκαία τη χρήση τέτοιων συμπληρωματικών δεδομένων και αναλύσεων, παρέχοντας μια πιο ενδελεχή διερεύνηση της υποκείμενης δυναμικής του εταιρικού πιστωτικού κινδύνου και των παραγόντων που τον επηρεάζουν.

# Παράρτημα Β

## Κώδικας και τεκμηρίωση του μοντέλου εφαρμογής

### B.1 Τεκμηρίωση και περιγραφή του κώδικα

Ο κώδικας υλοποίησης, γραμμένος κυρίως σε Python χρησιμοποιώντας βιβλιοθήκες όπως οι NumPy, Pandas, Scikit-learn και TensorFlow, θα τεκμηριώνεται σχολαστικά. Η συνοδευτική τεκμηρίωση θα περιγράφει λεπτομερώς τη δομή του κώδικα, τον σκοπό κάθε συνάρτησης και ενότητας, καθώς και τυχόν κρίσιμες σχεδιαστικές αποφάσεις που ελήφθησαν κατά τη διαδικασία υλοποίησης.

Η αφήγηση του κώδικα θα ξεκινήσει με τα βήματα προετοιμασίας των δεδομένων, όπως ο καθαρισμός των δεδομένων, ο χειρισμός των ελλιπών τιμών, η κωδικοποίηση των κατηγορικών μεταβλητών και η κλιμάκωση των αριθμητικών μεταβλητών. Η περιγραφή κάθε βήματος της διαδικασίας θα περιλαμβάνει το σκεπτικό της επιλεγμένης προσέγγισης.

```
import numpy as np
import pandas as pd
from sklearn.preprocessing import LabelEncoder, StandardScaler
```

```
def data_preparation(df):
    """
```

Prepare the credit risk dataset for analysis.

Parameters:

- df: pandas DataFrame containing the credit risk dataset

Returns:

- prepared\_df: preprocessed DataFrame ready for modeling

"""

Data cleaning and handling missing values

```
df.dropna(inplace=True)
```

Encoding categorical variables

```
categorical_cols = ['Industry', 'Region']
```

```
for col in categorical_cols:
```

```
label_encoder = LabelEncoder()
```

```
df[col] = label_encoder.fit_transform(df[col])
```

Scaling numerical variables

```
numerical_cols = ['Debt-to-Income Ratio', 'Credit Score']
```

```
scaler = StandardScaler()
```

```
df[numerical_cols] = scaler.fit_transform(df[numerical_cols])
```

Return the prepared DataFrame

```
return df
```

Load the credit risk dataset (mock data)

```
credit_risk_data = pd.read_csv('credit_risk_data.csv')
```

Perform data preparation

```
prepared_data = data_preparation(credit_risk_data)
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση `data_preparation` ορίζεται για να ενθυλακώσει τον απαραίτητο καθαρισμό των δεδομένων, τον χειρισμό των ελλιπών τιμών, την κωδικοποίηση των κατηγορικών μεταβλητών και την κλιμάκωση των

αριθμητικών μεταβλητών. Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού, των παραμέτρων και της τιμής επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση προετοιμάζει το σύνολο δεδομένων πιστωτικού κινδύνου για ανάλυση και επιστρέφει το προεξεργασμένο DataFrame. Η περιγραφή του κώδικα εξηγεί το σκεπτικό πίσω από κάθε βήμα της διαδικασίας προετοιμασίας των δεδομένων. Για παράδειγμα, οι ελλιπείς τιμές αντιμετωπίζονται με την

απόρριψη των γραμμών που περιέχουν ελλιπείς τιμές χρησιμοποιώντας τη μέθοδο `dropna`. Οι κατηγορικές μεταβλητές (στην προκειμένη περίπτωση, "Industry" και "Region") κωδικοποιούνται με τη χρήση του `LabelEncoder` από το `scikit-learn` για τη μετατροπή τους σε αριθμητικές αναπαραστάσεις. Οι αριθμητικές μεταβλητές (όπως οι "Debt-to-Income Ratio" και "Credit Score") κλιμακώνονται με τη χρήση του `StandardScaler` για να διασφαλιστεί ότι έχουν συγκρίσιμες κλίμακες και να αποτραπεί οποιαδήποτε μεροληψία στο μοντέλο. Αυτή η τεκμηρίωση του κώδικα και η περιγραφή παρέχουν σαφήνεια σχετικά με τον σκοπό κάθε βήματος και την επιλεγμένη προσέγγιση για την προετοιμασία του συνόλου δεδομένων πιστωτικού κινδύνου για περαιτέρω ανάλυση και μοντελοποίηση. Βοηθά στη διασφάλιση της αναπαραγωγιμότητας, της διαφάνειας και της ευκολίας κατανόησης τόσο για τον ερευνητή όσο και για κάθε μελλοντικό χρήστη του κώδικα.

## B.2 Επιλογή χαρακτηριστικών και μείωση διαστάσεων

Αυτό το μέρος θα περιλαμβάνει την υλοποίηση του μοντέλου αυτόματου κωδικοποιητή που χρησιμοποιείται για την επιλογή χαρακτηριστικών και τη μείωση της διαστατικότητας. Θα περιγράψει την αρχιτεκτονική του μοντέλου, τις επιλεγμένες υπερπαραμέτρους, τη διαδικασία εκπαίδευσης και τη μέθοδο εξαγωγής σημαντικών χαρακτηριστικών από το μοντέλο.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler

def autoencoder_feature_selection(X_train, X_test, encoding_dim=2, epochs=50, batch_size=32):
    """
```

Implement an autoencoder for feature selection and dimensionality reduction.

Parameters:

- X\_train: numpy array or pandas DataFrame of the training data
- X\_test: numpy array or pandas DataFrame of the test data
- encoding\_dim: dimensionality of the encoded representation (default: 2)
- epochs: number of training epochs (default: 50)
- batch\_size: batch size for training (default: 32)

Returns:

- encoded\_train: encoded representation of the training data
- encoded\_test: encoded representation of the test data

''''''

Scale the input data

```
scaler = StandardScaler()
```

```
X_train_scaled = scaler.fit_transform(X_train)
```

```
X_test_scaled = scaler.transform(X_test)
```

Define the autoencoder architecture

```
input_dim = X_
```

```
train_scaled.shape[1]
```

```
input_data = tf.keras.Input(shape=(input_dim,))
```

```
encoded = tf.keras.layers.Dense(encoding_dim, activation='relu')(input_data)
```

```
decoded = tf.keras.layers.Dense(input_dim, activation='sigmoid')(encoded)
```

```
autoencoder = tf.keras.Model(input_data, decoded)
```

Compile and train the autoencoder

```
autoencoder.compile(optimizer='adam', loss='mse')
```

```
autoencoder.fit(X_train_scaled, X_train_scaled, epochs=epochs, batch_size=batch_size, verbose=0)
```

Extract the encoded representations

```
encoder = tf.keras.Model(input_data, encoded)
encoded_train = encoder.predict(X_train_scaled)
encoded_test = encoder.predict(X_test_scaled)
```

Return the encoded representations of the training and test data

```
return encoded_train, encoded_test
```

Load the credit risk dataset (mock data)

```
credit_risk_data = pd.read_csv('credit_risk_data.csv')
```

Separate the features and the target variable

```
X = credit_risk_data.drop(columns=['Credit Risk'])
y = credit_risk_data['Credit Risk']
```

Split the data into training and test sets

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Perform feature selection and dimensionality reduction using autoencoder

```
encoded_train, encoded_test = autoencoder_feature_selection(X_train, X_test, encoding_dim=2,
epochs=50, batch_size=32)
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση `autoencoder_feature_selection` ορίζεται για την υλοποίηση ενός μοντέλου αυτόματου κωδικοποιητή για την επιλογή χαρακτηριστικών και τη μείωση της διαστατικότητας. Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού, των παραμέτρων και των τιμών επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση λαμβάνει τα δεδομένα εκπαίδευσης και δοκιμής, μαζί με προαιρετικές παραμέτρους για τη διάσταση κωδικοποίησης, τον αριθμό των εποχών και το μέγεθος της παρτίδας. Επιστρέφει τις κωδικοποιημένες αναπαραστάσεις των δεδομένων εκπαίδευσης και δοκιμής.

Η περιγραφή του κώδικα εξηγεί τα βήματα που απαιτούνται για την υλοποίηση του μοντέλου αυτόματου κωδικοποιητή. Τα δεδομένα εισόδου κλιμακώνονται πρώτα με τη χρήση

του StandardScaler από το scikit-learn για να διασφαλιστεί ότι τα χαρακτηριστικά έχουν συγκρίσιμες κλίμακες. Η αρχιτεκτονική του αυτόματου κωδικοποιητή ορίζεται χρησιμοποιώντας το API της Keras και αποτελείται από ένα στρώμα εισόδου, ένα στρώμα κωδικοποίησης με την καθορισμένη διαστατικότητα και ένα στρώμα αποκωδικοποίησης. Στη συνέχεια, ο αυτοκωδικοποιητής συντάσσεται και εκπαιδεύεται στα δεδομένα εκπαίδευσης. Μετά την εκπαίδευση, το τμήμα κωδικοποίησης του αυτόματου κωδικοποιητή χρησιμοποιείται για την εξαγωγή των κωδικοποιημένων αναπαραστάσεων των δεδομένων εκπαίδευσης και δοκιμής. Αυτές οι αναπαραστάσεις συλλαμβάνουν τα ουσιώδη χαρακτηριστικά και μειώνουν τη διαστατικότητα των δεδομένων. Τέλος, ο κώδικας επιδεικνύει τη χρήση της συνάρτησης `autoencoder_feature_selection` στο σύνολο δεδομένων πιστωτικού κινδύνου, χωρίζοντας τα δεδομένα σε σύνολα εκπαίδευσης και δοκιμής και λαμβάνοντας τις κωδικοποιημένες αναπαραστάσεις για περαιτέρω ανάλυση. Αυτή η τεκμηρίωση του κώδικα και η αφήγηση παρέχουν σαφή κατανόηση των λεπτομερειών υλοποίησης του μοντέλου αυτόματου κωδικοποιητή για την επιλογή χαρακτηριστικών και τη μείωση της διαστατικότητας. Συμβάλλει στη διασφάλιση της αναπαραγωγιμότητας και της διαφάνειας, επιτρέποντας σε άλλους να κατανοήσουν και να χρησιμοποιήσουν αποτελεσματικά τον κώδικα.

### B.3 Μοντέλα μηχανικής μάθησης και βαθιάς μάθησης

Η καρδιά της ενότητας υλοποίησης του μοντέλου θα αποτελείται από τα μοντέλα βαθιάς μάθησης που χρησιμοποιούνται για την πρόβλεψη του πιστωτικού κινδύνου. Θα ξεκινήσει με τον ορισμό της αρχιτεκτονικής του μοντέλου, καθορίζοντας τον αριθμό των στρωμάτων, τον τύπο των στρωμάτων (πυκνά, συνελκτικά, LSTM κ.λπ.), τον αριθμό των νευρώνων σε κάθε στρώμα και τις συναρτήσεις ενεργοποίησης. Θα ακολουθήσει η σύνταξη του μοντέλου, καθορίζοντας τον βελτιστοποιητή, τη συνάρτηση απώλειας και τις μετρικές.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
```

```
def create_credit_risk_model(input_dim):
```

```
''''''
```

Create a deep learning model for credit risk prediction.

Parameters:

- input\_dim: dimensionality of the input data

Returns:

- model: compiled deep learning model

```
''''''
```

Define the model architecture

```
model = tf.keras.Sequential()
```

```
model.add(tf.keras.layers.Dense(128, activation='relu', input_dim=input_dim))
```

```
model.add(tf.keras.layers.Dense(64, activation='relu'))
```

```
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
```

Compile the model

```
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

```
return model
```

Load the credit risk dataset (mock data)

```
credit_risk_data = pd.read_csv('credit_risk_data.csv')
```

Separate the features and the target variable

```
X = credit_risk_data.drop(columns=['Credit Risk']) y = credit_risk_data['Credit Risk']
```

Split the data into training and test sets

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Scale the input data

```
scaler = StandardScaler()
```

```
X_train_scaled = scaler.fit_transform(X_train)
```

```
X_test_scaled = scaler.transform(X_test)
```

Create and compile the credit risk prediction model

```
input_dim_ = X_train_scaled.shape[1]  
model = create_credit_risk_model(input_dim)
```

Train the model

```
model.fit(X_train_scaled, y_train, epochs=10, batch_size=32, verbose=0)
```

Evaluate the model on the test set

```
loss, accuracy = model.evaluate(X_test_scaled, y_test)  
print(f"Test Loss: loss:.4f, Test Accuracy: accuracy:.4f")
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση `create_credit_risk_model` ορίζεται για τη δημιουργία ενός μοντέλου βαθιάς μάθησης για την πρόβλεψη του πιστωτικού κινδύνου. Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού και της τιμής επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση λαμβάνει τη διαστατικότητα των δεδομένων εισόδου και επιστρέφει ένα μεταγλωττισμένο μοντέλο βαθιάς μάθησης. Η περιγραφή του κώδικα εξηγεί τα βήματα που απαιτούνται για τη δημιουργία του μοντέλου βαθιάς μάθησης. Η αρχιτεκτονική του μοντέλου ορίζεται χρησιμοποιώντας το Keras Sequential API, αποτελούμενο από μια σειρά πυκνών επιπέδων με καθορισμένες συναρτήσεις ενεργοποίησης. Σε αυτό το παράδειγμα, έχουμε δύο κρυφά στρώματα με 128 και 64 νευρώνες αντίστοιχα, ακολουθούμενα από ένα στρώμα εξόδου με έναν μόνο νευρώνα που χρησιμοποιεί τη σιγμοειδή συνάρτηση ενεργοποίησης. Στη συνέχεια, το μοντέλο μεταγλωττίζεται, καθορίζοντας τον βελτιστοποιητή ως "adam", τη συνάρτηση απώλειας ως "binary\_crossentropy" (κατάλληλη για προβλήματα δυαδικής ταξινόμησης) και τη μετρική ως "accuracy" για την αξιολόγηση της απόδοσης του μοντέλου. Στη συνέχεια, το σύνολο δεδομένων για τον πιστωτικό κίνδυνο προεπεξεργάζεται με την κλιμάκωση των δεδομένων εισόδου με τη χρήση του StandardScaler από το scikit-learn. Ο κώδικας επιδεικνύει τη χρήση της συνάρτησης `create_credit_risk_model` με τη δημιουργία του μοντέλου και την εκπαίδευσή του χρησιμοποιώντας τα κλιμακωτά δεδομένα εκπαίδευσης. Τέλος, το μοντέλο αξιολογείται στα κλιμα-

κωτά δεδομένα δοκιμής και εκτυπώνονται οι μετρικές απώλειας και ακρίβειας. Αυτή η τεκμηρίωση του κώδικα και η αφήγηση παρέχουν σαφή κατανόηση των λεπτομερειών υλοποίησης του μοντέλου βαθιάς μάθησης για την πρόβλεψη του πιστωτικού κινδύνου. Επιτρέπει την αναπαραγωγικότητα και τη διαφάνεια, δίνοντας τη δυνατότητα σε άλλους να κατανοήσουν και να χρησιμοποιήσουν αποτελεσματικά τον κώδικα.

## B.4 Εκπαίδευση και επικύρωση μοντέλων

Σε αυτή την ενότητα θα καταγραφεί η διαδικασία εκπαίδευσης των μοντέλων στο σύνολο εκπαίδευσης και επικύρωσής τους στο σύνολο επικύρωσης. Θα συζητηθεί ο αριθμός των εποχών, το μέγεθος της παρτίδας και η διαδικασία παρακολούθησης της απόδοσης του μοντέλου κατά τη διάρκεια της εκπαίδευσης για την αποφυγή της υπερβολικής προσαρμογής, όπως η πρόωρη διακοπή.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
```

```
def create_credit_risk_model(input_dim):
    """
```

Create a deep learning model for credit risk prediction.

Parameters:

- input\_dim: dimensionality of the input data

Returns:

- model: compiled deep learning model

```
"""
```

Define the model architecture

```
model = tf.keras.Sequential()
```

```
model.add(tf.keras.layers.Dense(128, activation='relu', input_dim=input_dim))
```

```
model.add(tf.keras.layers.Dense(64, activation='relu'))  
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
```

Compile the model

```
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

return model

Load the credit risk dataset (mock data)

```
credit_risk_data = pd.read_csv('credit_risk_data.csv')
```

Separate the features and the target variable

```
X = credit_risk_data.drop(columns=['Credit Risk'])
```

```
y = credit_risk_data['Credit Risk']
```

Split the data into training, validation, and test sets

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
X_val, X_test, y_val, y_test = train_test_split(X_val, y_val, test_size=0.5, random_state=42)
```

Scale the input data

```
scaler = StandardScaler()
```

```
X_train_scaled = scaler.fit_transform(X_train)
```

```
X_val_scaled = scaler.transform(X_val)
```

```
X_test_scaled = scaler.transform(X_test)
```

Create and compile the credit risk prediction model

```
input_dim = X_train_scaled.shape[1]
```

```
model = create_credit_risk_model(input_dim)
```

Train the model

```
history = model.fit(X_train_scaled, y_train, epochs=10, batch_size=32, validation_data=(X_val_scaled,  
y_val), verbose=0)
```

Evaluate the model on the test set

```
loss, accuracy = model.evaluate(X_test_scaled, y_test) print(f'Test Loss: loss:.4f, Test Accuracy: accuracy:.4f')
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση `create_credit_risk_model` ορίζεται για τη δημιουργία ενός μοντέλου βαθιάς μάθησης για την πρόβλεψη του πιστωτικού κινδύνου. Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού και της τιμής επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση λαμβάνει τη διαστατικότητα των δεδομένων εισόδου και επιστρέφει ένα μεταγλωττισμένο μοντέλο βαθιάς μάθησης. Η περιγραφή του κώδικα εξηγεί τα βήματα που απαιτούνται για την εκπαίδευση και την επικύρωση του μοντέλου βαθιάς μάθησης. Το σύνολο δεδομένων πιστωτικού κινδύνου χωρίζεται σε σύνολα εκπαίδευσης, επικύρωσης και δοκιμής χρησιμοποιώντας τη συνάρτηση `train_test_split` από το `scikit-learn`. Τα δεδομένα εισόδου κλιμακώνονται με τη χρήση του `StandardScaler`.

Το μοντέλο δημιουργείται και μεταγλωττίζεται χρησιμοποιώντας τη συνάρτηση `create_credit_risk_model`. Η διαδικασία εκπαίδευσης πραγματοποιείται με την κλήση της μεθόδου `fit` στο μοντέλο, περνώντας τα κλιμακωτά δεδομένα εκπαίδευσης, τη μεταβλητή-στόχο, τον αριθμό των εποχών και το μέγεθος της παρτίδας. Η παράμετρος `validation_data` χρησιμοποιείται για την παρακολούθηση της απόδοσης του μοντέλου στο σύνολο επικύρωσης κατά τη διάρκεια της εκπαίδευσης. Μετά την εκπαίδευση, το μοντέλο αξιολογείται στα κλιμακωτά δεδομένα δοκιμής με τη μέθοδο `evaluate`. Οι μετρικές απώλειας και ακρίβειας δοκιμής εκτυπώνονται. Αυτή η τεκμηρίωση του κώδικα και η αφήγηση παρέχουν σαφή κατανόηση των λεπτομερειών εφαρμογής της εκπαίδευσης και επικύρωσης του μοντέλου. Επιτρέπει την αναπαραγωγικότητα και τη διαφάνεια, επιτρέποντας σε άλλους να κατανοήσουν και να χρησιμοποιήσουν αποτελεσματικά τον κώδικα.

## B.5 Αξιολόγηση και σύγκριση μοντέλων

Εδώ θα περιγραφεί λεπτομερώς η διαδικασία δοκιμής του μοντέλου στο σύνολο δοκιμών και ο υπολογισμός των μετρήσεων απόδοσης. Θα γίνει επίσης σύγκριση των επιδόσεων των διαφόρων μοντέλων, καταδεικνύοντας τα δυνατά και τα αδύνατα σημεία τους.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score

def create_credit_risk_model(input_dim):
    """
    Create a deep learning model for credit risk prediction.

    Parameters:
    - input_dim: dimensionality of the input data

    Returns:
    - model: compiled deep learning model
    """
    Define the model architecture
    model = tf.keras.Sequential()
    model.add(tf.keras.layers.Dense(128, activation='relu', input_dim=input_dim))
    model.add(tf.keras.layers.Dense(64, activation='relu'))
    model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

    Compile the model
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

    return model

Load the credit risk dataset (mock data)
credit_risk_data = pd.read_csv('credit_risk_data.csv')

Separate the features and the target variable
```

```
X = credit_risk_data.drop(columns=['Credit Risk'])
y = credit_risk_data['Credit Risk']
```

Split the data into training, validation, and test sets

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2, random_state=42)
X_val, X_test, y_val, y_test = train_test_split(X_val, y_val, test_size=0.5, random_state=42)
```

Scale the input data

```
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_val_scaled = scaler.transform(X_val)
X_test_scaled = scaler.transform(X_test)
```

Create and compile the credit risk prediction model

```
input_dim = X_train_scaled.shape[1]
model = create_credit_risk_model(input_dim)
```

Train the model

```
model.fit(X_train_scaled, y_train, epochs=10, batch_size=32, verbose=0)
```

Evaluate the model on the test set

```
y_pred = model.predict(X_test_scaled)
y_pred = np.round(y_pred).flatten()
```

Compute performance metrics

```
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
```

Print the performance metrics

```
print(f'Test Accuracy: accuracy:.4f')
```

```
print(f'Precision: precision:.4f')  
print(f'Recall: recall:.4f')  
print(f'F1 Score: f1:.4f')
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση `create_credit_risk_model` ορίζεται για τη δημιουργία ενός μοντέλου βαθιάς μάθησης για την πρόβλεψη του πιστωτικού κινδύνου. Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού και της τιμής επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση λαμβάνει τη διαστατικότητα των δεδομένων εισόδου και επιστρέφει ένα μεταγλωττισμένο μοντέλο βαθιάς μάθησης.

Η περιγραφή του κώδικα εξηγεί τα βήματα που απαιτούνται για την αξιολόγηση και τη σύγκριση των επιδόσεων του μοντέλου. Το σύνολο δεδομένων για τον πιστωτικό κίνδυνο χωρίζεται σε σύνολα εκπαίδευσης, επικύρωσης και δοκιμής χρησιμοποιώντας τη συνάρτηση `train_test_split` της `scikit-learn`. Τα δεδομένα εισόδου κλιμακώνονται με τη χρήση του `StandardScaler`. Το μοντέλο δημιουργείται και μεταγλωττίζεται χρησιμοποιώντας τη συνάρτηση `create_credit_risk_model`. Το μοντέλο εκπαιδεύεται στα κλιμακωτά δεδομένα εκπαίδευσης. Μετά την εκπαίδευση, το μοντέλο αξιολογείται στα κλιμακωτά δεδομένα δοκιμών με την πραγματοποίηση προβλέψεων και τον υπολογισμό μετρήσεων επιδόσεων. Σε αυτό το παράδειγμα, χρησιμοποιούμε την ακρίβεια, την ανάκληση και το F1-score ως μετρικές αξιολόγησης. Οι μετρικές επιδόσεων υπολογίζονται χρησιμοποιώντας τις ετικέτες βασικής αλήθειας (`y_test`) και τις προβλεπόμενες ετικέτες (`y_pred`). Τέλος, ο κώδικας εκτυπώνει τις μετρήσεις απόδοσης, παρέχοντας μια αξιολόγηση της ακρίβειας, της ανάκλησης και του F1-score του μοντέλου στο σύνολο δοκιμών. Αυτή η τεκμηρίωση του κώδικα και η αφήγηση παρέχουν σαφή κατανόηση των λεπτομερειών εφαρμογής της αξιολόγησης και σύγκρισης μοντέλων. Επιτρέπει την αναπαραγωγικότητα και τη διαφάνεια, δίνοντας τη δυνατότητα σε άλλους να κατανοήσουν και να χρησιμοποιήσουν αποτελεσματικά τον κώδικα.

## B.6 Έλεγχοι αξιοπιστίας και ανάλυση ευαισθησίας

Σε αυτή την ενότητα θα εξηγηθούν τα βήματα που έγιναν για να εξασφαλιστεί η ευρωστία του μοντέλου, συμπεριλαμβανομένης της δοκιμής του σε διαφορετικά υποσύνολα

δεδομένων και της προσαρμογής των υπερπαραμέτρων. Η ανάλυση ευαισθησίας σε μακρο-οικονομικούς παράγοντες θα συμπεριληφθεί επίσης εδώ.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
```

```
def create_credit_risk_model(input_dim):
    """
```

Create a deep learning model for credit risk prediction.

Parameters:

- input\_dim: dimensionality of the input data

Returns:

- model: compiled deep learning model

```
"""
```

Define the model architecture

```
model = tf.keras.Sequential()
model.add(tf.keras.layers.Dense(128, activation='relu', input_dim=input_dim))
model.add(tf.keras.layers.Dense(64, activation='relu'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
```

Compile the model

```
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

```
return model
```

Load the credit risk dataset (mock data)

```
credit_risk_data = pd.read_csv('credit_risk_data.csv')
```

Separate the features and the target variable

```
X = credit_risk_data.drop(columns=['Credit Risk'])  
y = credit_risk_data['Credit Risk']
```

Perform robustness checks using different subsets of the data

```
subset_sizes = [0.5, 0.7, 0.9]
```

Subset sizes as a fraction of the original dataset

for size in subset\_sizes:

Split the data into training, validation, and test sets

```
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=1-size, random_state=42)  
X_val, X_test, y_val, y_test = train_test_split(X_val, y_val, test_size=0.5, random_state=42)
```

Scale the input data

```
scaler = StandardScaler()  
X_train_scaled = scaler.fit_transform(X_train)  
X_val_scaled = scaler.transform(X_val)  
X_test_scaled = scaler.transform(X_test)
```

Create and compile the credit risk prediction model

```
input_dim = X_train_scaled.shape[1]  
model = create_credit_risk_model(input_dim)
```

Train the model

```
model.fit(X_train_scaled, y_train, epochs=10, batch_size=32, verbose=0)
```

Evaluate the model on the test set

```
loss, accuracy = model.evaluate(X_test_scaled, y_test) print(f'Test Loss (Subset Size size):  
loss:.4f, Test Accuracy (Subset Size size): accuracy:.4f')
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση create\_credit\_risk\_model

ορίζεται για τη δημιουργία ενός μοντέλου βαθιάς μάθησης για την πρόβλεψη του πιστωτικού κινδύνου.

Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού και της τιμής επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση λαμβάνει τη διάσταση των δεδομένων εισόδου και επιστρέφει ένα μεταγλωττισμένο μοντέλο βαθιάς μάθησης. Η περιγραφή του κώδικα εξηγεί τα βήματα που απαιτούνται για τη διενέργεια ελέγχων αξιοπιστίας και ανάλυσης ευαισθησίας. Το σύνολο δεδομένων για τον πιστωτικό κίνδυνο χωρίζεται σε σύνολα εκπαίδευσης, επικύρωσης και δοκιμής χρησιμοποιώντας τη συνάρτηση `train_test_split` της `scikit-learn`. Στη συνέχεια, ο κώδικας πραγματοποιεί ελέγχους αξιοπιστίας με επανάληψη σε διαφορετικά μεγέθη υποσυνόλων. Για κάθε μέγεθος υποσυνόλου, πραγματοποιείται νέος διαχωρισμός και το μοντέλο εκπαιδεύεται και αξιολογείται χρησιμοποιώντας τα κλιμακωτά δεδομένα εκπαίδευσης και δοκιμής. Οι μετρικές απόδοσης, όπως η απώλεια και η ακρίβεια, εκτυπώνονται για κάθε μέγεθος υποσυνόλου, παρέχοντας πληροφορίες σχετικά με την ευρωστία του μοντέλου σε διαφορετικά υποσύνολα δεδομένων. Αυτή η τεκμηρίωση του κώδικα και η περιγραφή καταδεικνύουν την εφαρμογή των ελέγχων αξιοπιστίας και της ανάλυσης ευαισθησίας με τη χρήση διαφορετικών υποσυνόλων των δεδομένων. Επιτρέπει την αξιολόγηση των επιδόσεων του μοντέλου σε διάφορα υποσύνολα, παρέχοντας πληροφορίες για τη σταθερότητα και τη δυνατότητα γενίκευσής του.

## B.7 Ερμηνευσιμότητα και ερμηνευσιμότητα του μοντέλου

Τέλος, θα εξηγηθούν τα βήματα που ακολουθήθηκαν για την ερμηνεία και την εξήγηση των προβλέψεων του μοντέλου. Τεχνικές όπως οι SHapley Additive exPlanations (SHAP) μπορούν να χρησιμοποιηθούν εδώ για να δώσουν μια αίσθηση του ποια χαρακτηριστικά είχαν τον μεγαλύτερο αντίκτυπο στις αποφάσεις του μοντέλου.

```
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import shap
```

```
def create_credit_risk_model(input_dim):
    """
    Create a deep learning model for credit risk prediction.

    Parameters:
    - input_dim: dimensionality of the input data

    Returns:
    - model: compiled deep learning model
    """
    Define the model architecture
    model = tf.keras.Sequential()
    model.add(tf.keras.layers.Dense(128, activation='relu', input_dim=input_dim))
    model.add(tf.keras.layers.Dense(64, activation='relu'))
    model.add(tf.keras.layers.Dense(1, activation='sigmoid'))

    Compile the model
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

    return model

    Load the credit risk dataset (mock data)
    credit_risk_data = pd.read_csv('credit_risk_data.csv')

    Separate the features and the target variable
    X = credit_risk_data.drop(columns=['Credit Risk'])
    y = credit_risk_data['Credit Risk']

    Split the data into training and test sets
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

    Scale the input data
```

```
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train) X_test_scaled = scaler.transform(X_test)
```

Create and compile the credit risk prediction model

```
input_dim = X_train_scaled.shape[1]
model = create_credit_risk_model(input_dim)
```

Train the model

```
model.fit(X_train_scaled, y_train, epochs=10, batch_size=32, verbose=0)
```

Explain the model's predictions using SHAP

```
explainer = shap.Explainer(model, X_train_scaled)
shap_values = explainer(X_test_scaled)
```

Plot the SHAP values

```
shap.summary_plot(shap_values, X_test_scaled, feature_names=X.columns)
```

Analyze the impact of individual features on predictions

```
shap.summary_plot(shap_values, X_test_scaled, feature_names=X.columns, plot_type='bar')
```

Σε αυτό το παράδειγμα, υποθέτουμε ότι το σύνολο δεδομένων πιστωτικού κινδύνου φορτώνεται από ένα αρχείο CSV με όνομα "credit\_risk\_data.csv". Η συνάρτηση `create_credit_risk_model` ορίζεται για τη δημιουργία ενός μοντέλου βαθιάς μάθησης για την πρόβλεψη του πιστωτικού κινδύνου. Η τεκμηρίωση του κώδικα παρέχει μια επισκόπηση του σκοπού και της τιμής επιστροφής της συνάρτησης. Διευκρινίζει ότι η συνάρτηση λαμβάνει τη διάσταση των δεδομένων εισόδου και επιστρέφει ένα μεταγλωττισμένο μοντέλο βαθιάς μάθησης.

Η περιγραφή του κώδικα εξηγεί τα βήματα που απαιτούνται για την ερμηνεία και την επεξήγηση των προβλέψεων του μοντέλου με τη μέθοδο SHAP (SHapley Additive exPlanations). Το σύνολο δεδομένων για τον πιστωτικό κίνδυνο χωρίζεται σε σύνολα εκπαίδευσης και δοκιμής με τη χρήση της συνάρτησης `train_test_split` από το `scikit-learn`. Στη συνέχεια, τα δεδομένα εισόδου κλιμακώνονται με τη χρήση του `StandardScaler`. Το μοντέλο δημιουργείται και μεταγλωττίζεται χρησιμοποιώντας τη συνάρτηση `create_credit_risk_model`.

Το μοντέλο εκπαιδύεται στα κλιμακωτά δεδομένα εκπαίδευσης. Στη συνέχεια, οι προβλέψεις του μοντέλου εξηγούνται με τη χρήση της κλάσης `Explainer` από τη βιβλιοθήκη SHAP. Το αντικείμενο `explainer` αρχικοποιείται με το εκπαιδευμένο μοντέλο και τα κλιμακωτά δεδομένα εκπαίδευσης. Στη συνέχεια, το `explainer` χρησιμοποιείται για τον υπολογισμό των τιμών SHAP για το σύνολο δοκιμών. Ο κώδικας δείχνει πώς να απεικονίσετε τις τιμές SHAP χρησιμοποιώντας το `shap.summary_plot`. Το πρώτο διάγραμμα παρέχει μια σύνοψη των τιμών SHAP για κάθε χαρακτηριστικό, δείχνοντας την επίδραση των χαρακτηριστικών στις μεμονωμένες προβλέψεις. Το δεύτερο plot, με `plot_type='bar'`, εμφανίζει ένα ραβδόγραμμα που συνοψίζει την επίδραση κάθε χαρακτηριστικού στις προβλέψεις του μοντέλου. Αυτή η τεκμηρίωση του κώδικα και η αφήγηση παρουσιάζουν την εφαρμογή του SHAP για την ερμηνεία και την επεξήγηση των προβλέψεων του μοντέλου. Επιτρέπει την απόκτηση γνώσεων σχετικά με τη σημασία των χαρακτηριστικών και την κατανόηση της λογικής πίσω από τις αποφάσεις του μοντέλου.

Η εφαρμογή του κώδικα παρέχει μια συγκεκριμένη, απτή εκδήλωση των θεωρητικών δομών που παρουσιάζονται στη διατριβή και η κατάλληλη τεκμηρίωση είναι υψίστης σημασίας για την αναπαραγωγικότητα και τη διαφάνεια. Με την ενδεδειγμένη τεκμηρίωση κάθε βήματος, οι μελλοντικοί ερευνητές μπορούν να κατανοήσουν, να αναπαράγουν και να βασιστούν σε αυτή την εργασία.

# Παράρτημα Γ

## Δεοντολογία & Συγκατάθεση

### Γ.1 Έντυπα έγκρισης δεοντολογίας και συγκατάθεσης

Οι ηθικοί προβληματισμοί στο πλαίσιο της παρούσας διατριβής περιστρέφονται κυρίως γύρω από τη χρήση και το χειρισμό των δεδομένων. Τα ακατέργαστα δεδομένα που χρησιμοποιούνται για την παρούσα μελέτη πρέπει να είναι ανώνυμα και να έχουν αφαιρεθεί από κάθε αναγνωρίσιμη προσωπική ή εταιρική πληροφορία, ώστε να διασφαλίζεται η εμπιστευτικότητα. Επιπλέον, ως ερευνητής, είναι απαραίτητο να λάβει κανείς τις απαραίτητες άδειες και εγκρίσεις από τις αρμόδιες αρχές πριν από την έναρξη της μελέτης.

### Γ.2 Έγκριση δεοντολογίας

Συνήθως, οι έρευνες που αφορούν ανθρώπινα υποκείμενα ή ευαίσθητα δεδομένα απαιτούν έγκριση από μια επιτροπή δεοντολογίας, η οποία συνήθως συνδέεται με το πανεπιστήμιο στο οποίο σπουδάζει ο ερευνητής. Αυτό διασφαλίζει ότι η μελέτη συμμορφώνεται με τις συνήθειες δεοντολογικές κατευθυντήριες γραμμές για την έρευνα. Στην περίπτωσή μας, η πρωταρχική πηγή δεδομένων είναι εταιρικά χρηματοοικονομικά δεδομένα και μακροοικονομικά δεδομένα, τα οποία είναι συνήθως διαθέσιμα στο δημόσιο τομέα. Ωστόσο, η χρήση οποιωνδήποτε ιδιόκτητων δεδομένων ή δεδομένων που περιέχουν ευαίσθητες ή αναγνωρίσιμες πληροφορίες θα απαιτούσε έλεγχο δεοντολογίας. Το έγγραφο έγκρισης δεοντολογίας πρέπει να περιέχει τα εξής:

- Τίτλος του έργου

- Όνομα και στοιχεία επικοινωνίας του ερευνητή
- Σύντομη περιγραφή του έργου
- Λεπτομερής περιγραφή του είδους των δεδομένων που χρησιμοποιήθηκαν στην έρευνα
- Μέθοδοι συλλογής και αποθήκευσης δεδομένων
- Μέτρα που λαμβάνονται για τη διασφάλιση του απορρήτου και της ανωνυμίας των δεδομένων
- Πιθανοί δεοντολογικοί κίνδυνοι και μέτρα για τον μετριασμό τους

### Γ.3 Έντυπα συγκατάθεσης

Εάν η μελέτη απαιτεί δεδομένα ιδιοκτησίας από εταιρείες ή ιδιώτες, είναι απαραίτητο να ληφθεί η συγκατάθεση των ενδιαφερόμενων μερών. Το έντυπο συγκατάθεσης είναι ένα νομικά δεσμευτικό έγγραφο που περιγράφει λεπτομερώς τη φύση της μελέτης, τον τύπο των δεδομένων που απαιτούνται, τον τρόπο με τον οποίο θα χρησιμοποιηθούν και θα αποθηκευτούν τα δεδομένα, καθώς και διαβεβαιώσεις σχετικά με τα δικαιώματα του υποκειμένου, συμπεριλαμβανομένης της απόσυρσης από τη μελέτη. Το έντυπο συγκατάθεσης πρέπει να περιέχει τα εξής:

- Τίτλος του έργου
- Όνομα και στοιχεία επικοινωνίας του ερευνητή
- Λεπτομερείς πληροφορίες σχετικά με το σκοπό της μελέτης και τι συνεπάγεται η συμμετοχή
- ρητή δήλωση ότι η συμμετοχή είναι εθελοντική και ότι η συγκατάθεση μπορεί να ανακληθεί ανά πάσα στιγμή
- Δήλωση σχετικά με την εμπιστευτικότητα και τον χειρισμό δεδομένων
- θέση για να υπογράψει και να χρονολογήσει ο συμμετέχων το έντυπο

Τα έγγραφα αυτά δεν είναι μόνο ζωτικής σημασίας για τη δεοντολογική διεξαγωγή της έρευνας, αλλά παρέχουν επίσης διαφάνεια, διασφαλίζοντας ότι όλα τα μέρη που συμμετέχουν στην έρευνα κατανοούν τα δικαιώματά τους και τη φύση της μελέτης.

# Παράρτημα Δ

## Επικαιροποιήσεις βιβλιογραφίας (2023)

Υπό το πρίσμα των τελευταίων εξελίξεων της τεχνητής νοημοσύνης στο θέμα "Επίδραση μακροοικονομικών παραγόντων στην ανάλυση εταιρικού πιστωτικού κινδύνου με χρήση αλγορίθμων μηχανικής μάθησης και μοντέλων βαθιάς μάθησης", μπορούμε να αναπτύξουμε περαιτέρω τις ενότητες της μεθοδολογίας, λαμβάνοντας υπόψη τις τεχνολογίες αιχμής της τεχνητής νοημοσύνης, της μηχανικής μάθησης και της βαθιάς μάθησης που έχουν εμφανιστεί μεταξύ 2020 και 2023. Ορισμένες από αυτές τις τεχνολογίες βρίσκονται ακόμη σε εκκολαπτόμενα στάδια ή δοκιμάζονται σε πιλοτικά έργα. Η ακόλουθη επεξεργασία περιλαμβάνει παραπομπές σε σχετικές εργασίες από το arxiv.org, βάσεις δεδομένων του MIT και διακεκριμένους επιστήμονες TN.

### Δ.1 Νευρωνικά δίκτυα γραφημάτων (GNNs)

Τα GNN έχουν δείξει ότι μπορούν να αποτυπώσουν τις πολύπλοκες σχέσεις μεταξύ οντοτήτων και μπορούν να εφαρμοστούν στην ανάλυση πιστωτικού κινδύνου μοντελοποιώντας τη διασύνδεση μεταξύ επιχειρήσεων και μακροοικονομικών παραγόντων [114]. Κατασκευάζοντας μια αναπαράσταση των δεδομένων με βάση το γράφημα, τα GNN μπορούν να μάθουν χαρακτηριστικά υψηλού επιπέδου που συλλαμβάνουν τόσο τοπικές όσο και παγκόσμιες πληροφορίες στο δίκτυο, οδηγώντας σε δυνητικά καλύτερες προβλέψεις πιστωτικού κινδύνου.

## Δ.2 Εξηγήσιμη τεχνητή νοημοσύνη (XAI)

Οι επεξηγήσιμες τεχνικές τεχνητής νοημοσύνης, όπως οι LIME, SHAP και Integrated Gradients, έχουν κερδίσει δημοτικότητα για τη βελτίωση της ερμηνευσιμότητας και της επεξηγηματικότητας των μοντέλων βαθιάς μάθησης [115]. Με την ενσωμάτωση του XAI στον αγωγό ανάλυσης πιστωτικού κινδύνου, τα μοντέλα που προκύπτουν μπορούν να παρέχουν πιο διαφανείς και εφαρμόσιμες πληροφορίες για τους υπεύθυνους λήψης αποφάσεων στα χρηματοπιστωτικά ιδρύματα.

## Δ.3 Μάθηση μεταφοράς και προσαρμογή στον τομέα

Οι τεχνικές μάθησης μεταφοράς και προσαρμογής στον τομέα επιτρέπουν στα μοντέλα να αξιοποιούν τη γνώση που έχουν μάθει από έναν τομέα ή μια εργασία για να βελτιώσουν την απόδοσή τους σε έναν διαφορετικό, αλλά συναφή τομέα ή εργασία [116]. Αυτό μπορεί να είναι ιδιαίτερα χρήσιμο στην ανάλυση πιστωτικού κινδύνου, καθώς τα μοντέλα μπορούν να προ-εκπαιδευτούν σε σύνολα μακροοικονομικών δεδομένων μεγάλης κλίμακας και να ρυθμιστούν με τη χρήση χρηματοοικονομικών δεδομένων συγκεκριμένων επιχειρήσεων, βελτιώνοντας ενδεχομένως την απόδοση πρόβλεψης.

## Δ.4 Transformer-based Models

Μοντέλα βασισμένα σε μετασχηματιστές και οι παραλλαγές τους, έχουν δείξει εντυπωσιακά αποτελέσματα σε διάφορες εργασίες NLP και πρόσφατα επεκτάθηκαν σε άλλους τομείς, όπως τα δεδομένα σε πίνακες [117]. Αυτά τα μοντέλα μπορούν δυνητικά να προσαρμοστούν στην ανάλυση πιστωτικού κινδύνου με την κωδικοποίηση των μακροοικονομικών παραγόντων και των χρηματοοικονομικών δεδομένων συγκεκριμένης επιχείρησης ως ακολουθίες εισόδου και την εκπαίδευση των μοντέλων για την πρόβλεψη των αποτελεσμάτων του πιστωτικού κινδύνου.

## Δ.5 Μετρικές αμεροληψίας

Δεδομένης της αυξανόμενης ανησυχίας σχετικά με τις πιθανές προκαταλήψεις στα μοντέλα τεχνητής νοημοσύνης και μηχανικής μάθησης, οι μετρικές δικαιοσύνης, όπως η δημο-

γραφική ισοτιμία, οι εξισωμένες πιθανότητες και οι ίσες ευκαιρίες, μπορούν να ενσωματωθούν στη διαδικασία αξιολόγησης για την αξιολόγηση της δικαιοσύνης των μοντέλων ανάλυσης πιστωτικού κινδύνου [118]. Η διασφάλιση ότι τα μοντέλα είναι δίκαια και αμερόληπτα μπορεί να βοηθήσει τα χρηματοπιστωτικά ιδρύματα να λαμβάνουν πιο δίκαιες πιστωτικές αποφάσεις.





## Παράρτημα Ζ

### Πίνακας Συστημικής Ανασκόπησης

Τύπος Κινδύνου	Μέθοδος / Εργαλείο Διαχείρισης	Βιβλιογραφική Αναφορά	Αλγόριθμος
Compliance Risk Management	Risk Monitoring	Mainelli and Yeandle 2006	SVM
Credit Risk Management— Concentration Risk	Stress Testing	Pavlenko and Chernyak 2009	Bayesian Networks
Credit Risk Management— Consumer Credit	Exposure (PD, LGD, EAD)	Yeh and Lien 2009	Bayesclassifier, Nearest neighbor, ANN, Classification trees
Credit Risk Management— Consumer Credit	Scoring Models	Bellotti and Crook 2009	SVM
Credit Risk Management— Consumer Credit	Scoring Models	Galindo and Tamayo 2000	CART, NN, KNN
Credit Risk Management— Consumer Credit	Scoring Models	Wang et al. 2015	Lasso logistic regression
Credit Risk Management— Consumer Credit	Scoring Models	Hamori et al. 2018	Bagging, Random Forest, Boosting
Credit Risk Management— Consumer Credit	Scoring Models	Harris 2013	SVM
Credit Risk Management— Consumer Credit	Scoring Models	Huang et al. 2007	SVM
Credit Risk Management— Consumer Credit	Scoring Models	Keremati and Yousefi 2011	NN, Bayesian Classifier, DA, Logistic Regression, KNN, Decision tree, Survival Analysis, Fuzzy Rule based system, SVM, Hybrid mode
Credit Risk Management— Consumer Credit	Scoring Models	Khandani et al. 2010	CART
Credit Risk Management— Consumer Credit	Scoring Models	Lai et al. 2006	SVM
Credit Risk Management— Consumer Credit	Scoring Models	Lessmann et al. 2015	Multiple algos assessed
Credit Risk Management— Consumer Credit	Scoring Models	Van-Sang and Nguyen 2016	Deep Learning
Credit Risk Management— Consumer Credit	Scoring Models	Yu et al. 2016	Deep belief network, Extreme Machine Learning
Credit Risk Management— Consumer Credit	Scoring Models	Y. Wang et al. 2005	SVM, Fuzzy SVM

# Παράρτημα Ε

## Βιβλιογραφία

1. Ala'raj, Maher, and Maysam F. Abbod. 2016a. A New Hybrid Ensemble Credit Scoring Model Based on Classifiers Consensus System Approach. *Expert Systems with Applications* 64: 36–55. [Google Scholar] [CrossRef]
2. Ala'Raj, Maher, and Maysam F. Abbod. 2016b. Classifiers Consensus System Approach for Credit Scoring. *Knowledge-Based Systems* 104: 89–105. [Google Scholar] [CrossRef]
3. Apostolik, Richard, Christopher Donohue, Peter Went, and Global Association of Risk Professionals. 2009. *Foundations of Banking Risk: An Overview of Banking, Banking Risks, and Risk-Based Banking Regulation*. New York: John Wiley. [Google Scholar]
4. Arezzo, Maria, and Giuseppina Guagnano. 2018. Response-Based Sampling for Binary Choice Models with Sample Selection. *Econometrics* 6: 12. [Google Scholar] [CrossRef]
5. Awad, Mariette, and Rahul Khanna. 2015. *Machine Learning in Action: Examples. Efficient Learning Machines*. [Google Scholar] [CrossRef]
6. Aziz, Saqib, and Michael M. Dowling. 2018. AI and Machine Learning for Risk Management. *SSRN Electronic Journal*. [Google Scholar] [CrossRef]
7. Bacham, Dinesh, and Janet Zhao. 2017. *Machine Learning: Challenges and Opportunities in Credit Risk Modeling*. Available online: <https://www.moodysanalytics.com/risk-perspectives-magazine/managing-disruption/spotlight/machine-learning-challenges-lessons-and-opportunities-in-credit-risk-modeling> (accessed on 2 April 2018).
8. Barboza, Flavio, Herbert Kimura, and Edward Altman. 2017. Machine learning models and bankruptcy prediction. *Expert Systems with Applications* 83: 405–17. [Google Scholar] [CrossRef]

9. Basel Committee on Banking Supervision. 2005a. Guidance on Paragraph 468 of the Framework Document. Basel: Bank for International Settlements. [Google Scholar]
10. Basel Committee on Banking Supervision. 2005b. An Explanatory Note on the Basel II IRB Risk Weight Functions. Basel: Bank for International Settlements. [Google Scholar]
11. Basel Committee on Banking Supervision. 2006. Minimum Capital Requirements for Market Risk. Basel: Bank for International Settlements. [Google Scholar]
12. Basel Committee on Banking Supervision. 2008. Principles for Sound Liquidity Risk Management and Supervision. Basel: Bank for International Settlements. [Google Scholar]
13. Basel Committee on Banking Supervision. 2011. Principles for the Sound Management of Operational Risk. Basel: Bank for International Settlements, pp. 1–27. [Google Scholar]
14. Bastos, João A. 2014. Ensemble Predictions of Recovery Rates. *Journal of Financial Services Research* 46: 177–93. [Google Scholar] [CrossRef]
15. Bauguess, Scott W. 2015. The Hope and Limitations of Machine Learning in Market Risk Assessment. Washington, DC: U.S. Securities and Exchange Commission. [Google Scholar]
16. Bellotti, Tony, and Jonathan Crook. 2009. Support Vector Machines for Credit Scoring and Discovery of Significant Features. *Expert Systems with Applications*. [Google Scholar] [CrossRef]
17. Blom, Tineke. 2015. Top down Stress Testing: An Application of Adaptive Lasso to Forecasting Credit Loss Rates. Master's Thesis, Faculty of Science, Hongkong, China. [Google Scholar]
18. Brown, Iain, and Christophe Mues. 2012. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications* 39: 3446–53. [Google Scholar] [CrossRef]
19. Cao, Jie, Hongke Lu, Weiwei Wang, and Jian Wang. 2013. A Loan Default Discrimination Model Using Cost-Sensitive Support Vector Machine Improved by PSO. *Information Technology and Management* 14: 193–204. [Google Scholar] [CrossRef]
20. Chan-Lau, Jorge. 2017. Lasso Regressions and Forecasting Models in Applied Stress Testing. *IMF Working Papers* 17: 1. [Google Scholar] [CrossRef]
21. Chen, Ning, Bernardete Ribeiro, and An Chen. 2016. Financial Credit Risk Assessment: A Recent Review. *Artificial Intelligence Review* 45: 1–23. [Google Scholar] [CrossRef]

- 
22. Dal Pozzolo, Andrea. 2015. Adaptive Machine Learning for Credit Card Fraud Detection. Unpublished doctoral dissertation, Université libre de Bruxelles, Faculté des Sciences—Informatique, Bruxelles. [Google Scholar]
23. Deloitte University Press. 2017. Global Risk Management Survey, 10th ed. Deloitte University Press: Available online: <https://www2.deloitte.com/tr/en/pages/risk/articles/global-risk-management-survey-10th-ed.html> (accessed on 4 October 2018).
24. Financial Stability Board. 2017. Artificial Intelligence and Machine Learning in Financial Services. Market Developments and Financial Stability Implications. Financial Stability Board. November 1. Available online: <http://www.fsb.org/2017/11/artificial-intelligence-and-machine-learning-in-financial-service/> (accessed on 2 July 2018).
25. Galindo, Jorge, and Pablo Tamayo. 2000. Credit Risk Assessment Using Statistical and Machine Learning: Basic Methodology and Risk Modeling Applications. *Computational Economics* 15: 107–43. [Google Scholar] [CrossRef]
26. Gotoh, Jun-ya, Akiko Takeda, and Rei Yamamoto. 2014. Interaction between financial risk measures and machine learning methods. *Computational Management Science* 11: 365–402. [Google Scholar] [CrossRef]
27. Greene, William H. 1992. A Statistical Model for Credit Scoring. NYU Working Paper No. EC-92-29. Available online: <https://ssrn.com/abstract=1867088> (accessed on 8 April 1992).
28. Guegan, Dominique, Peter Addo, and Bertrand Hassani. 2018. Credit risk analysis using machine and deep learning models. *Risks* 6: 38. [Google Scholar]
29. Hamori, Shigeyuki, Minami Kawai, Takahiro Kume, Yuji Murakami, and Chikara Watanabe. 2018. Ensemble Learning or Deep Learning? Application to Default Risk Analysis. *Journal of Risk and Financial Management* 11: 12. [Google Scholar] [CrossRef]
30. Hand, David J., and William E. Henley. 1997. Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of the Royal Statistical Society Series A: Statistics in Society*. [Google Scholar] [CrossRef]
31. Harris, Terry. 2013. Quantitative credit risk assessment using support vector machines: Broad versus Narrow default definitions. *Expert Systems with Applications* 40: 4404–13. [Google Scholar] [CrossRef]
32. Helbekkmo, Hans, Alok Kshirsagar, Andreas Schlosser, Francesco Selandari, Uwe Stegemann, and Joyce Vorholt. 2013. Enterprise Risk Management—Shaping the Risk Rev-

olution. New York: McKinsey & Co., Available online: [www.rmahq.org](http://www.rmahq.org) (accessed on 18 June 2018).

33. Huang, Cheng Lung, Mu Chen Chen, and Chieh Jen Wang. 2007. Credit Scoring with a Data Mining Approach Based on Support Vector Machines. *Expert Systems with Applications* 33: 847–56. [Google Scholar] [CrossRef]

34. Hull, John. 2012. *Risk Management and Financial Institutions*. New York: John Wiley and Sons, vol. 733. [Google Scholar]

35. Islam, Tushith, Christos Vasilopoulos, and Erik Pruyt. 2013. Stress—Testing Banks under Deep Uncertainty. Paper presented at the 31st International Conference of the System Dynamics Society, Cambridge, MA, USA, July 21–25; Available online: <http://repository.tudelft.nl/islandora/4235-4d29-8eed-3246df87e119?collection=education> (accessed on 17 July 2018).

36. Jacobs, Michael, Jr. 2018. The validation of machine-learning models for the stress testing of credit risk. *Journal of Risk Management in Financial Institutions* 11: 218–43. [Google Scholar]

37. Jorion, Philippe. 2007. *Value at Risk: The New Benchmark for Managing Financial Risk*. New York: McGraw-Hill. [Google Scholar]

38. Kanevski, Mikhail F., and Vadim Timonin. 2010. Machine learning analysis and modeling of interest rate curves. Paper presented at the 18th European Symposium on Artificial Neural Networks ESANN, Bruges, Belgium, April 28–30; Available online: <https://www.elen.ucl.ac.be/Proc/17.pdf> (accessed on 18 June 2018).

39. Kannan, Somasundaram, and K. Somasundaram. 2017. Autoregressive-Based Outlier Algorithm to Detect Money Laundering Activities. *Journal of Money Laundering Control* 20: 190–202. [Google Scholar] [CrossRef]

40. Keramati, Abbas, and Niloofar Yousefi. 2011. A proposed classification of data mining techniques in credit scoring. Paper presented at the 2011 International Conference of Industrial Engineering and Operations Management, Kuala Lumpur, Malaysia, January 22–24. [Google Scholar]

41. Khandani, Amir E., Adlar J. Kim, and Andrew W. Lo. 2010. Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance* 34: 2767–87. [Google Scholar]

42. Khrestina, Marina Pavlovna, Dmitry Ivanovich Dorofeev, Polina Andreevna Kachurina, Timur Rinatovich Usubaliev, and Aleksey Sergeevich Dobrotvorskiy. 2017. Develop-

---

ment of Algorithms for Searching, Analyzing and Detecting Fraudulent Activities in the Financial Sphere. *European Research Studies Journal* 20: 484–98. [Google Scholar]

43. Lai, Kin Keung, Lean Yu, Ligang Zhou, and Shouyang Wang. 2006. Credit risk evaluation with least square support vector machine. In *International Conference on Rough Sets and Knowledge Technology*. Berlin/Heidelberg: Springer, pp. 490–95. [Google Scholar]

44. Lessmann, Stefan, Bart Baesens, Hsin Vonn Seow, and Lyn C. Thomas. 2015. Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research. *European Journal of Operational Research* 247: 124–36. [Google Scholar] [CrossRef]

45. Mahdavi-Damghani, Babak, and Stephen Roberts. 2017. A Proposed Risk Modeling Shift from the Approach of Stochastic Differential Equation towards Machine Learning Clustering: Illustration with the Concepts of Anticipative and Responsible VaR. *SSRN Electronic Journal*. [Google Scholar] [CrossRef]

46. Mainelli, Michael, and Mark Yeandle. 2006. Best execution compliance: New techniques for managing compliance risk. *The Journal of Risk Finance* 7: 301–12. [Google Scholar] [CrossRef]

47. Malhotra, Rashmi, and D. K. Malhotra. 2003. Evaluating Consumer Loans Using Neural Networks. *Omega* 31: 83–96. [Google Scholar] [CrossRef]

48. MetricStream. 2018. The Chief Risk Officer's Role in 2018 and Beyond Managing the Challenges and Opportunities of a Digital Era New Roles of the CRO. Available online: <https://www.metricstream.com/insights/chief-risk-officer-role-2018.htm> (accessed on 23 June 2018).

49. Monfared, Soheil Almasi, and David Enke. 2014. Volatility Forecasting Using a Hybrid GJR-GARCH Neural Network Model. *Procedia Computer Science* 36: 246–53. [Google Scholar] [CrossRef]

50. Ngai, Eric W. T., Yong Hu, Yiu Hing Wong, Yijun Chen, and Xin Sun. 2011. The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and an Academic Review of Literature. *Decision Support Systems* 50: 569. [Google Scholar] [CrossRef]

51. Oliver Wyman. 2017. Next Generation Risk Management. Available online: [https://www.oliverwyman.com/~/media/OwensCorning/~/media/2017/08/Next\\_Generation\\_Risk\\_Management\\_Targeting\\_A-Technology\\_Dividend.pdf](https://www.oliverwyman.com/~/media/OwensCorning/~/media/2017/08/Next_Generation_Risk_Management_Targeting_A-Technology_Dividend.pdf)

52. Pavlenko, Tatjana, and Oleksandr Chernyak. 2009. Bayesian Networks for Modeling and Assessment of Credit Concentration Risks. *International Statistical Conference Prague*. Available online: [http://www.czso.cz/conference2009/proceedings/data/methods/pavlenko\\_paper.pdf](http://www.czso.cz/conference2009/proceedings/data/methods/pavlenko_paper.pdf) (accessed on 23 June 2018).

53. Peters, Gareth, Pavel V. Shevchenko, Ruben Cohen, and Diane Maurice. 2017. Statistical Machine Learning Analysis of Cyber Risk Data: Event Case Studies. Available online: <https://ssrn.com/abstract=3073704> (accessed on 18 June 2018).
54. Proofpoint. 2010. MLX Whitepaper “Machine Learning to Beat Spam Today and Tomorrow”. Available online: <https://www.excelmicro.com/datasheets/Proofpoint-White-Paper-MLX-Technology.pdf> (accessed on 2 May 2018).
55. Pun, Joseph, and Yuri Lawryshyn. 2012. Improving credit card fraud detection using a meta-classification strategy. *International Journal of Computer Applications* 56: 41–46. [Google Scholar] [CrossRef]
56. Raei, Reza, Mahdi Saeidi Kousha, Saeid Fallahpour, and Mohammad Fadaeinejad. 2016. A Hybrid Model for Estimating the Probability of Default of Corporate Customers. *Iranian Journal of Management Studies* 9: 651–73. [Google Scholar]
57. Ray, Sunil. 2015. Understanding Support Vector Machine Algorithm from Examples (along with Code). Available online: <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/> (accessed on 16 August 2018).
58. Sala, Jordi Petchamé. 2011. Liquidity Risk Modeling Using Artificial Neural Network. Master’s thesis, Universitat Politècnica de Catalunya, Barcelona, Spain. [Google Scholar]
59. Saunders, Anthony, Marcia Millon Cornett, and Patricia Anne McGraw. 2006. *Financial Institutions Management: A Risk Management Approach*. New York: McGraw-Hill. [Google Scholar]
60. Shalev-Shwartz, Shai, and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press. [Google Scholar] [CrossRef]
61. Sharma, Shashank, and Arjun Roy Choudhury. 2016. Fraud Analytics: A Survey on Bank Fraud Prediction Using Unsupervised Learning Based Approach. *International Journal of Innovation in Engineering Research and Technology* 3: 1–9. [Google Scholar]
62. Sudjianto, Agus, Sheela Nair, Ming Yuan, Aijun Zhang, Daniel Kern, and Fernando Cela-Díaz. 2010. Statistical Methods for Fighting Financial Crimes. *Technometrics* 52: 5–19. [Google Scholar] [CrossRef]
63. Tavana, Madjid, Amir Reza Abtahi, Debora Di Caprio, and Maryam Poortarigh. 2018. An Artificial Neural Network and Bayesian Network Model for Liquidity Risk Assessment in Banking. *Neurocomputing* 275: 2525–54. [Google Scholar] [CrossRef]

---

64. Vaidya, Avanti H., and Sudhir W. Mohod. 2014. Internet Banking Fraud Detection using HMM and BLAST-SSAHA Hybridization. *International Journal of Science and Research (IJSR)* 3: 574–9. [Google Scholar]

65. Van Gestel, Ir Tony, Bart Baesens, Ir Joao Garcia, and Peter Van Dijcke. 2003. A support vector machine approach to credit scoring. In *Forum Financier—Revue Bancaire Et Financieraire Bank En Financiewezen*. Bruxelles: Larcier, pp. 73–82. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.6492&rep=rep1&type=pdf> (accessed on 7 July 2018).

66. Van Liebergen, Bart. 2017. Machine Learning: A Revolution in Risk Management and Compliance? *Journal of Financial Transformation* 45: 60–67. [Google Scholar]

67. Van-Sang, Ha, and Ha-Nam Nguyen. 2016. Credit Scoring with a Feature Selection Approach Based Deep Learning. In *MATEC Web of Conferences*. Les Ulis: EDP Sciences, volume 54, p. 05004. Available online: <https://www.matec-conferences.org/articles/mateconf/abs/2016/17/>

68. Villalobos, Miguel Agustín, and Eliud Silva. 2017. A Statistical and Machine Learning Model to Detect Money Laundering: An Application. Available online: <http://hddavii.eventos.cimat.mx/>

69. Wang, Yongqiao, Shouyang Wang, and Kin Keung Lai. 2005. A New Fuzzy Support Vector Machine to Evaluate Credit Risk. *IEEE Transactions on Fuzzy Systems* 13: 820–31. [Google Scholar] [CrossRef]

70. Wang, Hong, Qingsong Xu, and Lifeng Zhou. 2015. Large Unbalanced Credit Scoring Using Lasso-Logistic Regression Ensemble. *PLoS ONE* 10: e0117844. [Google Scholar] [CrossRef] [PubMed]

71. Wójcicka, Aleksandra. 2017. Neural Networks vs. Discriminant Analysis in the Assessment of Default. *Electronic Economy*, 339–49. [Google Scholar] [CrossRef]

72. Yang, Zijiang, Wenjie You, and Guoli Ji. 2011. Using partial least squares and support vector machines for bankruptcy prediction. *Expert Systems with Applications* 38: 8336–42. [Google Scholar] [CrossRef]

73. Yao, Xiao, Jonathan Crook, and Galina Andreeva. 2015. Support vector regression for loss given default modelling. *European Journal of Operational Research* 240: 528–38. [Google Scholar] [CrossRef]

74. Yao, Xiao, Jonathan Crook, and Galina Andreeva. 2017. Enhancing two-stage modelling methodology for loss given default with support vector machines. *European Journal of Operational Research* 263: 679–89. [Google Scholar] [CrossRef]

75. Yeh, I. Cheng, and Chehui Lien. 2009. The Comparisons of Data Mining Techniques for the Predictive Accuracy of Probability of Default of Credit Card Clients. *Expert Systems with Applications*. [Google Scholar] [CrossRef]
76. Yu, Lean, Zebin Yang, and Ling Tang. 2016. A Novel Multistage Deep Belief Network Based Extreme Learning Machine Ensemble Learning Paradigm for Credit Risk Assessment. *Flexible Services and Manufacturing Journal* 28: 576–92. [Google Scholar] [CrossRef]
77. Zareapoor, Masoumeh, and Pourya Shamsolmoali. 2015. Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier. *Procedia Computer Science* 48: 679–86. [Google Scholar] [CrossRef]
78. Zhang, Wenhao. 2017. Machine Learning Approaches to Predicting Company Bankruptcy. *Journal of Financial Risk Management* 6: 364–74. [Google Scholar] [CrossRef]
79. Zhang, Heng Guo, Chi Wei Su, Yan Song, Shuqi Qiu, Ran Xiao, and Fei Su. 2017. Calculating Value-at-Risk for High-Dimensional Time Series Using a Nonlinear Random Mapping Model. *Economic Modelling* 67: 355–67. [Google Scholar] [CrossRef]
80. Zhou, Lifeng, and Hong Wang. 2012. Loan Default Prediction on Large Imbalanced Data Using Random Forests. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. [Google Scholar] [CrossRef]
81. Richardson, A.; Mulder, T. Nowcasting New Zealand GDP using machine learning algorithms. In *Proceedings of the Use of Big Data Analytics and Artificial Intelligence in Central Banking*, Bali, Indonesia, 23–26 July 2018.
82. Goudarzi, S.; Hickok, E.; Sinha, A. AI in Banking and Finance. *The Center for Internet and Society*. 2018. Available online: <https://journals.sagepub.com/doi/10.1177/09722629221087371> (accessed on 13 Jan 2023).
83. Awan, T. (Producer). *How Artificial Intelligence Helps Pakistan to Fight Its Battle*. Technology Times.
84. Liu, L.; Wang, W. Exchange rates forecasting with least squares support vector machine. In *Proceedings of the the 2008 International Conference on Computer Science and Software Engineering*, Washington, DC, USA, 12–14 December 2008.
85. Collinson, S.; Rugman, A.M. The regional character of Asian multinational enterprises. *Asia Pac. J. Manag.* 2007, 24, 429–446. [CrossRef]
86. Ramakrishnan, S.; Butt, S.; Chohan, M.A.; Ahmad, H. Forecasting Malaysian exchange rate using machine learning techniques based on commodities prices. In *Proceedings*

---

of the 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), Langkawi, Malaysia, 16–17 July 2017.

87. Jena, P.R.; Majhi, R.; Majhi, B. Development and performance evaluation of a novel knowledge guided artificial neural network (KGANN) model for exchange rate prediction. *J. King Saud Univ. Comput. Inf. Sci.* 2015, 27, 450–457. [CrossRef]

88. Shah, D.; Campbell, W.; Zulkernine, F.H. A comparative study of LSTM and DNN for stock market forecasting. In *Proceedings of the 2018 IEEE International Conference on Big Data (big data)*, Seattle, WA, USA, 10–13 December 2018.

89. Raj, D.A. Spotlight on the Remarkable Potential of AI in KYC.

90. Navyasri, S.; Hafsa, N.; Nayana, S.; Bilimale, N. Predicting the GDP of India using Machine Learning. *Int. J. Progress. Res. Sci. Eng.* 2020, 1, 78–81.

91. Yoon, J. Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Comput. Econ.* 2021, 57, 247–265. [CrossRef]

92. Maehashi, K.; Shintani, M. Macroeconomic forecasting using factor models and machine learning: An application to Japan. *J. Jpn Int. Econ.* 2020, 58, 101104. [CrossRef]

93. Ülke, V.; Sahin, A.; Subasi, A. A comparison of time series and machine learning models for inflation forecasting: Empirical evidence from the USA. *Neural Comput. Appl.* 2018, 30, 1519–1527. [CrossRef]

94. Syed, A.A.S.; Lee, K.H. Macroeconomic forecasting for Pakistan in a data-rich environment. *Appl. Econ.* 2021, 53, 1077–1091 [CrossRef]

95. Coulombe, P.G.; Leroux, M.; Stevanovic, D.; Surprenant, S. How is machine learning useful for macroeconomic forecasting? *arXiv* 2020, arXiv:2008.12477. [CrossRef]

96. Gogas, P.; Papadimitriou, T.; Takli, E. Comparison of simple sum and Divisia monetary aggregates in GDP forecasting: A support vector machines approach. *Econ. Bull.* 2013, 33, 1101–1115.

97. Aziz, S.; Dowling, M. Machine learning and AI for risk management. In *Disrupting Finance*; Lynn, T., Mooney, J.G., Rosati, P., Cummins, M., Eds.; Palgrave Pivot Cham: New York, NY, USA, 2019; pp. 30–50.

98. Medeiros, M.C.; Vasconcelos, G.F.; Veiga, Á.; Zilberman, E. Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *J. Bus. Econ. Stat.* 2021, 39, 98–119. [CrossRef]

99. Kou, G.; Chao, X.; Peng, Y.; Alsaadi, F.E.; Herrera-Viedma, E. Machine Learning

Methods for Systemic Risk Analysis in Financial Sectors. *Technol. Econ. Dev. Econ.* 2019, 25, 716–742. [CrossRef]

100. Huang, Y.; Kou, G.; Peng, Y. Nonlinear manifold learning for early warnings in financial markets. *Eur. J. Oper. Res.* 2017, 2, 692–702. [CrossRef]

101. Chao, X.; Peng, Y. A cost-sensitive multi-criteria quadratic programming model for imbalanced data. *J. Oper. Res. Soc.* 2017, 69, 500–516. [CrossRef]

102. Gensler, G.; Bailey, L. Deep Learning and Financial Stability. 2020. Available online: <https://www.semanticscholar.org/paper/Deep-Learning-and-Financial-Stability-Gensler-Bailey/daeb93c0419ee9af5d6c785ebe1e75c9fb3ea73f> (accessed on 13 Jan 2023).

103. Danielsson, J.; Macrae, R.; Uthemann, A. Artificial intelligence and systemic risk. *J. Bank. Financ. Forthcom.* 2021, 140, 106290. [CrossRef]

104. Boukherouaa, E.B.; Shabsigh, G.; AlAjmi, K.; Deodoro, J.; Farias, A.; Iskender, E.S.; Mirestean, A.T.; Ravikumar, R. Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance; International Monetary Fund: Washington, DC, USA, 2021.

105. Cicceri, G.; de Vita, F.; Bruneo, D.; Merlino, G.; Puliafito, A. A deep learning approach for pressure ulcer prevention using wearable computing. *Hum.-Cent. Comput. Inf. Sci.* 2020, 10, 5. [CrossRef]

106. AlShorman, O.; Alkhatni, F.; Masadeh, M.; Irfan, M.; Glowacz, A.; Althobiani, F.; Kozik, J.; Glowacz, W. Sounds and acoustic emission-based early fault diagnosis of induction motor: A review study. *Adv. Mech. Eng.* 2021, 13, 1–19. [CrossRef]

107. Ahmed, E.; Hamdan, A. The impact of corporate governance on firm performance: Evidence from Bahrain Bourse. *Marietta* 2015, 11, 21–37.

108. G. Zioviris, "Soft Computing: An Intelligent Fraud Detection Model based on Deep Learning Ensembles," in 2023.

109. G. Zioviris, K. Kolomvatsos, and G. Stamoulis, "Credit card fraud detection using a deep learning multistage model," 2022.

110. G. Zioviris, K. Kolomvatsos, and G. Stamoulis, "On the Use of a Sequential Deep Learning Scheme for Financial Fraud Detection," 2022.

111. B. Bernanke, M. Gertler, and S. Gilchrist, "The financial accelerator in a quantitative business cycle framework," *Handbook of Macroeconomics*, vol. 1, pp. 1341-1393, 1999.

112. S. Sundaresan, "Credit risk transfer, macroeconomic implications, and policy," *Jour-*

nal of Financial Stability, vol. 26, pp. 1-13, 2016.

113. C. P. Kindleberger and R. Z. Aliber, "Manias, Panics, and Crashes: A History of Financial Crises," Palgrave Macmillan, 2011.

114. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P. S. (2023). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*.

115. Adadi, A., Berrada, M. (2023). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138-52160.

116. Pan, S. J., Yang, Q. (2023). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-135/

117. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2023). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998-6008.

118. Verma, S., Rubin, J. (2023). Fairness definitions explained. In *Proceedings of the International Workshop*

119. Richardson, A.; Mulder, T. Nowcasting New Zealand GDP using machine learning algorithms. In *Proceedings of the Use of Big Data Analytics and Artificial Intelligence in Central Banking*, Bali, Indonesia, 23–26 July 2018.

120. Goudarzi, S.; Hickok, E.; Sinha, A. AI in Banking and Finance. *The Center for Internet and Society*. 2018. Available online: <https://journals.sagepub.com/doi/10.1177/09722629221087371> (accessed on 13 Jan 2023).

121. Awan, T. (Producer). How Artificial Intelligence Helps Pakistan to Fight Its Battle. *Technology Times*.

122. Liu, L.; Wang, W. Exchange rates forecasting with least squares support vector machine. In *Proceedings of the the 2008 International Conference on Computer Science and Software Engineering*, Washington, DC, USA, 12–14 December 2008.

123. Collinson, S.; Rugman, A.M. The regional character of Asian multinational enterprises. *Asia Pac. J. Manag.* 2007, 24, 429–446. [CrossRef]

124. Ramakrishnan, S.; Butt, S.; Chohan, M.A.; Ahmad, H. Forecasting Malaysian exchange rate using machine learning techniques based on commodities prices. In *Proceedings of the 2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, Langkawi, Malaysia, 16–17 July 2017.

125. Jena, P.R.; Majhi, R.; Majhi, B. Development and performance evaluation of a novel knowledge guided artificial neural network (KGANN) model for exchange rate prediction. *J. King Saud Univ. Comput. Inf. Sci.* 2015, 27, 450–457. [CrossRef]

126. Shah, D.; Campbell, W.; Zulkernine, F.H. A comparative study of LSTM and DNN for stock market forecasting. In Proceedings of the 2018 IEEE International Conference on Big Data (big data), Seattle, WA, USA, 10–13 December 2018.
127. Raj, D.A. Spotlight on the Remarkable Potential of AI in KYC.
128. Navyasri, S.; Hafsa, N.; Nayana, S.; Bilimale, N. Predicting the GDP of India using Machine Learning. *Int. J. Progress. Res. Sci. Eng.* 2020, 1, 78–81.
129. Yoon, J. Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Comput.Econ.* 2021, 57, 247–265. [CrossRef]
130. Maehashi, K.; Shintani, M. Macroeconomic forecasting using factor models and machine learning: An application to Japan. *J. Jpn Int. Econ.* 2020, 58, 101104. [CrossRef]
131. Ülke, V.; Sahin, A.; Subasi, A. A comparison of time series and machine learning models for inflation forecasting: Empirical evidence from the USA. *Neural Comput. Appl.* 2018, 30, 1519–1527. [CrossRef]
132. Syed, A.A.S.; Lee, K.H. Macroeconomic forecasting for Pakistan in a data-rich environment. *Appl. Econ.* 2021, 53, 1077–1091 [CrossRef]
133. Coulombe, P.G.; Leroux, M.; Stevanovic, D.; Surprenant, S. How is machine learning useful for macroeconomic forecasting? arXiv 2020, arXiv:2008.12477. [CrossRef]
134. Gogas, P.; Papadimitriou, T.; Takli, E. Comparison of simple sum and Divisia monetary aggregates in GDP forecasting: A support vector machines approach. *Econ. Bull.* 2013, 33, 1101–1115.
135. Aziz, S.; Dowling, M. Machine learning and AI for risk management. In *Disrupting Finance*; Lynn, T., Mooney, J.G., Rosati, P., Cummins, M., Eds.; Palgrave Pivot Cham: New York, NY, USA, 2019; pp. 30–50.
136. Medeiros, M.C.; Vasconcelos, G.F.; Veiga, Á.; Zilberman, E. Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *J. Bus. Econ. Stat.* 2021, 39, 98–119. [CrossRef]
137. Kou, G.; Chao, X.; Peng, Y.; Alsaadi, F.E.; Herrera-Viedma, E. Machine Learning Methods for Systemic Risk Analysis in Financial Sectors. *Technol. Econ. Dev. Econ.* 2019, 25, 716–742. [CrossRef]
138. Huang, Y.; Kou, G.; Peng, Y. Nonlinear manifold learning for early warnings in financial markets. *Eur. J. Oper. Res.* 2017, 2, 692–702. [CrossRef]
139. Chao, X.; Peng, Y. A cost-sensitive multi-criteria quadratic programming model for

---

imbalanced data. *J. Oper. Res. Soc.* 2017, 69, 500–516. [CrossRef]

140. Gensler, G.; Bailey, L. Deep Learning and Financial Stability. 2020. Available online: <https://www.semanticscholar.org/paper/Deep-Learning-and-Financial-Stability-Gensler-Bailey/daeb93c0419ee9af5d6c785ebe1e75c9fb3ea73f> (accessed on 13 Jan 2023).

141. Danielsson, J.; Macrae, R.; Uthemann, A. Artificial intelligence and systemic risk. *J. Bank. Financ. Forthcom.* 2021, 140, 106290. [CrossRef]

142. Boukherouaa, E.B.; Shabsigh, G.; AlAjmi, K.; Deodoro, J.; Farias, A.; Iskender, E.S.; Mirestean, A.T.; Ravikumar, R. Powering the Digital Economy: Opportunities and Risks of Artificial Intelligence in Finance; International Monetary Fund: Washington, DC, USA, 2021.

143. Cicceri, G.; de Vita, F.; Bruneo, D.; Merlino, G.; Puliafito, A. A deep learning approach for pressure ulcer prevention using wearable computing. *Hum.-Cent. Comput. Inf. Sci.* 2020, 10, 5. [CrossRef]

144. AlShorman, O.; Alkhatni, F.; Masadeh, M.; Irfan, M.; Glowacz, A.; Althobiani, F.; Kozik, J.; Glowacz, W. Sounds and acoustic emission-based early fault diagnosis of induction motor: A review study. *Adv. Mech. Eng.* 2021, 13, 1–19. [CrossRef]

145. Ahmed, E.; Hamdan, A. The impact of corporate governance on firm performance: Evidence from Bahrain Bourse. *Marietta* 2015, 11, 21–37.

146. Xiang, S., Zhu, M., Cheng, D., Li, E., Zhao, R., Ouyang, Y., ... & Zheng, Y. (2023, June). Semi-supervised credit card fraud detection via attribute-driven graph representation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 12, pp. 14557-14565).

147. Zhu, Honghao, et al. "NUS: Noisy-Sample-Removed Undersampling Scheme for Imbalanced Classification and Application to Credit Card Fraud Detection." *IEEE Transactions on Computational Social Systems* (2023).

148. Leevy, Joffrey L., John Hancock, and Taghi M. Khoshgoftaar. "Comparative analysis of binary and one-class classification techniques for credit card fraud data." *Journal of Big Data* 10.1 (2023): 118.

149. Prabhakaran, N., and R. Nedunchelian. "Oppositional Cat Swarm Optimization-Based Feature Selection Approach for Credit Card Fraud Detection." *Computational Intelligence and Neuroscience* 2023 (2023).

150. Jiang, Shanshan, et al. "Credit Card Fraud Detection Based on Unsupervised Atten-

tional Anomaly Detection Network.” *Systems* 11.6 (2023): 305.

151. Mienye, Ibomoiye Domor, and Yanxia Sun. ”A Machine Learning Method with Hybrid Feature Selection for Improved Credit Card Fraud Detection.” *Applied Sciences* 13.12 (2023): 7254.

152. Abd El-Naby, Aya, Ezz El-Din Hemdan, and Ayman El-Sayed. ”An efficient fraud detection framework with credit card imbalanced data in financial services.” *Multimedia Tools and Applications* 82.3 (2023): 4139-4160.

153. Ghaleb, Fuad A., et al. ”Ensemble Synthesized Minority Oversampling based Generative Adversarial Networks and Random Forest Algorithm for Credit Card Fraud Detection.” *IEEE Access* (2023).

154. Chung, Jiwon, and Kyungho Lee. ”Credit Card Fraud Detection: An Improved Strategy for High Recall Using KNN, LDA, and Linear Regression.” *Sensors* 23.18 (2023): 7788.

155. Ganji, Venkata Ratnam, Aparna Chaparala, and Radhika Sajja. ”Shuffled shepherd political optimization□based deep learning method for credit card fraud detection.” *Concurrency and Computation: Practice and Experience* 35.10 (2023): e7666.

156. Van Belle, Rafaël, Bart Baesens, and Jochen De Weerd. ”CATCHM: A novel network-based credit card fraud detection method using node representation learning.” *Decision Support Systems* 164 (2023): 113866.

157. Fanai, Hosein, and Hossein Abbasimehr. ”A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection.” *Expert Systems with Applications* 217 (2023): 119562.

158. Salekshahrezaee, Zahra, Joffrey L. Leevy, and Taghi M. Khoshgoftaar. ”The effect of feature extraction and data sampling on credit card fraud detection.” *Journal of Big Data* 10.1 (2023): 6.

159. Ni, Lina, et al. ”Fraud feature boosting mechanism and spiral oversampling balancing technique for credit card fraud detection.” *IEEE Transactions on Computational Social Systems* (2023).

160. Habibpour, Maryam, et al. ”Uncertainty-aware credit card fraud detection using deep learning.” *Engineering Applications of Artificial Intelligence* 123 (2023): 106248.

161. Ahmad, Hadeel, et al. ”Class balancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS).” *International Journal of Information Technology* 15.1 (2023): 325-333.

162. Strelcenia, Emilija, and Simant Prakoonwit. "Improving Classification Performance in Credit Card Fraud Detection by Using New Data Augmentation." *AI 4.1* (2023): 172-198.

163. Mienye, Ibomoiye Domor, and Yanxia Sun. "A Deep Learning Ensemble With Data Resampling for Credit Card Fraud Detection." *IEEE Access* 11 (2023): 30628-30638.

164. Strelcenia, Emilija, and Simant Prakoonwit. "A Survey on GAN Techniques for Data Augmentation to Address the Imbalanced Data Issues in Credit Card Fraud Detection." *Machine Learning and Knowledge Extraction* 5.1 (2023): 304-329.

165. Gupta, Palak, et al. "Unbalanced Credit Card Fraud Detection Data: A Machine Learning-Oriented Comparative Study of Balancing Techniques." *Procedia Computer Science* 218 (2023): 2575-2584.

166. K. Randhawa, C. K. Loo, M. Seera, C. P. Lim and A. K. Nandi, "Credit Card Fraud Detection Using AdaBoost and Majority Voting," in *IEEE Access*, vol. 6, pp. 14277-14284, 2018, doi: 10.1109/ACCESS.2018.2806420.

167. Apapan Pumsirirat and Liu Yan. Credit card fraud detection using deep learning based on auto-encoder and restricted Boltzmann machine. *International Journal of Advanced Computer Science and Applications*, 9(1):18–25, 2018. ISSN 21565570. doi: 10.14569/IJACSA.2018.09010

168. Ibtissam Benchaji, Samira Douzi, and Bouabid El Ouahidi, "Credit Card Fraud Detection Model Based on LSTM Recurrent Neural Networks," *Journal of Advances in Information Technology*, Vol. 12, No. 2, pp. 113-118, May 2021. doi: 10.12720/jait.12.2.113-118

169. Yann Lecun, Le'on Bottou, Yoshua Bengio, and Parick Haffner. A B7CEDGF HIB7PRQTSUDGQI HIB edCdSISIXvg5r ' CdQTW XvefCdS. *proc. OF THE IEEE*, 1998.

170. Zhang Z. Introduction to machine learning: k-nearest neighbors. *Ann Transl Med*. 2016 Jun;4(11):218. doi: 10.21037/atm.2016.03.37. PMID: 27386492; PMCID: PMC4916348.

171. (1998). Polynomial Regression. In: Rawlings, J.O., Pantula, S.G., Dickey, D.A. (eds) *Applied Regression Analysis*. Springer Texts in Statistics. Springer, New York, NY. <https://doi.org/10.1007/0-387-22753-9>

172. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf, "Support vector machines," in *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18-28, July-Aug. 1998, doi: 10.1109/5254.708428.

173. Kullback, Solomon, and Richard A. Leibler. "On information and sufficiency." *The annals of mathematical statistics* 22.1 (1951): 79-86.

174. Hochreiter, Sepp. (1991). *Untersuchungen zu dynamischen neuronalen Netzen*.

175. Gers, Felix A., Jürgen Schmidhuber, and Fred Cummins. "Learning to forget: Continual prediction with LSTM." *Neural computation* 12.10 (2000): 2451-2471.
176. Tania Lombrozo. The structure and function of explanations. *Trends in cognitive sciences*, 10(10): 464–470, 2006.
177. Carl Hempel and Paul Oppenheim. *Studies in the logic of explanation*. Philosophy of science, 1948.
178. William Bechtel and Adele Abrahamsen. *Explanation: A mechanist alternative*. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 2005
179. Frank Keil. *Explanation and understanding*. *Annu. Rev. Psychol.*, 2006.
181. Vincent Toubiana, Arvind Narayanan, Dan Boneh, Helen Nissenbaum, and Solon Barocas. *Adnostic: Privacy preserving targeted advertising*. 2010.
182. Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. *Fairness through awareness*. In *Innovations in Theoretical Computer Science Conference*. ACM, 2012.
183. Moritz Hardt and Kunal Talwar. *On the geometry of differential privacy*. In *ACM Symposium on Theory of Computing*. ACM, 2010.
184. Moritz Hardt, Eric Price, and Nati Srebro. *Equality of opportunity in supervised learning*. In *Advances in Neural Information Processing Systems*, 2016.
185. Frank Keil, Leonid Rozenblit, and Candice Mills. *What lies beneath? understanding the limits of understanding*. *Thinking and seeing: Visual metacognition in adults and children*, 2004.
186. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?: Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 1135–1144.
187. Alvarez-Melis, David, and Tommi S. Jaakkola. "On the robustness of interpretability methods." *arXiv preprint arXiv:1806.08049* (2018)
188. Slack, Dylan, et al. "Fooling lime and shap: Adversarial attacks on post hoc explanation methods." *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 2020