

ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

Εξηγήσιμη Τεχνητή Νοημοσύνη: Πώς οι κατηγοριοποιητές επιλέγουν τις κλάσεις σε δεδομένα χρηματιστηριακών χρονοσειρών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Του Παρασκευά Γ. Σεργιάννη
Επιβλέπων: Βασίλειος Μεγαλοοικονόμου

Πάτρα - Ιούνιος 2025

Copyright © Παρασκευάς Β. Σεργιάννης, 2025
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, η αποθήκευση και η διανομή της παρούσας εργασίας, εξ' ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, η αποθήκευση και η διανομή για σκοπό μη-κερδοσκοπικό, εκπαίδευσης ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Πατρών.

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία έχει ως αντικείμενο τη διερεύνηση και εφαρμογή μεθόδων Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI) για την ερμηνεία των αποφάσεων ενός μοντέλου βαθιάς μάθησης σε προβλήματα ταξινόμησης χρηματοοικονομικών χρονοσειρών. Ειδικότερα, εστιάζει στην ανάλυση ενός δικτύου Μακράς Βραχυπρόθεσμης Μνήμης (LSTM), το οποίο εκπαιδεύεται να προβλέπει σήματα συναλλαγών (Αγορά, Πώληση, Διακράτηση) για τη μετοχή της Apple ($AAPL)

Βασικός στόχος είναι διττός: αφενός, η ανάπτυξη ενός αποτελεσματικού μοντέλου LSTM ικανού να αναγνωρίζει δυνητικά κερδοφόρες συνθήκες συναλλαγών , και αφετέρου, ο κύριος στόχος, η ενδελεχής ανάλυση των εσωτερικών μηχανισμών λήψης αποφάσεών του. Η έρευνα αποσκοπεί να απαντήσει στο κρίσιμο ερώτημα του γιατί το μοντέλο προτείνει μια συγκεκριμένη δράση, διασφαλίζοντας ότι οι προβλέψεις του βασίζονται σε οικονομικά εύλογα σήματα και όχι σε στατιστικές ανωμαλίες.

Για την επίτευξη των παραπάνω, αρχικά αναπτύχθηκε ένα πολυ-εισοδικό (multi-input) μοντέλο LSTM που επεξεργάζεται ταυτόχρονα βραχυπρόθεσμα, μεσοπρόθεσμα και μακροπρόθεσμα χρονικά παράθυρα δεδομένων. Στο πλαίσιο της εργασίας, διερευνήθηκαν και συγκρίθηκαν δύο διακριτές στρατηγικές σήμανσης των ιστορικών δεδομένων: μία προσαρμοσμένη στρατηγική βασισμένη σε κανόνες τεχνικής ανάλυσης και η μέθοδος του «Τριπλού Φράγματος» (Triple Barrier). Για την ερμηνεία του εκπαιδευμένου μοντέλου, εφαρμόστηκε ένα σύνολο πέντε διαφορετικών τεχνικών XAI: SHAP, LIME, DeepLIFT, Integrated Gradients και Permutation Feature Importance, επιτρέποντας μια σφαιρική και διασταυρωμένη αξιολόγηση της συμπεριφοράς του.

Η αξία της παρούσας εργασίας έγκειται στην ανάδειξη του καθοριστικού ρόλου που διαδραματίζει η στρατηγική σήμανσης στη διαμόρφωση της «σκέψης» του μοντέλου, καθώς και στην πρακτική εφαρμογή και συγκριτική αξιολόγηση πολλαπλών μεθόδων XAI. Τα ευρήματα παρέχουν χρήσιμη γνώση για τους παράγοντες που οδηγούν τις αλγοριθμικές προβλέψεις και θέτουν τις βάσεις για την ανάπτυξη πιο διαφανών, αξιόπιστων και ελέγξιμων συστημάτων υποστήριξης επενδυτικών αποφάσεων.

Λέξεις Κλειδιά: Εξηγήσιμη Τεχνητή Νοημοσύνη, XAI, Μηχανική Μάθηση, Δίκτυα LSTM, Χρηματοοικονομικές Χρονοσειρές, Πρόβλεψη Μετοχών, SHAP, LIME, DeepLIFT, Integrated Gradients, Permutation Feature Importance, Στρατηγικές Σήμανσης, Triple Barrier Method.

“If you can't explain it simply, you don't understand it well enough.”

— Albert Einstein

Ευχαριστίες

Φτάνοντας στο τέλος αυτού του σημαντικού ακαδημαϊκού κύκλου, αισθάνομαι την ανάγκη να ευχαριστήσω τους ανθρώπους που στάθηκαν δίπλα μου.

Εκφράζω τη βαθιά μου ευγνωμοσύνη στον επιβλέποντα καθηγητή μου, κ Βασίλειο Μεγαλοοικονόμου. Η εμπιστοσύνη που μου έδειξε, οι εποικοδομητικές μας συζητήσεις και η συνεχής του ενθάρρυνση αποτέλεσαν πηγή έμπνευσης και ήταν θεμελιώδους σημασίας για την ολοκλήρωση αυτής της προσπάθειας.

Θα ήταν παράλειψή μου να μην αναφερθώ στην οικογένειά μου. Ευχαριστώ από καρδιάς τους γονείς μου για την αγάπη, την κατανόηση στις δύσκολες στιγμές και την πίστη τους σε εμένα. Χωρίς τη δική τους στήριξη, τίποτα δεν θα ήταν το ίδιο.

- Σεργιάννης Γ. Παρασκευάς

Πίνακας Περιεχομένων

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ....1
ΠΕΡΙΛΗΨΗ.... 3
ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ.... 11
1.1 Στόχος Διπλωματικής.... 11
1.2 Περιγραφή Προβλήματος.... 12
1.3 Προτεινόμενη Λύση.... 15
ΚΕΦΑΛΑΙΟ 2: ΒΑΣΙΚΕΣ ΓΝΩΣΕΙΣ.... 19
2.1 Ιστορία Χρηματιστηριακών Αγορών.... 19
2.2 Ιστορία Μετοχής Apple ($AAPL).... 23
2.3 Πώς πραγματοποιούνται τρέχουσες προβλέψεις της χρηματιστηριακής αγοράς από αναλυτές.... 26
2.4 Εξηγησιμότητα και η Σημασία της στις Προβλέψεις της Χρηματιστηριακής Αγοράς.... 30
ΚΕΦΑΛΑΙΟ 3: ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ.... 34
3.1 Μοντέλα Μηχανικής Μάθησης.... 34
3.2 Μοντέλα Μηχανικής Μάθησης για Προβλέψεις Δεδομένων Χρονοσειρών.... 36
3.3 Μοντέλα LSTM: Εις Βάθος Ανάλυση, Λειτουργία και Μαθηματική Θεμελίωση.... 39
3.4 Εφαρμογή των LSTMs στις Προβλέψεις της Χρηματιστηριακής Αγοράς.... 43
3.5 Πολυ-εισοδικά (Multi-input) LSTMs και η Λογική Πίσω από Αυτά.... 45
ΚΕΦΑΛΑΙΟ 4: ΕΞΗΓΗΣΙΜΟΤΗΤΑ.... 49
4.1 Εισαγωγή στην Εξηγησιμότητα.... 49
4.2 Κατηγορίες Μεθόδων Εξηγησιμότητας.... 51
4.3 SHAP (SHapley Additive exPlanations).... 54
4.3.1 Θεωρητική Θεμελίωση.... 54
4.3.2 Μαθηματικό Υπόβαθρο.... 55
4.3.3 Πλεονεκτήματα του SHAP.... 56
4.3.4 Περιορισμοί του SHAP.... 57
4.4 LIME (Local Interpretable Model-agnostic Explanations).... 58
4.4.1 Θεωρητική Θεμελίωση.... 58
4.4.2 Μαθηματικό Υπόβαθρο.... 59
4.4.3 Πλεονεκτήματα του LIME.... 61
4.4.4 Περιορισμοί του LIME.... 62
4.5 DeepLIFT (Deep Learning Important FeaTures).... 63
4.5.1 Θεωρητική Θεμελίωση.... 63
4.5.2 Μαθηματικό Υπόβαθρο.... 64
4.5.3 Πλεονεκτήματα του DeepLIFT.... 66
4.5.4 Περιορισμοί του DeepLIFT.... 66
4.6 Integrated Gradients (Ολοκληρωμένες Κλίσεις).... 67

4.6.1 Θεωρητική Θεμελίωση..... 67
4.6.2 Μαθηματικό Υπόβαθρο..... 68
4.6.3 Πλεονεκτήματα του Integrated Gradients..... 69
4.6.4 Περιορισμοί του Integrated Gradients..... 70
4.7 Permutation Feature Importance (Σημασία Χαρακτηριστικών μέσω Μετάθεσης)..... 71
4.7.1 Θεωρητική Θεμελίωση..... 71
4.7.2 Μαθηματικό Υπόβαθρο..... 72
4.7.3 Πλεονεκτήματα του Permutation Feature Importance..... 73
4.7.4 Περιορισμοί του Permutation Feature Importance..... 73
ΚΕΦΑΛΑΙΟ 5: ΜΕΘΟΔΟΛΟΓΙΑ..... 75
5.1 Συλλογή Δεδομένων..... 75
5.1.2 Σήμανση Δεδομένων (Labeling)..... 76
5.2 Οπτικοποίηση Δεδομένων..... 80
5.3 Προεπεξεργασία Δεδομένων..... 81
5.4 Αρχιτεκτονική Μοντέλου LSTM..... 83
5.5 Προσομοίωση Στρατηγικής Συναλλαγών..... 87
5.5.1 Λογική Σεναρίου και Παραδοχές..... 87
5.5.2 Εκτέλεση Προσομοίωσης και Αποτελέσματα..... 88
5.6 Εξηγησιμότητα Μοντέλου (XAI)..... 90
5.6.1 KernelSHAP..... 90
5.6.2 LIME (Local Interpretable Model-Agnostic Explanations)..... 93
5.6.3 DeepLIFT (Deep Learning Important FeaTures)..... 95
5.6.4 Integrated Gradients (Ενοποιημένες Κλίσεις)..... 97
5.6.5 Permutation Feature Importance (Σημαντικότητα Χαρακτηριστικών με Μεταθέσεις)..... 98
ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ..... 101
6.1 Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης (Custom Labeling Strategy)..... 101
6.1.1 Περιγραφή Πειράματος..... 101
6.1.2 Ευρήματα Πειράματος Custom Labeling..... 106
6.1.2.1 DeepLIFT..... 106
6.1.2.2 Integrated Gradients..... 107
6.1.2.3 Kernel SHAP..... 109
6.1.2.4 LIME..... 115
6.1.2.5 Permutation Feature Importance..... 121
6.2 Πείραμα με Στρατηγική Σήμανσης Τριπλού Φράγματος (Triple Barrier Labeling Strategy)..... 125
6.2.1 Περιγραφή Πειράματος..... 125
6.2.2 Ευρήματα Πειράματος Triple Barrier Labeling..... 138
6.2.2.1 DeepLIFT..... 138

6.2.2.2 Integrated Gradients....139
6.2.2.3 Kernel SHAP....141
6.2.2.4 LIME....146
6.2.2.5 Permutation Feature Importance....152
ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ....157
7.1 Συγκριτική Αξιολόγηση Πειραμάτων και Εξηγησιμότητας....157
7.2 Ερευνητική Συνεισφορά και Μελλοντικές Κατευθύνσεις....159
ΚΕΦΑΛΑΙΟ 8: ΒΙΒΛΙΟΓΡΑΦΙΑ....161

Πίνακας Εικόνων

Εικόνα 3.1: Αρχιτεκτονική ενός μοντέλου LSTM....42
Εικόνα 4.1: Τοπική Ερμηνεία με LIME....61
Εικόνα 5.1: Αρχιτεκτονική του τελικού μοντέλου LSTM....86
Εικόνα 6.1: Heatmap σημαντικότητας DeepLIFT στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....106
Εικόνα 6.2: Heatmap σημαντικότητας Integrated Gradients στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....107
Εικόνα 6.4: SHAP Summary Plot στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....111
Εικόνα 6.5: SHAP Time Importance Plot (50 days window) στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....112
Εικόνα 6.6: SHAP Time Importance Plot (30 days window) στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....113
Εικόνα 6.7: SHAP Time Importance Plot (10 days window) στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....114
Εικόνα 6.8: LIME Aggregate Importance Plot for Buy action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....115
Εικόνα 6.9: LIME Aggregate Importance Plot for Hold action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....116
Εικόνα 6.10: LIME Aggregate Importance Plot for Sell action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....117
Εικόνα 6.11: LIME Time Specific Importance Plot for Buy action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....118
Εικόνα 6.12: LIME Time Specific Importance Plot for Hold action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....119
Εικόνα 6.13: LIME Time Specific Importance Plot for Sell action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....120
Εικόνα 6.14: PFI Importance Plot στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης....122

Εικόνα 6.15: PFI Time Importance Plot στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.....123
Εικόνα 6.16: Η trading απόδοση του 1ου μοντέλου, σε σύγκριση με την απλή επένδυση (Buy & Hold) 100$ από 1/1/2000.....126
Εικόνα 6.17: Η trading απόδοση του 2ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.....127
Εικόνα 6.18: Η trading απόδοση του 3ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.....128
Εικόνα 6.19: Η trading απόδοση του 4ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.....129
Εικόνα 6.20: Η trading απόδοση του 5ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.....130
Εικόνα 6.23: Η trading απόδοση του 8ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.....133
Εικόνα 6.24: Η trading απόδοση του 9ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.....134
Εικόνα 6.26: Συνολική αξιολόγηση της απόδοσης των 10 LSTM μοντέλων..... 136
Εικόνα 6.27: Heatmap σημαντικότητας DeepLIFT στο Πείραμα με Triple Barrier Labeling.....138
Εικόνα 6.28: Heatmap σημαντικότητας Integrated Gradients στο Πείραμα με Triple Barrier Labeling.....139
Εικόνα 6.29: SHAP Feature Importance Graph στο Πείραμα με Triple Barrier Labeling.....141
Εικόνα 6.30: SHAP Summary Plot στο Πείραμα με Triple Barrier Labeling..... 143
Εικόνα 6.31: SHAP Time Importance Plot (50 days window) στο Πείραμα με με Triple Barrier Labeling..... 143
Εικόνα 6.32: SHAP Time Importance Plot (30 days window) στο Πείραμα με με Triple Barrier Labeling..... 144
Εικόνα 6.33: SHAP Time Importance Plot (10 days window) στο Πείραμα με Triple Barrier Labeling.....145
Εικόνα 6.34: LIME Aggregate Importance Plot for Buy action στο Πείραμα με Triple Barrier Labeling.....146
Εικόνα 6.35: LIME Aggregate Importance Plot for Hold action στο Πείραμα με Triple Barrier Labeling.....147
Εικόνα 6.36: LIME Aggregate Importance Plot for Sell action στο Πείραμα με Triple Barrier Labeling.....148
Εικόνα 6.37: LIME Time Specific Importance Plot for Buy action στο Πείραμα με Triple Barrier Labeling..... 149
Εικόνα 6.38: LIME Time Specific Importance Plot for Hold action στο Πείραμα με Triple Barrier Labeling..... 150
Εικόνα 6.39: LIME Time Specific Importance Plot for Sell action στο Πείραμα

με Triple Barrier Labeling..... 151
Εικόνα 6.40: PFI Importance Plot στο Πείραμα με Triple Barrier Labeling... 152
Εικόνα 6.41: PFI Time Varying Importance Plot στο Πείραμα με Triple Barrier Labeling..... 153

Πίνακας Πινάκων

Πίνακας 2.2: Ιστορικό Stock Splits της Apple (AAPL)..... 26
Πίνακας 2.3: Σύγκριση Τεχνικής και Θεμελιώδους Ανάλυσης..... 30
Πίνακας 4.1: Κατηγοριοποίηση μεθόδων εξηγησιμότητας..... 53

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

Η παρούσα διπλωματική εργασία διερευνά το πεδίο της Εξηγήσιμης Τεχνητής Νοημοσύνης (Explainable Artificial Intelligence - XAI) στο πλαίσιο της πρόβλεψης χρηματοοικονομικών χρονοσειρών. Ειδικότερα, εστιάζει στην κατανόηση του τρόπου με τον οποίο πολύπλοκα μοντέλα μηχανικής μάθησης, όπως τα δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (Long Short-Term Memory - LSTM), λαμβάνουν αποφάσεις κατηγοριοποίησης σημάτων συναλλαγών (αγορά, πώληση, διακράτηση) για μετοχές. Ως μελέτη περίπτωσης χρησιμοποιούνται ιστορικά δεδομένα της μετοχής της Apple Inc. ($AAPL). Η έρευνα αυτή αποσκοπεί στο να ρίξει φως στους εσωτερικούς μηχανισμούς των εν λόγω «μαύρων κουτιών», εφαρμόζοντας και αξιολογώντας ποικίλες μεθόδους XAI. Η σημασία της εξηγησιμότητας είναι ιδιαίτερα κρίσιμη σε τομείς υψηλού ρίσκου, όπως οι χρηματοοικονομικές επενδύσεις, όπου η κατανόηση της αιτίας πίσω από τις αλγοριθμικές προτάσεις μπορεί να ενισχύσει την εμπιστοσύνη, να βελτιώσει τη λήψη αποφάσεων και να διευκολύνει τη συμμόρφωση με κανονιστικά πλαίσια.

1.1 Στόχος Διπλωματικής

Ο κεντρικός στόχος της παρούσας διπλωματικής εργασίας είναι διττός και πολυεπίπεδος. Αφενώς, στόχος είναι η ανάπτυξη και βελτιστοποίηση της ικανότητας πρόβλεψης ενός εξειδικευμένου μοντέλου μηχανικής μάθησης, συγκεκριμένα ενός δικτύου LSTM, για την σχεδίαση εν δυνάμη κερδοφόρων στρατηγικών συναλλαγών που αφορούν την μετοχή της Apple ($AAPL). Η επιλογή του μοντέλου LSTM δεν είναι τυχαία, καθώς τα δίκτυα αυτά έχουν επιδείξει σημαντικές επιδόσεις στην ανάλυση και πρόβλεψη δεδομένων χρονοσειρών, χάρη στην ικανότητά τους να συλλαμβάνουν μακροπρόθεσμες εξαρτήσεις και μη γραμμικά πρότυπα, τα οποία είναι χαρακτηριστικά στις χρηματοοικονομικές αγορές.

Αφετέρου, και κύριος στόχος της εργασίας είναι η ενδελεχής αξιολόγηση της αξιοπιστίας και της εγκυρότητας των προβλέψεων που παράγει το μοντέλο LSTM. Αυτή η αξιολόγηση δεν θα περιοριστεί σε παραδοσιακές μετρικές απόδοσης (όπως η ακρίβεια ή το F1-score), αλλά θα επεκταθεί στην ερμηνεία των εσωτερικών μηχανισμών λήψης αποφάσεων του μοντέλου. Για τον σκοπό αυτό, θα αξιοποιηθεί ένα σύνολο πέντε σύγχρονων μεθόδων Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI): SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), DeepLIFT (Deep Learning Important FeaTures), Integrated Gradients (Ενοποιημένες Κλίσεις) και Permutation Feature Importance (Σπουδαιότητα Χαρακτηριστικών μέσω Μεταθέσεων). Η εφαρμογή αυτών των μεθόδων αποσκοπεί

στο να απαντήσει στο κρίσιμο ερώτημα του *γιατί* το μοντέλο προτείνει μια συγκεκριμένη ενέργεια (αγορά, πώληση ή διακράτηση) και όχι απλώς στο *τι* προτείνει.

Η δυνατότητα εξήγησης των αποφάσεων του μοντέλου είναι θεμελιώδους σημασίας. Επιτρέπει τον έλεγχο του κατά πόσον το μοντέλο βασίζει τις προβλέψεις του σε οικονομικά εύλογα και αναγνωρίσιμα σήματα της αγοράς ή αν, αντίθετα, εκμεταλλεύεται τυχόν στατιστικές ανωμαλίες, θόρυβο στα δεδομένα εκπαίδευσης ή ακόμα και ακούσια διαρροή πληροφορίας (data leakage). Η κατανόηση αυτή είναι απαραίτητη για την αποφυγή ψευδώς ενθαρρυντικών αποτελεσμάτων που δεν θα γενικεύονταν σε πραγματικές συνθήκες αγοράς.

Επιπλέον, η εξηγησιμότητα διαδραματίζει καίριο ρόλο στην οικοδόμηση εμπιστοσύνης από την πλευρά του τελικού χρήστη (π.χ. ενός επενδυτή ή διαχειριστή κεφαλαίων) προς το προτεινόμενο σύστημα τεχνητής νοημοσύνης. Σε τομείς όπου το οικονομικό ρίσκο είναι υψηλό και οι αποφάσεις πρέπει να είναι αιτιολογημένες και υπεύθυνες, όπως οι χρηματοοικονομικές συναλλαγές, η αποδοχή και υιοθέτηση συστημάτων ΤΝ εξαρτάται σε μεγάλο βαθμό από την ικανότητά τους να παρέχουν διαφανείς και κατανοητές εξηγήσεις για τις προτάσεις τους. Ένα σύστημα που λειτουργεί ως «μαύρο κουτί», όσο ακριβές κι αν είναι στις προβλέψεις του υπό ελεγχόμενες συνθήκες, είναι δύσκολο να εμπνεύσει την απαραίτητη εμπιστοσύνη για την εφαρμογή του σε πραγματικές επενδυτικές στρατηγικές. Η XAI, παρέχοντας αυτή τη διαφάνεια, μπορεί να μετατρέψει ένα πολύπλοκο μοντέλο από ένα ακατανόητο εργαλείο σε έναν κατανοητό και ελέγξιμο συνεργάτη στη διαδικασία λήψης αποφάσεων.

1.2 Περιγραφή Προβλήματος

Η πρόβλεψη των βραχυπρόθεσμων κινήσεων της τιμής μιας μετοχής αποτελεί ένα από τα πιο διαβόητα δύσκολα προβλήματα στον χώρο της χρηματοοικονομικής και της μηχανικής μάθησης. Η δυσκολία αυτή πηγάζει από πολλαπλούς παράγοντες, κυριότεροι των οποίων είναι η εγγενής στοχαστική φύση των χρηματοοικονομικών αγορών, η παρουσία σημαντικού θορύβου στα δεδομένα τιμών και η επίδραση ενός τεράστιου αριθμού, συχνά μη παρατηρήσιμων ή μη ποσοτικοποιήσιμων, παραγόντων (όπως μακροοικονομικές ειδήσεις, γεωπολιτικές εξελίξεις, ψυχολογία των επενδυτών, ακόμα και τυχαία γεγονότα).

Στο πλαίσιο της παρούσας διπλωματικής εργασίας, το πρόβλημα ορίζεται συγκεκριμένα ως ένα πρόβλημα multi-class classification με ορίζοντα πρόβλεψης μίας ημέρας. Για κάθε διαθέσιμο ημερήσιο σημείο δεδομένων της μετοχής $AAPL,

στόχος είναι το μοντέλο μηχανικής μάθησης να αποφασίσει ποια είναι η βέλτιστη ενέργεια που θα έπρεπε να αναλάβει ένας επενδυτής για την *επόμενη* ημέρα συναλλαγών. Οι πιθανές ενέργειες (κλάσεις) είναι τρεις:

1. **Αγορά (Buy):** Το μοντέλο προβλέπει ότι η ανάληψη θέσης αγοράς στη μετοχή είναι η ενδεδειγμένη ενέργεια.
2. **Διακράτηση (Hold):** Το μοντέλο προβλέπει ότι η διατήρηση της υπάρχουσας θέσης (ή η μη ανάληψη νέας θέσης) είναι η βέλτιστη στρατηγική.
3. **Πώληση (Sell):** Το μοντέλο προβλέπει ότι η ρευστοποίηση της υπάρχουσας θέσης πώλησης (ή η αποφυγή αγοράς) είναι η κατάλληλη ενέργεια.

Μία από τις σημαντικότερες και πιο θεμελιώδεις προκλήσεις στην αντιμετώπιση αυτού του προβλήματος είναι ο ορισμός του κατάλληλου σήματος-στόχου (target variable) για την εκπαίδευση του μοντέλου μηχανικής μάθησης. Σε αντίθεση με πολλά άλλα προβλήματα κατηγοριοποίησης (π.χ., αναγνώριση εικόνων, όπου η ετικέτα "γάτα" ή "σκύλος" είναι αντικειμενική), δεν υπάρχει εκ των προτέρων ένα καθολικά αποδεκτό ή "σωστό" σηματοδοτημένο σύνολο δεδομένων που να υποδεικνύει ποια ήταν η "ιδανική" ενέργεια (Buy, Hold, Sell) για κάθε ημέρα στο παρελθόν. Η βέλτιστη ενέργεια εξαρτάται από τον επενδυτικό ορίζοντα, την ανοχή στον κίνδυνο και, κυρίως, από τις μελλοντικές κινήσεις της τιμής, οι οποίες είναι εξ ορισμού άγνωστες τη στιγμή της απόφασης.

Αρχικά, στην προσπάθεια δημιουργίας ετικετών για την εκπαίδευση, εξετάστηκε μια ευρετική προσέγγιση (heuristic approach). Αυτή περιελάμβανε τη δημιουργία κανόνων βασισμένων σε γνωστούς τεχνικούς δείκτες και αρχές της τεχνικής ανάλυσης. Για παράδειγμα, μια ημέρα θα μπορούσε να χαρακτηριστεί ως σήμα "Buy" εάν πληρούνταν συγκεκριμένες συνθήκες, όπως:

- Ο δείκτης σχετικής ισχύος (Relative Strength Index - RSI) να υποδείκνυε συνθήκες υπερπώλησης (π.χ. RSI < 40) και ταυτόχρονα ο δείκτης ορμής (Momentum) να είναι θετικός.
- Ο βραχυπρόθεσμος κινητός μέσος όρος (π.χ. SMA10) περνάει πάνω από τον μακροπρόθεσμο (π.χ. SMA30), κάτι που συνήθως δείχνει ανοδική τάση, ενώ ταυτόχρονα παρατηρείται αύξηση στον δείκτη On-Balance Volume (OBV).
- Η τιμή κλεισίματος της μετοχής να πέφτει κάτω από το κάτω όριο των ζωνών Bollinger (Bollinger Bands) με ταυτόχρονη αύξηση του όγκου συναλλαγών.

Αντίστοιχοι κανόνες θα μπορούσαν να οριστούν για τον χαρακτηρισμό μιας ημέρας ως σήμα "Sell" (π.χ., RSI>60 με αρνητικό Momentum, πτωτική διάσχιση των SMAs με πτώση του OBV, κλείσιμο πάνω από το άνω όριο των Bollinger Bands με υψηλό όγκο). Επιπρόσθετα, στην ευρετική αυτή προσέγγιση, θα μπορούσαν να

συνυπολογιστούν και ευρύτεροι παράγοντες της αγοράς, όπως η ημερήσια απόδοση του γενικού δείκτη της αγοράς (π.χ., S&P 500) ή ενός σχετικού κλαδικού δείκτη (π.χ., τεχνολογίας), για την ενίσχυση ή αποδυνάμωση των σημάτων.

Παρόλο που οι παραπάνω κανόνες βασίζονται σε καθιερωμένες αρχές της χρηματοοικονομικής τεχνικής ανάλυσης, η προσέγγιση αυτή ενέχει δύο σημαντικά προβλήματα:

1. Οι κανόνες αυτοί είναι εκ φύσεως στατικοί και συχνά υποκειμενικοί, αντανακλώντας τις παραδοχές και τις προτιμήσεις του ατόμου που τους σχεδιάζει. Είναι πιθανό να μην καλύπτουν όλες τις πιθανές συνθήκες της αγοράς ή να μην είναι βέλτιστοι για όλες τις περιόδους. Η αγορά εξελίσσεται, και κανόνες που μπορεί να ήταν αποτελεσματικοί στο παρελθόν ενδέχεται να μην είναι στο μέλλον.
2. Η εφαρμογή τέτοιων κανόνων συχνά οδηγεί στην παραγωγή σημαντικού θορύβου στις ετικέτες. Είναι συχνό φαινόμενο, ειδικά σε οριακές ή πλάγιες καταστάσεις της αγοράς, διαφορετικοί δείκτες να παρέχουν αντικρουόμενα σήματα (π.χ., κάποιοι να υποδεικνύουν αγορά ενώ άλλοι πώληση για την ίδια ημέρα). Ένα μοντέλο μηχανικής μάθησης που θα εκπαιδευόταν σε τέτοια θορυβώδη και δυνητικά αντιφατικά labels, πολύ πιθανό να εμφάνιζε μειωμένη απόδοση και ικανότητα γενίκευσης.

Συνοψίζοντας, η πρόκληση έγκειται στην εκπαίδευση ενός μοντέλου ικανού να μαθαίνει από ιστορικά δεδομένα (τιμές, όγκοι, τεχνικοί δείκτες, μακροοικονομικοί παράγοντες κ.λπ.) πώς να προβλέπει αυτά τα "ιδανικά" σήματα Buy/Hold/Sell. Στη συνέχεια, το εκπαιδευμένο μοντέλο θα χρησιμοποιηθεί για την παραγωγή προβλέψεων σε νέα, μη παρατηρηθέντα δεδομένα (μελλοντικές ημέρες), με απώτερο στόχο τη βελτιστοποίηση του κέρδους από τις συναλλαγές. Ταυτόχρονα, και αυτό είναι κεντρικής σημασίας για την παρούσα εργασία, απαιτείται το μοντέλο αυτό να είναι επεξηγήσιμο. Η εξηγησιμότητα είναι απαραίτητη για να διαπιστωθεί εάν οι προβλέψεις του βασίζονται σε ορθολογικά, οικονομικά εύλογα πρότυπα της αγοράς ή εάν παρατηρούνται ανεπιθύμητες συμπεριφορές που χρήζουν περαιτέρω διερεύνησης, όπως πιθανή διαρροή δεδομένων (data leakage) από το μέλλον προς το παρελθόν κατά την εκπαίδευση, ή υπερβολική προσαρμογή (overfitting) σε συγκεκριμένα, ιδιόμορφα μοτίβα των δεδομένων εκπαίδευσης που δεν έχουν ικανότητα γενίκευσης. Η κατανόηση αυτών των πτυχών είναι κρίσιμη για την ανάπτυξη ενός πραγματικά αξιόπιστου και χρήσιμου συστήματος πρόβλεψης.

1.3 Προτεινόμενη Λύση

Για την αντιμετώπιση του σύνθετου προβλήματος που περιγράφηκε στην προηγούμενη ενότητα, σχεδιάστηκε και υλοποιήθηκε μια ολοκληρωμένη λύση που αποτελείται από πέντε κύρια συστατικά, τα οποία καλύπτουν τον πλήρη κύκλο από τη συλλογή δεδομένων έως την ερμηνεία και αξιολόγηση των προβλέψεων.

1. Συλλογή και Επεξεργασία Δεδομένων:
   Πρώτο βήμα είναι η εηασφάλιση της ποιότητας και της πληρότητας των δεδομένων. Για τον σκοπό αυτό, κατασκευάστηκε ένα σύνολο δεδομένων (dataset) που περιλαμβάνει ιστορικά στοιχεία για τη μετοχή της Apple (AAPL). Αυτά τα δεδομένα περιλαμβάνουν τις βασικές ημερήσιες χρηματιστηριακές τιμές (τιμή ανοίγματος, υψηλότερη τιμή ημέρας, χαμηλότερη τιμή ημέρας, τιμή κλεισίματος) καθώς και τον ημερήσιο όγκο συναλλαγών. Πέραν αυτών των βασικών δεδομένων, υπολογίστηκε ένας αριθμός τεχνικών δεικτών, οι οποίοι χρησιμοποιούνται ευρέως στην τεχνική ανάλυση για την πρόβλεψη των τάσεων της αγοράς (π.χ., κινητοί μέσοι όροι, RSI, MACD, Bollinger Bands, OBV κ.ά.). Επιπλέον, για να δοθεί στο μοντέλο μια ευρύτερη εικόνα του οικονομικού περιβάλλοντος, ενσωματώθηκαν στο dataset μακροοικονομικοί δείκτες και δείκτες της αγοράς, όπως η πορεία του δείκτη S&P 500 (ως μέτρο της γενικής τάσης της αμερικανικής αγοράς), ο δείκτης μεταβλητότητας VIX (γνωστός και ως "δείκτης φόβου"), και ένας σχετικός κλαδικός δείκτης τεχνολογίας. Η συλλογή αυτών των ποικίλων δεδομένων στοχεύει στην παροχή μιας πολυδιάστατης πληροφόρησης στο μοντέλο.
2. Στρατηγική Labeling:
   Όπως αναφέρθηκε προηγουμένως, ο ορισμός των ετικετών (Buy, Hold, Sell) αποτελεί μια σημαντική πρόκληση. Για το labeling χρησιμοποιήθηκαν 2 ξεχωριστές στρατηγικές.
   i) Η πρώτη στρατηγική labeling βασίζεται σε προκαθορισμένους κανόνες που συνδυάζουν τεχνικούς δείκτες (RSI, MACD, κινητούς μέσους όρους, Bollinger Bands) με δεδομένα αγοράς και κλάδου για τη δημιουργία των σημάτων Buy, Hold, Sell. Η μέθοδος αυτή ορίζει πέντε διακριτές συνθήκες για κάθε κατηγορία δράσης, επιδιώκοντας να δημιουργήσει ερμηνεύσιμα σήματα συναλλαγών που μπορούν να αιτιολογηθούν μέσω των τεχνικών explainability.
   ii) Αντί για την ευρετική προσέγγιση που βασίζεται σε στατικούς κανόνες τεχνικών δεικτών, στην παρούσα εργασία υιοθετήθηκε η προσέγγιση "Triple Barrier" (Τριπλό Φράγμα) για τη δημιουργία των action labels στο σύνολο δεδομένων εκπαίδευσης. Αυτή η μέθοδος, που προτάθηκε από τον Lopez de Prado (2018), ορίζει τις ετικέτες με βάση την κίνηση της τιμής της μετοχής μετά από ένα συγκεκριμένο σημείο στο χρόνο, λαμβάνοντας υπόψη τρία πιθανά

ενδεχόμενα: (α) η τιμή φτάνει ένα άνω όριο (profit target), (β) η τιμή φτάνει ένα κάτω όριο (stop loss), ή (γ) παρέρχεται ένας συγκεκριμένος χρονικός ορίζοντας χωρίς να έχει επιτευχθεί κανένα από τα δύο όρια. Θεωρείται ότι αυτή η προσέγγιση παράγει πιο αξιόπιστα, αντικειμενικά και τελικά πιο χρήσιμα σήματα συναλλαγών σε σύγκριση με απλούς, συχνά υποκειμενικούς, κανόνες, επιτρέποντας στο μοντέλο να μάθει να αναγνωρίζει καταστάσεις όπου πράγματι θα ήταν επωφελής μια συγκεκριμένη ενέργεια.

3. Μοντέλο Μηχανικής Μάθησης (LSTM):
   Ως βασικό μοντέλο για την πρόβλεψη των σημάτων συναλλαγών επιλέχθηκε ένα νευρωνικό δίκτυο τύπου Long Short-Term Memory (LSTM). Τα δίκτυα LSTM είναι μια ειδική κατηγορία επαναληπτικών νευρωνικών δικτύων (RNNs) που είναι ιδιαίτερα κατάλληλα για την επεξεργασία και πρόβλεψη δεδομένων χρονοσειρών. Η αρχιτεκτονική τους τους επιτρέπει να μαθαίνουν ακολουθιακά πρότυπα και μακροπρόθεσμες εξαρτήσεις μεταξύ χρονικά διατεταγμένων δεδομένων, αντιμετωπίζοντας αποτελεσματικά το πρόβλημα της εξασθένισης ή έκρηξης των κλίσεων (vanishing/exploding gradients) που ταλαιπωρεί τα απλά RNNs. Το μοντέλο LSTM τροφοδοτείται με μια "παραθυρική" ιστορική ακολουθία (sequence) δεδομένων από τις πιο πρόσφατες ημέρες και παράγει την πρόβλεψη για την συνιστώμενη δράση την επόμενη ημέρα. Ένα ιδιαίτερο χαρακτηριστικό της υλοποίησης στην παρούσα εργασία είναι η χρήση πολλαπλών χρονικών "παραθύρων" ως παράλληλες είσοδοι στο μοντέλο. Αυτό επιτρέπει στο μοντέλο να αξιολογεί ταυτόχρονα βραχυπρόθεσμα, μεσοπρόθεσμα και πιο μακροπρόθεσμα μοτίβα στην πορεία της μετοχής πριν καταλήξει σε μια απόφαση. Οι διαφορετικές αυτές ακολουθίες εισόδου επεξεργάζονται από ξεχωριστές υπο-μονάδες LSTM και οι αναπαραστάσεις που παράγουν συνδυάζονται στη συνέχεια εντός του νευρωνικού δικτύου για την τελική πρόβλεψη. Αυτή η πολυεπίπεδη χρονική ανάλυση μιμείται τον τρόπο με τον οποίο οι αναλυτές συχνά συνδυάζουν πληροφορίες από διαφορετικούς χρονικούς ορίζοντες.
4. Εξηγησιμότητα Προβλέψεων (XAI):
   Μετά την εκπαίδευση του μοντέλου LSTM, και δεδομένης της φύσης του ως "μαύρο κουτί", εφαρμόστηκαν πέντε διακριτές μέθοδοι Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI) με σκοπό την ερμηνεία των προβλέψεών του. Κάθε μία από αυτές τις μεθόδους προσφέρει μια διαφορετική οπτική γωνία και προσέγγιση στο πώς ο ταξινομητής LSTM λαμβάνει τις αποφάσεις του:
   - **SHAP (SHapley Additive exPlanations):** Μια μέθοδος βασισμένη στη θεωρία παιγνίων που αποδίδει σε κάθε χαρακτηριστικό μια τιμή που αντικατοπτρίζει τη συμβολή του στην πρόβλεψη.
   - **LIME (Local Interpretable Model-agnostic Explanations):** Model-agnostic

τεχνική που εξηγεί μεμονωμένες προβλέψεις προσεγγίζοντας τοπικά το πολύπλοκο μοντέλο με ένα απλούστερο, ερμηνεύσιμο μοντέλο.

- **DeepLIFT (Deep Learning Important FeaTures):** Ειδικά σχεδιασμένη για βαθιά νευρωνικά δίκτυα που συγκρίνει την ενεργοποίηση των νευρώνων με μια ενεργοποίηση αναφοράς για να αποδώσει σκορ συμβολής.
- **Integrated Gradients (Ενοποιημένες Κλίσεις):** Αποδίδει σημασία στα χαρακτηριστικά εισόδου ολοκληρώνοντας τις κλίσεις (gradients) κατά μήκος μιας διαδρομής από μια είσοδο αναφοράς (baseline) στην πραγματική είσοδο.
- **Permutation Feature Importance (Σπουδαιότητα Χαρακτηριστικών μέσω Μεταθέσεων):** Επίσης model-agnostic μέθοδος που αξιολογεί τη σπουδαιότητα ενός χαρακτηριστικού μετρώντας την πτώση στην απόδοση του μοντέλου όταν οι τιμές αυτού του χαρακτηριστικού ανακατεύονται τυχαία. Μέσω αυτών των τεχνικών, είναι δυνατόν να προσδιοριστεί, για κάθε μεμονωμένη πρόβλεψη του μοντέλου, ποια συγκεκριμένα χαρακτηριστικά εισόδου (π.χ., ποιος τεχνικός δείκτης ή ποια τιμή σε μια συγκεκριμένη χρονική στιγμή εντός της ακολουθίας εισόδου) συνέβαλαν περισσότερο ή λιγότερο στο τελικό αποτέλεσμα (Buy, Hold ή Sell). Η συνδυαστική χρήση πολλαπλών μεθόδων XAI επιτρέπει τη διασταύρωση των ευρημάτων και την εξαγωγή πιο στιβαρών συμπερασμάτων σχετικά με τη συμπεριφορά του μοντέλου.

5. Αξιολόγηση μέσω Προσομοίωσης Συναλλαγών (Backtesting):
   Τέλος, για την εκτίμηση της απόδοσης και της οικονομικής αξίας του εκπαιδευμένου μοντέλου LSTM, ενσωματώθηκε ένας απλός εξομοιωτής συναλλαγών (trading simulator). Ο εξομοιωτής αυτός λαμβάνει ως είσοδο τις προβλέψεις (σήματα Buy/Hold/Sell) που παράγει το μοντέλο σε μια test period, δηλαδή σε δεδομένα που δεν χρησιμοποιήθηκαν κατά την εκπαίδευση και προσομοιώνει τις συναλλαγές που θα πραγματοποιούσε έναςautomated trader ακολουθώντας αυτά τα σήματα. Ξεκινώντας με ένα υποθετικό αρχικό κεφάλαιο (π.χ., €100), ο εξομοιωτής εκτελεί τις αντίστοιχες αγοραπωλησίες και παρακολουθεί την εξέλιξη της αξίας του χαρτοφυλακίου. Αυτή η διαδικασία επιτρέπει την παραγωγή σημαντικών μετρικών απόδοσης, όπως η σωρευτική απόδοση του χαρτοφυλακίου με την πάροδο του χρόνου, ο λόγος Sharpe (που μετρά την προσαρμοσμένη στον κίνδυνο απόδοση), η μέγιστη υποχώρηση (maximum drawdown, δηλαδή η μέγιστη ποσοστιαία πτώση της αξίας του χαρτοφυλακίου από ένα προηγούμενο υψηλό σημείο), και άλλες. Οι επιδόσεις αυτές στη συνέχεια συγκρίνονται με την παθητική στρατηγική buy-and-hold της μετοχής $AAPL για την ίδια περίοδο. Αυτή η ποσοτική αξιολόγηση είναι κρίσιμη για να διαπιστωθεί κατά πόσο η χρήση του μοντέλου LSTM προσφέρει

πραγματική προστιθέμενη αξία σε σχέση με την πιο απλή επενδυτική στρατηγική.

ΚΕΦΑΛΑΙΟ 2: ΒΑΣΙΚΕΣ ΓΝΩΣΕΙΣ

Το κεφάλαιο αυτό παρέχει τις θεμελιώδεις γνώσεις που είναι απαραίτητες για την κατανόηση του ευρύτερου πλαισίου εντός του οποίου εντάσσεται η παρούσα διπλωματική εργασία. Θα εξεταστεί η ιστορία και η λειτουργία των χρηματιστηριακών αγορών, θα δοθεί έμφαση στην ιστορική πορεία και τους παράγοντες που επηρεάζουν τη μετοχή της Apple Inc. ($AAPL), θα αναλυθούν οι τρέχουσες μέθοδοι πρόβλεψης που χρησιμοποιούνται από τους αναλυτές, και τέλος, θα εισαχθεί η έννοια της Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI) και θα τονιστεί η σημασία της στις χρηματοοικονομικές προβλέψεις.

2.1 Ιστορία Χρηματιστηριακών Αγορών

Οι χρηματιστηριακές αγορές, ως οργανωμένοι χώροι για την αγοραπωλησία μετοχών εταιρειών και άλλων χρηματοοικονομικών μέσων, έχουν μια μακρά και συναρπαστική ιστορία που εκτείνεται για αιώνες. Οι ρίζες τους εντοπίζονται ήδη από τον 17ο αιώνα, με το Χρηματιστήριο του Άμστερνταμ, που ιδρύθηκε το 1602 από την Ολλανδική Εταιρεία Ανατολικών Ινδιών, να αναφέρεται συχνά ως η πρώτη επίσημη χρηματιστηριακή αγορά στον κόσμο. Η εταιρεία αυτή ήταν η πρώτη που εξέδωσε μετοχές και ομόλογα στο ευρύ κοινό, εισάγοντας θεμελιώδη στοιχεία της σύγχρονης χρηματιστηριακής πρακτικής, όπως η αρχική δημόσια προσφορά (Initial Public Offering - IPO). Ωστόσο, η ιδέα της διαπραγμάτευσης εταιρικών μεριδίων υπήρχε ακόμη και νωρίτερα, με τη Ρωμαϊκή Δημοκρατία να διαθέτει ένα σύστημα για τη διαπραγμάτευση μεριδίων σε εταιρείες δημοσίων έργων (publicani). Πιο ανεπίσημες μορφές αγορών για μετοχές και χρεόγραφα είχαν αρχίσει να εμφανίζονται σε πόλεις όπως η Αμβέρσα και το Λονδίνο ήδη από τον 16ο αιώνα. Με την πάροδο του χρόνου, χρηματιστηριακές αγορές άρχισαν να εμφανίζονται και σε άλλες ευρωπαϊκές πόλεις κατά τα τέλη του 18ου αιώνα. Το Χρηματιστήριο του Λονδίνου (London Stock Exchange - LSE) ιδρύθηκε επίσημα το 1773, εξελισσόμενο από άτυπες συναντήσεις χρηματιστών σε καφενεία σε έναν εξαιρετικά οργανωμένο και διεθνώς σημαντικό θεσμό. Στις Ηνωμένες Πολιτείες, το Χρηματιστήριο της Νέας Υόρκης (New York Stock Exchange - NYSE) ιδρύθηκε στη Wall Street της Νέας Υόρκης στις 17 Μαΐου 1792. Αρχικά, η συναλλακτική δραστηριότητα αφορούσε κυρίως κρατικά ομόλογα και τραπεζικές μετοχές, με τη Συμφωνία του Buttonwood, υπογεγραμμένη από 24 χρηματιστές, να θέτει τους κανόνες για τη διαπραγμάτευση κινητών αξιών. Το NYSE οργανώθηκε επίσημα το 1817 και έκτοτε έχει εξελιχθεί στο μεγαλύτερο χρηματιστήριο του κόσμου βάσει κεφαλαιοποίησης.

Ο 19ος αιώνας υπήρξε περίοδος σημαντικής επέκτασης της χρηματιστηριακής

διαπραγμάτευσης στις δυτικές χώρες, τροφοδοτούμενη από τη Βιομηχανική Επανάσταση και την ανάπτυξη των σιδηροδρόμων. Οι νέες τεχνολογίες επικοινωνίας, όπως ο τηλέγραφος, διευκόλυναν τις συναλλαγές σε μεγάλες αποστάσεις, συμβάλλοντας στη δημιουργία πιο ενοποιημένων εθνικών χρηματιστηριακών αγορών. Κατά τη διάρκεια αυτής της περιόδου, πολλά ιδιωτικά χρηματιστήρια συγχωνεύθηκαν σε πιο ρυθμιζόμενα δημόσια χρηματιστήρια. Το 1896, ο Charles Dow δημιούργησε τον βιομηχανικό δείκτη Dow Jones (Dow Jones Industrial Average - DJIA) για τη μέτρηση των βιομηχανικών μετοχών που διαπραγματεύονταν στο NYSE, ο οποίος παραμένει μέχρι σήμερα ένας από τους πιο γνωστούς χρηματιστηριακούς δείκτες παγκοσμίως.

Ένα κρίσιμο σημείο καμπής στην ιστορία των χρηματιστηρίων ήταν το χρηματιστηριακό κραχ του 1929 (Wall Street Crash of 1929), το οποίο πυροδότησε τη Μεγάλη Ύφεση. Το γεγονός αυτό οδήγησε στην εφαρμογή αυστηρότερων ελέγχων με στόχο τον περιορισμό της υπερβολικής κερδοσκοπίας και της απάτης. Ως αποτέλεσμα, ιδρύθηκε η Επιτροπή Κεφαλαιαγοράς των ΗΠΑ (Securities and Exchange Commission - SEC) το 1934 για τη ρύθμιση των χρηματιστηρίων. Ακολούθησαν και άλλες νομοθετικές ρυθμίσεις, όπως ο νόμος Glass-Steagall του 1933, ο οποίος ίδρυσε την Ομοσπονδιακή Αρχή Ασφάλισης Καταθέσεων (FDIC), ο νόμος περί κινητών αξιών του 1933 (Securities Act of 1933), ο οποίος απαιτούσε από τις εταιρείες να καταχωρούν νέες κινητές αξίες και να δημοσιοποιούν οικονομικά στοιχεία, και ο νόμος περί επενδυτικών εταιρειών του 1940 (Investment Company Act of 1940), ο οποίος ρύθμιζε τις επενδυτικές εταιρείες όπως τα αμοιβαία κεφάλαια.

Από τα μέσα του 20ου αιώνα και μετά, οι χρηματιστηριακές αγορές γνώρισαν αξιοσημείωτη ανάπτυξη, χαρακτηριζόμενη από την ευρεία συμμετοχή του κοινού, τις τεχνολογικές καινοτομίες όπως οι ηλεκτρονικές πλατφόρμες συναλλαγών, την παγκόσμια επέκταση και την αυξημένη ρύθμιση. Οι σύγχρονες χρηματιστηριακές αγορές είναι εξαιρετικά τεχνολογικές και παγκοσμιοποιημένες, με τις ηλεκτρονικές πλατφόρμες συναλλαγών, τις συνδέσεις διαδικτύου υψηλής ταχύτητας και τους αλγορίθμους υπολογιστών να κυριαρχούν στις συναλλαγές.

Οι βασικές λειτουργίες των χρηματιστηριακών αγορών μπορούν να συνοψιστούν ως εξής:

1. **Άντληση Κεφαλαίων για Εταιρείες:** Οι εταιρείες εκδίδουν μετοχές για να αντλήσουν κεφάλαια από επενδυτές, τα οποία χρηματοδοτούν τις λειτουργίες τους, τα σχέδια επέκτασης και την έρευνα και ανάπτυξη. Αυτός είναι ένας κρίσιμος μηχανισμός για την ανάπτυξη των επιχειρήσεων.

2. **Αποδόσεις για Επενδυτές:** Οι επενδυτές αγοράζουν μετοχές εταιρειών και μπορούν να αποκομίσουν κέρδη μέσω υπεραξιών (εάν η τιμή της μετοχής αυξηθεί) και μερισμάτων. Αυτό παρέχει μια σημαντική οδό για τη δημιουργία πλούτου και εισοδήματος μακροπρόθεσμα.
3. **Κατανομή Πόρων:** Οι χρηματιστηριακές αγορές βοηθούν στην κατεύθυνση των επενδύσεων προς καινοτόμες και αναπτυσσόμενες εταιρείες, βελτιώνοντας έτσι τη συνολική κατανομή των πόρων εντός της οικονομίας.

Επιπλέον, οι χρηματιστηριακές αγορές λειτουργούν ως δείκτης της οικονομικής υγείας, καθώς η απόδοσή τους συνδέεται στενά με τα εταιρικά κέρδη. Παρέχουν μια ρυθμιζόμενη, κεντρική αγορά για την αγορά και πώληση επενδύσεων, διασφαλίζοντας τη δικαιοσύνη και τη διαφάνεια μέσω αυστηρών κανόνων και εποπτείας. Οι τιμές των μετοχών αντανακλούν τις αντιλήψεις των επενδυτών για τη μελλοντική αξία μιας εταιρείας, αντιπροσωπεύοντας τη συλλογική γνώση, τις προσδοκίες και τις προοπτικές εκατομμυρίων επενδυτών. Η ιστορική εξέλιξη από άτυπες συναθροίσεις σε αυστηρά ρυθμιζόμενους και τεχνολογικά προηγμένους οργανισμούς υπογραμμίζει τη διαρκώς αυξανόμενη σημασία τους και την ανάγκη για διαφάνεια και έλεγχο, μια ανάγκη που η σύγχρονη τεχνολογία, συμπεριλαμβανομένης της ΤΝ και της ΧΑΙ, καλείται να υπηρετήσει. Η τεχνολογία, από τον τηλέγραφο έως τους σημερινούς αλγορίθμους υψηλής συχνότητας, έπαιζε πάντα καταλυτικό ρόλο, και η τρέχουσα εποχή της ΤΝ απλώς συνεχίζει αυτή την παράδοση, εισάγοντας νέες δυνατότητες αλλά και την επιτακτική ανάγκη για κατανόηση των πολύπλοκων αποφάσεων που λαμβάνονται.

Ο παρακάτω πίνακας συνοψίζει ορισμένους σημαντικούς σταθμούς στην ιστορία των χρηματιστηρίων:

Ημερομηνία/Περίοδος	**Γεγονός**	**Σημασία**
17ος αιώνας (1602)	Ίδρυση του Χρηματιστηρίου του Άμστερνταμ	Συχνά αναφέρεται ως το πρώτο επίσημο χρηματιστήριο, έκδοση μετοχών/ομολόγων από την Ολλανδική Εταιρεία Ανατολικών Ινδιών.

1773	Ίδρυση του Χρηματιστηρίου του Λονδίνου (LSE)	Εξέλιξη από άτυπες συναντήσεις σε οργανωμένο χρηματιστήριο.
1792	Ίδρυση του Χρηματιστηρίου της Νέας Υόρκης (NYSE)	Αρχικά διαπραγμάτευση κυβερνητικών ομολόγων, μετέπειτα εξέλιξη στο μεγαλύτερο χρηματιστήριο του κόσμου.
19ος αιώνας	Βιομηχανική Επανάσταση, ανάπτυξη σιδηροδρόμων, τηλέγραφος	Σημαντική επέκταση της χρηματιστηριακής διαπραγμάτευσης, διευκόλυνση συναλλαγών σε μεγάλες αποστάσεις.
1929	Χρηματιστηριακό Κραχ της Wall Street	Πυροδότησε τη Μεγάλη Ύφεση, οδήγησε σε αυξημένη ρύθμιση.
1934	Ίδρυση της Επιτροπής Κεφαλαιαγοράς των ΗΠΑ (SEC)	Εισαγωγή αυστηρότερων ελέγχων και ρύθμισης στις χρηματιστηριακές αγορές για την προστασία των επενδυτών και την αποφυγή απάτης.

Πίνακας 2.1: Σημαντικοί Σταθμοί στην Ιστορία των Χρηματιστηρίων

2.2 Ιστορία Μετοχής Apple ($AAPL)

Η Apple Inc. ($AAPL) αποτελεί μία από τις μεγαλύτερες και πιο αναγνωρίσιμες εταιρείες τεχνολογίας παγκοσμίως, με τεράστια επιρροή τόσο στην καταναλωτική αγορά όσο και στις χρηματοοικονομικές αγορές. Η ιστορία της μετοχής της είναι άρρηκτα συνδεδεμένη με την καινοτομία, τις στρατηγικές της αποφάσεις και την ευρύτερη δυναμική του τεχνολογικού κλάδου.

Η Apple εισήχθη στο χρηματιστήριο (IPO) στις 12 Δεκεμβρίου 1980, με τιμή $22.00 ανά μετοχή. Από τότε, η μετοχή έχει γνωρίσει σημαντικές περιόδους ανάπτυξης αλλά και διακυμάνσεων, αντανακλώντας την πορεία της εταιρείας. Για την καλύτερη κατανόηση της ιστορικής της πορείας, είναι σημαντικό να ληφθούν υπόψη οι διασπάσεις μετοχών (stock splits) που έχουν πραγματοποιηθεί. Η Apple έχει προχωρήσει σε πέντε διασπάσεις μετοχών από την αρχική της δημόσια προσφορά:

1. **16 Ιουνίου 1987:** Διάσπαση 2-προς-1 (2-for-1).
2. **21 Ιουνίου 2000:** Διάσπαση 2-προς-1 (2-for-1).
3. **28 Φεβρουαρίου 2005:** Διάσπαση 2-προς-1 (2-for-1).
4. **9 Ιουνίου 2014:** Διάσπαση 7-προς-1 (7-for-1).
5. **28 Αυγούστου 2020** (ημερομηνία καταγραφής, η διαπραγμάτευση με τη νέα τιμή ξεκίνησε στις 31 Αυγούστου 2020): Διάσπαση 4-προς-1 (4-for-1). Οι διασπάσεις μετοχών πραγματοποιούνται συχνά για να καταστήσουν την τιμή της μετοχής πιο προσιτή σε ένα ευρύτερο φάσμα επενδυτών, ειδικά όταν η τιμή έχει αυξηθεί σημαντικά. Η προσαρμοσμένη για τις διασπάσεις τιμή IPO της Apple ήταν μόλις $0.10.

Η τιμή της μετοχής της Apple έχει επηρεαστεί σημαντικά από την κυκλοφορία και την επιτυχία των προϊόντων της. Η εισαγωγή του iPhone το 2007, για παράδειγμα, σηματοδότησε την έναρξη μιας περιόδου εκρηκτικής ανάπτυξης για την εταιρεία και τη μετοχή της. Η ικανότητα της Apple να καινοτομεί και να δημιουργεί μια τεράστια καταναλωτική βάση οδήγησε σε σημαντικές αυξήσεις της τιμής της μετοχής με την πάροδο του χρόνου. Πέρα από τις πωλήσεις υλικού (hardware), η ανάπτυξη του τομέα υπηρεσιών της Apple (όπως Apple Music, Apple TV+, iCloud, App Store) έχει αναδειχθεί σε σημαντικό παράγοντα σταθερότητας και ανάπτυξης για τη μετοχή. Αυτή η διαφοροποίηση βοηθά την Apple να μειώσει την εξάρτησή της από τις κυκλικές πωλήσεις υλικού και να δημιουργήσει πιο προβλέψιμες, επαναλαμβανόμενες ροές εσόδων, κάτι που εκτιμάται από τους επενδυτές.

Οι παράγοντες που επηρεάζουν την τιμή της μετοχής $AAPL είναι πολυάριθμοι και

περιλαμβάνουν:

- **Κυκλοφορίες Προϊόντων:** Οι ανακοινώσεις νέων προϊόντων (iPhones, Macs, Apple Watch κ.λπ.) και η αποδοχή τους από την αγορά έχουν άμεσο αντίκτυπο. Ισχυρές πωλήσεις συνήθως οδηγούν σε άνοδο της μετοχής.
- **Τριμηνιαίες Αναφορές Κερδών:** Οι οικονομικές επιδόσεις της εταιρείας, όπως η αύξηση των εσόδων, η κερδοφορία και άλλοι βασικοί χρηματοοικονομικοί δείκτες, παρακολουθούνται στενά από τους επενδυτές.
- **Γενικές Τάσεις Αγοράς και Οικονομικές Συνθήκες:** Η ευρύτερη πορεία του χρηματιστηρίου, οι οικονομικές υφέσεις, οι αλλαγές στα επιτόκια και άλλοι μακροοικονομικοί παράγοντες επηρεάζουν το επενδυτικό κλίμα και, κατά συνέπεια, την τιμή της μετοχής.
- **Καινοτομία και Στρατηγική:** Η συνεχής εστίαση της Apple στην καινοτομία και η μακροπρόθεσμη στρατηγική της αποτελούν βασικούς μοχλούς της αξίας της.
- **Τάσεις της Βιομηχανίας:** Η υιοθέτηση νέων τεχνολογιών, όπως η τεχνητή νοημοσύνη (AI), από την Apple και η θέση της έναντι των ανταγωνιστών σε αυτές τις τάσεις.
- **Εφοδιαστική Αλυσίδα:** Ως εταιρεία που βασίζεται σε μεγάλο βαθμό στο υλικό, η Apple είναι ευάλωτη σε διαταραχές στην παγκόσμια εφοδιαστική αλυσίδα, ελλείψεις ημιαγωγών και αύξηση του κόστους εργασίας σε βασικές περιοχές παραγωγής.
- **Επέκταση σε Νέες Αγορές:** Η είσοδος της Apple σε νέες γεωγραφικές αγορές ή νέους τομείς προϊόντων/υπηρεσιών.
- **Ανταγωνισμός:** Η Apple λειτουργεί σε ένα εξαιρετικά ανταγωνιστικό περιβάλλον, αντιμετωπίζοντας ισχυρούς αντιπάλους όπως η Samsung, η Google και η Microsoft.
- **Ρυθμιστικό Πλαίσιο:** Έλεγχοι σχετικά με την προστασία δεδομένων, τις πολιτικές του App Store και ζητήματα αντιμονοπωλιακής νομοθεσίας μπορούν να επηρεάσουν τις δραστηριότητες και την κερδοφορία της εταιρείας.

Η κατανόηση αυτών των παραγόντων και της ιστορικής τους επίδρασης στη μετοχή $AAPL παρέχει ένα κρίσιμο πλαίσιο για την ανάπτυξη και αξιολόγηση μοντέλων πρόβλεψης. Ένα μοντέλο μηχανικής μάθησης, αν και βασίζεται κυρίως σε ποσοτικά δεδομένα, προσπαθεί εμμέσως να συλλάβει τα πρότυπα που δημιουργούνται από την αλληλεπίδραση αυτών των πολύπλοκων δυνάμεων.

Ο παρακάτω πίνακας συνοψίζει το ιστορικό των διασπάσεων της μετοχής της Apple:

Ημερομηνία Split	**Αναλογία Split**	**Τιμή Πριν το Split (κατά προσέγγιση)**	**Σύντομη Σημείωση/Λόγος**
16 Ιουνίου 1987	2-για-1	$79 (31 Μαΐου 1987)	Διατήρηση της τιμής της μετοχής σε προσιτά επίπεδα για τους επενδυτές.
21 Ιουνίου 2000	2-για-1	$111 (31 Μαΐου 2000)	Εν μέσω της φούσκας των dot-com, αν και ακολούθησε πτώση λόγω ευρύτερων συνθηκών αγοράς.
28 Φεβρουαρίου 2005	2-για-1	$90 (31 Ιανουαρίου 2005)	Μετά από σημαντική αύξηση της αξίας της μετοχής, εν μέρει λόγω της επιτυχίας του iPod και του iTunes.
9 Ιουνίου 2014	7-για-1	$656 (31 Μαΐου 2014)	Η πιο σημαντική διάσπαση, μείωσε δραστικά την ονομαστική τιμή για μεγαλύτερη προσιτότητα και πιθανώς για διευκόλυνση της εισαγωγής στον δείκτη Dow Jones.
31 Αυγούστου	4-για-1	-	Περαιτέρω μείωση της

2020			ονομαστικής τιμής μετά από συνεχή άνοδο.

Πίνακας 2.2: Ιστορικό Stock Splits της Apple (AAPL)

2.3 Πώς πραγματοποιούνται τρέχουσες προβλέψεις της χρηματιστηριακής αγοράς από αναλυτές

Η πρόβλεψη των κινήσεων της χρηματιστηριακής αγοράς είναι μια διαρκής προσπάθεια που απασχολεί επενδυτές, αναλυτές και ακαδημαϊκούς. Παραδοσιακά, οι αναλυτές βασίζονται σε δύο κύριες σχολές σκέψης για την προσέγγιση αυτής της πρόκλησης: την Τεχνική Ανάλυση και τη Θεμελιώδη Ανάλυση. Πολλοί επαγγελματίες, ωστόσο, χρησιμοποιούν έναν συνδυασμό και των δύο προσεγγίσεων για να διαμορφώσουν τις επενδυτικές τους στρατηγικές.

1. Τεχνική Ανάλυση (Technical Analysis)

Η τεχνική ανάλυση είναι μια μεθοδολογία αξιολόγησης κινητών αξιών μέσω της ανάλυσης στατιστικών που προκύπτουν από τη δραστηριότητα της αγοράς, όπως οι ιστορικές τιμές και ο όγκος συναλλαγών. Οι τεχνικοί αναλυτές δεν προσπαθούν να μετρήσουν την εγγενή αξία μιας μετοχής, αλλά χρησιμοποιούν διαγράμματα και άλλα εργαλεία για να εντοπίσουν μοτίβα και τάσεις που μπορεί να υποδηλώνουν μελλοντικές κινήσεις της τιμής.

- **Υποκείμενες Παραδοχές :**
  - **Η αγορά προεξοφλεί τα πάντα:** Η τρέχουσα τιμή μιας μετοχής αντανακλά όλες τις διαθέσιμες πληροφορίες, συμπεριλαμβανομένων θεμελιωδών παραγόντων, οικονομικών συνθηκών και ψυχολογίας της αγοράς.
  - **Οι τιμές κινούνται σε τάσεις:** Οι τιμές των μετοχών τείνουν να ακολουθούν αναγνωρίσιμες τάσεις (ανοδικές, πτωτικές, πλάγιες) για κάποιο χρονικό διάστημα.
  - **Η ιστορία επαναλαμβάνεται:** Ορισμένα μοτίβα τιμών (chart patterns) τείνουν να επαναλαμβάνονται με την πάροδο του χρόνου, κυρίως λόγω της προβλέψιμης ανθρώπινης ψυχολογίας (φόβος, απληστία) που οδηγεί τις αποφάσεις της αγοράς.
- **Εργαλεία και Δείκτες:**
  - **Τάσεις Τιμών (Price Trends):** Εντοπισμός της γενικής κατεύθυνσης της αγοράς ή μιας μετοχής (π.χ., γραμμές τάσης).

- **Μοτίβα Διαγραμμάτων (Chart Patterns):** Αναγνώριση σχηματισμών στα διαγράμματα τιμών (π.χ., κεφάλι και ώμοι, διπλοί πυθμένες/κορυφές, τρίγωνα) που μπορεί να σηματοδοτούν συνέχιση ή αντιστροφή της τάσης.
- **Δείκτες Όγκου και Ορμής (Volume and Momentum Indicators):** Ο όγκος συναλλαγών χρησιμοποιείται για την επιβεβαίωση της ισχύος μιας τάσης. Δείκτες ορμής (π.χ., Momentum, Rate of Change - ROC) μετρούν την ταχύτητα των μεταβολών των τιμών.
- **Ταλαντωτές (Oscillators):** Δείκτες που κυμαίνονται μεταξύ δύο ακραίων τιμών και χρησιμοποιούνται για τον εντοπισμό συνθηκών υπεραγοράς (overbought) ή υπερπώλησης (oversold) (π.χ., Relative Strength Index - RSI, Stochastic Oscillator).
- **Κινητοί Μέσοι Όροι (Moving Averages - MAs):** Εξομαλύνουν τις διακυμάνσεις των τιμών για να αναδείξουν την υποκείμενη τάση. Συνηθισμένοι τύποι είναι ο Απλός Κινητός Μέσος Όρος (SMA) και ο Εκθετικός Κινητός Μέσος Όρος (EMA). Οι διασταυρώσεις κινητών μέσων διαφορετικών περιόδων (π.χ., "χρυσός σταυρός", "σταυρός θανάτου") θεωρούνται σημαντικά σήματα.
- **Επίπεδα Στήριξης και Αντίστασης (Support and Resistance Levels):** Επίπεδα τιμών στα οποία η τιμή μιας μετοχής τείνει να σταματά την πτώση της (στήριξη) ή την άνοδό της (αντίσταση).

Η διπλωματική εργασία, χρησιμοποιώντας ένα μοντέλο LSTM που τροφοδοτείται με ιστορικές τιμές και τεχνικούς δείκτες, εντάσσεται σε μεγάλο βαθμό στο πλαίσιο της τεχνικής ανάλυσης. Η προσθήκη, ωστόσο, μακροοικονομικών δεικτών όπως ο S&P 500 και ο VIX, προσδίδει στοιχεία ευρύτερης ανάλυσης της αγοράς.

2. Θεμελιώδης Ανάλυση (Fundamental Analysis)

Η θεμελιώδης ανάλυση επικεντρώνεται στην αξιολόγηση της "πραγματικής" ή "εγγενούς" αξίας (intrinsic value) μιας εταιρείας και των μετοχών της. Οι θεμελιώδεις αναλυτές εξετάζουν ένα ευρύ φάσμα ποσοτικών και ποιοτικών παραγόντων για να καθορίσουν εάν μια μετοχή είναι υποτιμημένη ή υπερτιμημένη από την αγορά.

- **Ποσοτικοί Παράγοντες:**
  - **Οικονομικές Καταστάσεις:** Ανάλυση του ισολογισμού, της κατάστασης αποτελεσμάτων χρήσης και της κατάστασης ταμειακών ροών για την κατανόηση της οικονομικής υγείας, της κερδοφορίας και της ρευστότητας της εταιρείας.
  - **Δείκτες Αποτίμησης:**

        - Λόγος Τιμής προς Κέρδη (Price-to-Earnings Ratio - P/E): Συγκρίνει την τρέχουσα τιμή της μετοχής με τα κέρδη ανά μετοχή.
        - Λόγος Τιμής προς Λογιστική Αξία (Price-to-Book Ratio - P/B): Συγκρίνει την αγοραία αξία της εταιρείας με τη λογιστική της αξία.
    - **Δείκτες Κερδοφορίας:**
        - Απόδοση Ιδίων Κεφαλαίων (Return on Equity - ROE): Μετρά την κερδοφορία σε σχέση με τα ίδια κεφάλαια των μετόχων.
        - Απόδοση Ενεργητικού (Return on Assets - ROA): Μετρά την κερδοφορία σε σχέση με το σύνολο του ενεργητικού της εταιρείας.
    - **Επίπεδα Χρέους:** Δείκτης Χρέους προς Ίδια Κεφάλαια (Debt-to-Equity Ratio - D/E) για την αξιολόγηση της μόχλευσης της εταιρείας.
    - **Μοντέλα Αποτίμησης:**
        - Μοντέλο Προεξόφλησης Ταμειακών Ροών (Discounted Cash Flow - DCF): Εκτιμά την αξία μιας εταιρείας με βάση τις προβλεπόμενες μελλοντικές ελεύθερες ταμειακές ροές της, προεξοφλημένες στο παρόν.
        - Μοντέλο Προεξόφλησης Μερισμάτων (Dividend Discount Model - DDM): Εκτιμά την αξία μιας μετοχής με βάση την παρούσα αξία των αναμενόμενων μελλοντικών μερισμάτων.
- **Ποιοτικοί Παράγοντες:**
    - **Ποιότητα Διοίκησης (Management Quality):** Η ικανότητα και η εμπειρία της διοικητικής ομάδας.
    - **Ανταγωνιστικό Πλεονέκτημα (Economic Moat):** Μοναδικά πλεονεκτήματα που προστατεύουν την εταιρεία από τον ανταγωνισμό (π.χ., ισχυρό brand, πατέντες, υψηλό κόστος αλλαγής για τους πελάτες).
    - **Φήμη Μάρκας (Brand Reputation):** Η αναγνωρισιμότητα και η αξιοπιστία της μάρκας.
    - **Συνθήκες Κλάδου και Αγοράς:** Οι προοπτικές ανάπτυξης του κλάδου, το ρυθμιστικό περιβάλλον και οι μακροοικονομικές συνθήκες.

Άλλες Προσεγγίσεις

Πέρα από τις δύο κύριες σχολές, υπάρχουν και άλλες θεωρίες ή στρατηγικές που χρησιμοποιούνται, όπως:

- **Mean Reversion (Επιστροφή στον Μέσο Όρο):** Η υπόθεση ότι οι τιμές των μετοχών και οι ιστορικές αποδόσεις τείνουν να επιστρέφουν μακροπρόθεσμα στον μακροχρόνιο μέσο όρο ή μέση τιμή τους.
- **Momentum Investing (Επένδυση Ορμής):** Η στρατηγική αγοράς μετοχών που έχουν παρουσιάσει ισχυρή ανοδική τάση πρόσφατα, με την προσδοκία ότι αυτή η τάση θα συνεχιστεί.

- **Martingale Theory (Θεωρία Martingale):** Μια μαθηματική έννοια από τη θεωρία πιθανοτήτων, όπου η καλύτερη πρόβλεψη για την επόμενη τιμή μιας σειράς είναι η τρέχουσα τιμή. Στις αγορές, αυτό συχνά μεταφράζεται στην ιδέα του "τυχαίου περίπατου" (random walk).

Η Εξηγήσιμη Τεχνητή Νοημοσύνη (XAI) μπορεί να διαδραματίσει σημαντικό ρόλο στη γεφύρωση του χάσματος μεταξύ των ποσοτικών μοντέλων, όπως το LSTM που χρησιμοποιείται εδώ, και της ανθρώπινης διαίσθησης των αναλυτών. Εξηγώντας ποιοι τεχνικοί δείκτες ή ποια χρονικά μοτίβα οδηγούν σε μια συγκεκριμένη πρόβλεψη, το μοντέλο καθίσταται λιγότερο "μαύρο κουτί". Οι εξηγήσεις του μπορούν να συγκριθούν με τη λογική που θα ακολουθούσε ένας έμπειρος τεχνικός αναλυτής. Για παράδειγμα, εάν ένας αναλυτής τείνει να αγοράζει όταν ο δείκτης RSI υποδεικνύει υπερπουλημένες συνθήκες και οι κινητοί μέσοι όροι δείχνουν ανοδική δυναμική, και οι μέθοδοι XAI αποκαλύψουν ότι το μοντέλο LSTM βασίζει μια απόφαση "Buy" σε παρόμοιους παράγοντες, αυτό ενισχύει την εμπιστοσύνη και την κατανόηση της λειτουργίας του μοντέλου. Με την εξήγηση των αποφάσεων μπορούμε να ελέγξουμε εάν το μοντέλο βασίζεται σε οικονομικά εύλογα σήματα.

Ο παρακάτω πίνακας συνοψίζει τις βασικές διαφορές μεταξύ Τεχνικής και Θεμελιώδους Ανάλυσης:

Κριτήριο	**Τεχνική Ανάλυση**	**Θεμελιώδης Ανάλυση**
Βασική Εστίαση	Ιστορικές τιμές και όγκος συναλλαγών, μοτίβα διαγραμμάτων.	Εγγενής αξία εταιρείας, οικονομικές καταστάσεις, μακροοικονομικοί παράγοντες, ποιότητα διοίκησης.
Χρονικός Ορίζοντας	Κυρίως βραχυπρόθεσμος έως μεσοπρόθεσμος.	Κυρίως μεσοπρόθεσμος έως μακροπρόθεσμος.
Εργαλεία/Δείκτες	Διαγράμματα τιμών, κινητοί μέσοι όροι, RSI, MACD, Bollinger Bands, επίπεδα	Οικονομικές καταστάσεις, P/E, P/B, ROE, D/E, DCF, DDM, ανάλυση

	στήριξης/αντίστασης, όγκος συναλλαγών.	κλάδου, ποιότητα διοίκησης.
Βασικές Υποθέσεις	Η αγορά προεξοφλεί τα πάντα, οι τιμές κινούνται σε τάσεις, η ιστορία επαναλαμβάνεται.	Η αγορά μπορεί να υποτιμά ή να υπερτιμά μετοχές, η εγγενής αξία θα αναγνωριστεί μακροπρόθεσμα.
Στόχος	Πρόβλεψη μελλοντικών κινήσεων τιμών και εντοπισμός σημείων εισόδου/εξόδου.	Εντοπισμός υποτιμημένων ή υπερτιμημένων μετοχών για μακροπρόθεσμη επένδυση.

Πίνακας 2.3: Σύγκριση Τεχνικής και Θεμελιώδους Ανάλυσης

2.4 Εξηγησιμότητα και η Σημασία της στις Προβλέψεις της Χρηματιστηριακής Αγοράς

Καθώς η τεχνητή νοημοσύνη (ΤΝ) και οι αλγόριθμοι μηχανικής μάθησης (ΜΜ) ενσωματώνονται όλο και περισσότερο σε κρίσιμους τομείς λήψης αποφάσεων, όπως οι χρηματοοικονομικές αγορές, η ανάγκη για κατανόηση και διαφάνεια αυτών των συστημάτων γίνεται επιτακτική. Εδώ παρεμβαίνει η Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable Artificial Intelligence - XAI).

Η XAI ορίζεται ως ένα σύνολο διαδικασιών και μεθόδων που επιτρέπουν στους ανθρώπινους χρήστες να κατανοήσουν και να εμπιστευτούν τα αποτελέσματα και τις εξόδους που παράγονται από αλγορίθμους μηχανικής μάθησης. Ο κεντρικός στόχος της XAI είναι να καταστήσει τη συμπεριφορά αυτών των, συχνά πολύπλοκων, συστημάτων κατανοητή στους ανθρώπους, διευκρινίζοντας τους υποκείμενους μηχανισμούς των διαδικασιών λήψης αποφάσεών τους. Ιδανικά, η XAI θα πρέπει να περιλαμβάνει την ικανότητα να εξηγεί τις αρμοδιότητες και την κατανόηση του συστήματος, τις προηγούμενες ενέργειές του, τις τρέχουσες διαδικασίες και τα επόμενα βήματα, καθώς και να αποκαλύπτει τις σχετικές πληροφορίες στις οποίες βασίζονται οι ενέργειές του.

Το θεμελιώδες πρόβλημα που η XAI προσπαθεί να αντιμετωπίσει είναι το φαινόμενο του "μαύρου κουτιού" (black box effect). Πολλά σύγχρονα μοντέλα ΜΜ, ειδικά τα

βαθιά νευρωνικά δίκτυα όπως τα LSTMs, λειτουργούν με τρόπους που είναι αδιαφανείς ακόμη και για τους δημιουργούς τους. Οι αποφάσεις τους προκύπτουν από πολύπλοκες, μη γραμμικές αλληλεπιδράσεις εκατομμυρίων παραμέτρων, καθιστώντας εξαιρετικά δύσκολη την παρακολούθηση της λογικής που οδηγεί σε ένα συγκεκριμένο αποτέλεσμα.

Οι κίνδυνοι από τέτοια αδιαφανή συστήματα ΤΝ είναι σημαντικοί:

- **Απόκρυψη Ελλείψεων:** Η αδιαφάνεια μπορεί να κρύψει εγγενείς αδυναμίες στα συστήματα ΤΝ, όπως συστηματικές μεροληψίες (biases) που προέρχονται από τα δεδομένα εκπαίδευσης, ανακρίβειες στις προβλέψεις ή ακόμα και τις λεγόμενες "παραισθήσεις" (hallucinations), όπου το μοντέλο παράγει εξόδους που φαίνονται εύλογες αλλά είναι εσφαλμένες ή άσχετες.
- **Παραπλανητική Εμπιστοσύνη ή Υπερβολική Εξάρτηση:** Οι χρήστες μπορεί να εμπιστευτούν υπερβολικά τις αποφάσεις ενός "μαύρου κουτιού" χωρίς να κατανοούν τους περιορισμούς του ή, αντίστροφα, να το απορρίψουν εντελώς λόγω έλλειψης κατανόησης.
- **Δυσκολία Λογοδοσίας:** Όταν οι αποφάσεις λαμβάνονται από αλγορίθμους των οποίων η λογική δεν είναι σαφής, γίνεται δύσκολο να αποδοθούν ευθύνες, ειδικά αν οι αποφάσεις αυτές έχουν αρνητικές συνέπειες.
- **Πιθανές Διακρίσεις ή Επιβλαβή Αποτελέσματα:** Κακοσχεδιασμένοι ή πλημμελώς εκπαιδευμένοι αλγόριθμοι μπορούν να οδηγήσουν σε αποτελέσματα που εισάγουν διακρίσεις εις βάρος συγκεκριμένων ομάδων ατόμων.

Στο πλαίσιο των χρηματοοικονομικών προβλέψεων και της διαπραγμάτευσης μετοχών, η σημασία της XAI είναι πολλαπλή:

1. **Οικοδόμηση Εμπιστοσύνης:** Όπως αναφέρθηκε και στην ενότητα 1.1, οι επενδυτές και οι διαχειριστές κεφαλαίων διστάζουν να βασιστούν σε συστήματα που δεν κατανοούν, ειδικά όταν διακυβεύονται σημαντικά οικονομικά ποσά. Η XAI, παρέχοντας εξηγήσεις για το *γιατί* ένα μοντέλο προτείνει μια συγκεκριμένη ενέργεια (π.χ., "αγορά της $AAPL διότι ο δείκτης RSI είναι χαμηλός και ο όγκος συναλλαγών αυξάνεται"), μπορεί να αυξήσει σημαντικά την εμπιστοσύνη στο σύστημα. Αυτή η εμπιστοσύνη είναι απαραίτητη για την υιοθέτηση της ΤΝ σε εφαρμογές υψηλού ρίσκου.
2. **Υποβοήθηση Στρατηγικής Λήψης Αποφάσεων:** Οι εξηγήσεις που παρέχονται από τις μεθόδους XAI μπορούν να προσφέρουν πολύτιμες γνώσεις στους επαγγελματίες των χρηματοοικονομικών αγορών. Κατανοώντας τους παράγοντες που το μοντέλο θεωρεί σημαντικούς για μια πρόβλεψη, οι αναλυτές μπορούν να αξιολογήσουν την ορθότητα αυτών των παραγόντων, να τους συγκρίνουν με τη δική τους εμπειρία και γνώση, και τελικά να λάβουν πιο

τεκμηριωμένες επενδυτικές αποφάσεις. Η XAI μπορεί να αποκαλύψει εάν το μοντέλο έχει μάθει εύλογες οικονομικές σχέσεις ή εάν βασίζεται σε τυχαίες συσχετίσεις.[1]

3. **Διαχείριση Κινδύνου:** Η κατανόηση του τρόπου με τον οποίο ένα μοντέλο αντιδρά σε διάφορες συνθήκες της αγοράς μπορεί να βοηθήσει στον εντοπισμό και τον μετριασμό των κινδύνων. Εάν, για παράδειγμα, οι εξηγήσεις αποκαλύψουν ότι το μοντέλο είναι υπερβολικά ευαίσθητο σε έναν συγκεκριμένο δείκτη που είναι γνωστό ότι είναι ασταθής, μπορούν να ληφθούν μέτρα για τον περιορισμό της εξάρτησης από αυτόν τον δείκτη ή για την προσαρμογή της στρατηγικής.
4. **Κανονιστική Συμμόρφωση:** Ο χρηματοοικονομικός τομέας υπόκειται σε αυστηρό ρυθμιστικό πλαίσιο. Η διαφάνεια και η ικανότητα εξήγησης των αλγοριθμικών αποφάσεων καθίστανται όλο και περισσότερο νομικές απαιτήσεις, ειδικά σε κρίσιμους τομείς όπως η παροχή επενδυτικών συμβουλών, η δανειοδότηση και οι επενδύσεις. Η XAI παρέχει τα εργαλεία για την τεκμηρίωση και την αιτιολόγηση των αποφάσεων που λαμβάνονται από συστήματα ΤΝ, διευκολύνοντας τη συμμόρφωση.
5. **Βελτίωση και Εντοπισμός Σφαλμάτων Μοντέλων:** Η XAI μπορεί να βοηθήσει στον εντοπισμό αδυναμιών ή ανεπιθύμητων συμπεριφορών στα μοντέλα ΜΜ. Για παράδειγμα, εάν οι εξηγήσεις δείχνουν ότι το μοντέλο βασίζεται σε ένα χαρακτηριστικό που δεν θα έπρεπε να είναι σημαντικό (π.χ., λόγω data leakage), αυτό μπορεί να οδηγήσει σε διόρθωση του μοντέλου ή των δεδομένων εκπαίδευσης.

Η XAI συνδέεται στενά με τις αρχές της Υπεύθυνης Τεχνητής Νοημοσύνης (Responsible AI), οι οποίες περιλαμβάνουν τη διαφάνεια, τη δικαιοσύνη, την ασφάλεια, την προστασία της ιδιωτικής ζωής και τη λογοδοσία. Η διαφάνεια, ειδικότερα, είναι θεμελιώδης για την επίτευξη λογοδοσίας.

Η έλλειψη διαφάνειας στα χρηματοοικονομικά μοντέλα ΤΝ δεν είναι απλώς ένα ακαδημαϊκό ζήτημα, αλλά μια πραγματική ανησυχία με άμεσες οικονομικές, νομικές και ηθικές προεκτάσεις. Η XAI δεν αποσκοπεί στο να καταστήσει κάθε πτυχή ενός πολύπλοκου μοντέλου πλήρως κατανοητή, αλλά στο να παρέχει επαρκείς εξηγήσεις ώστε οι άνθρωποι να μπορούν να λαμβάνουν τεκμηριωμένες αποφάσεις σχετικά με τη χρήση και την εμπιστοσύνη σε αυτά τα συστήματα. Επιπλέον, η XAI μπορεί να διευκολύνει μια πιο αποτελεσματική συνεργασία ανθρώπου-μηχανής. Αντί οι αλγόριθμοι να λειτουργούν ως απόλυτοι αντικαταστάτες των ανθρώπινων αναλυτών, μπορούν να εξελιχθούν σε ισχυρά εργαλεία υποστήριξης αποφάσεων. Σε αυτό το πλαίσιο, η XAI παρέχει την απαραίτητη "γλώσσα" επικοινωνίας, επιτρέποντας στους

ανθρώπους να κατανοούν τις προτάσεις της μηχανής, να τις αξιολογούν κριτικά και να τις ενσωματώνουν στη δική τους διαδικασία λήψης αποφάσεων. Αυτή η συνέργεια είναι ιδιαίτερα πολύτιμη σε έναν τομέα τόσο δυναμικό και αβέβαιο όσο οι χρηματοοικονομικές αγορές.

ΚΕΦΑΛΑΙΟ 3: ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Το κεφάλαιο αυτό εισάγει τις βασικές έννοιες και τεχνικές της μηχανικής μάθησης που είναι θεμελιώδεις για την κατανόηση της παρούσας διπλωματικής εργασίας. Αρχικά, γίνεται μια γενική επισκόπηση των μοντέλων μηχανικής μάθησης. Στη συνέχεια, η εστίαση μετατοπίζεται στα μοντέλα που είναι ειδικά σχεδιασμένα για την ανάλυση και πρόβλεψη δεδομένων χρονοσειρών. Ιδιαίτερη έμφαση δίνεται στα μοντέλα Μακράς Βραχυπρόθεσμης Μνήμης (LSTM), με λεπτομερή ανάλυση της αρχιτεκτονικής, της μαθηματικής τους βάσης και του τρόπου λειτουργίας τους. Τέλος, εξετάζεται η εφαρμογή των LSTMs στις προβλέψεις της χρηματιστηριακής αγοράς και αναλύεται η λογική πίσω από τα πολυ-εισοδικά (multi-input) LSTMs, μια προσέγγιση που υιοθετείται στην παρούσα εργασία.

3.1 Μοντέλα Μηχανικής Μάθησης

Η μηχανική μάθηση (Machine Learning - ML) είναι ένας υποκλάδος της τεχνητής νοημοσύνης (Artificial Intelligence - AI) που επικεντρώνεται στην ανάπτυξη αλγορίθμων οι οποίοι επιτρέπουν στα συστήματα υπολογιστών να "μαθαίνουν" από δεδομένα, χωρίς να είναι ρητά προγραμματισμένα για κάθε συγκεκριμένη εργασία. Αντί να ακολουθούν ένα στατικό σύνολο οδηγιών, τα μοντέλα μηχανικής μάθησης αναγνωρίζουν πρότυπα (patterns) μέσα στα δεδομένα και προσαρμόζουν τη συμπεριφορά τους ανάλογα, βελτιώνοντας την απόδοσή τους με την πάροδο του χρόνου καθώς εκτίθενται σε περισσότερα δεδομένα.

Τα μοντέλα μηχανικής μάθησης μπορούν γενικά να κατηγοριοποιηθούν σε τρεις κύριους τύπους, ανάλογα με τη φύση των δεδομένων εκπαίδευσης και τον τρόπο με τον οποίο μαθαίνουν:

1. **Επιβλεπόμενη Μάθηση (Supervised Learning):** Σε αυτή την κατηγορία, το μοντέλο εκπαιδεύεται χρησιμοποιώντας ένα σύνολο δεδομένων όπου κάθε δείγμα εισόδου (input sample) συνοδεύεται από μια σωστή ετικέτα εξόδου (output label). Ο στόχος του μοντέλου είναι να μάθει μια συνάρτηση απεικόνισης (mapping function) που μπορεί να προβλέψει την ετικέτα εξόδου για νέα, μη παρατηρηθέντα δεδομένα εισόδου. Η επιβλεπόμενη μάθηση περιλαμβάνει δύο κύριους τύπους προβλημάτων:
    - **Ταξινόμηση (Classification):** Η ετικέτα εξόδου είναι μια κατηγορική μεταβλητή (π.χ., πρόβλεψη εάν μια μετοχή θα ανέβει, θα πέσει ή θα παραμείνει σταθερή – "Buy", "Sell", "Hold", όπως στην παρούσα διπλωματική).
    - **Παλινδρόμηση (Regression):** Η ετικέτα εξόδου είναι μια συνεχής

μεταβλητή (π.χ., πρόβλεψη της ακριβούς τιμής μιας μετοχής).

2. **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning):** Εδώ, το μοντέλο εκπαιδεύεται χρησιμοποιώντας δεδομένα που δεν έχουν προκαθορισμένες ετικέτες εξόδου. Ο στόχος είναι το μοντέλο να ανακαλύψει από μόνο του κρυμμένες δομές, πρότυπα ή σχέσεις μέσα στα δεδομένα. Οι συνήθεις εργασίες περιλαμβάνουν:
    - **Ομαδοποίηση (Clustering):** Ομαδοποίηση παρόμοιων δειγμάτων δεδομένων μαζί (π.χ., τμηματοποίηση πελατών με βάση τη συμπεριφορά τους).
    - **Μείωση Διαστατικότητας (Dimensionality Reduction):** Μείωση του αριθμού των μεταβλητών εισόδου διατηρώντας παράλληλα τις σημαντικές πληροφορίες (π.χ., Ανάλυση Κύριων Συνιστωσών - PCA).
3. **Ενισχυτική Μάθηση (Reinforcement Learning):** Σε αυτό το παράδειγμα μάθησης, ένας "πράκτορας" (agent) μαθαίνει να λαμβάνει αποφάσεις αλληλεπιδρώντας με ένα "περιβάλλον" (environment). Ο πράκτορας λαμβάνει ανταμοιβές (rewards) ή ποινές (penalties) για τις ενέργειές του και ο στόχος του είναι να μάθει μια στρατηγική (policy) που μεγιστοποιεί τη συνολική του ανταμοιβή με την πάροδο του χρόνου. Χρησιμοποιείται συχνά σε προβλήματα όπως η ρομποτική, τα παιχνίδια και η βελτιστοποίηση συστημάτων.

Υπάρχει μια πληθώρα συγκεκριμένων αλγορίθμων μηχανικής μάθησης, καθένας με τα δικά του χαρακτηριστικά, πλεονεκτήματα και περιορισμούς. Μερικά κοινά παραδείγματα περιλαμβάνουν:

- Γραμμική Παλινδρόμηση (Linear Regression)
- Λογιστική Παλινδρόμηση (Logistic Regression)
- Δέντρα Απόφασης (Decision Trees)
- Τυχαία Δάση (Random Forests)
- Μηχανές Υποστήριξης Διανυσμάτων (Support Vector Machines - SVMs)
- Νευρωνικά Δίκτυα (Neural Networks), συμπεριλαμβανομένων των Πολυεπίπεδων Perceptron (Multilayer Perceptrons - MLPs) και των πιο εξειδικευμένων αρχιτεκτονικών όπως τα Επαναληπτικά Νευρωνικά Δίκτυα (RNNs) και τα LSTMs.

Η διαδικασία εκπαίδευσης ενός μοντέλου μηχανικής μάθησης τυπικά περιλαμβάνει τα ακόλουθα βήματα:

1. **Συλλογή και Προετοιμασία Δεδομένων:** Συγκέντρωση των σχετικών δεδομένων, καθαρισμός, μετασχηματισμός και διαχωρισμός σε σύνολα εκπαίδευσης (training set), επικύρωσης (validation set) και ελέγχου (test set).
2. **Επιλογή Μοντέλου:** Επιλογή της κατάλληλης αρχιτεκτονικής μοντέλου με βάση

τη φύση του προβλήματος και των δεδομένων.

3. **Ορισμός Συνάρτησης Απώλειας (Loss Function):** Μια συνάρτηση που ποσοτικοποιεί το σφάλμα του μοντέλου, δηλαδή πόσο μακριά είναι οι προβλέψεις του από τις πραγματικές τιμές.
4. **Εκπαίδευση (Training):** Χρήση ενός αλγορίθμου βελτιστοποίησης (π.χ., Gradient Descent και οι παραλλαγές του) για την προσαρμογή των παραμέτρων (βαρών) του μοντέλου έτσι ώστε να ελαχιστοποιείται η συνάρτηση απώλειας στα δεδομένα εκπαίδευσης.
5. **Αξιολόγηση (Evaluation):** Αξιολόγηση της απόδοσης του εκπαιδευμένου μοντέλου στο σύνολο ελέγχου (και ενδεχομένως στο σύνολο επικύρωσης κατά τη διάρκεια της εκπαίδευσης) χρησιμοποιώντας κατάλληλες μετρικές (π.χ., ακρίβεια, μέσο τετραγωνικό σφάλμα).
6. **Βελτιστοποίηση Υπερπαραμέτρων (Hyperparameter Tuning):** Περαιτέρω βελτίωση της απόδοσης του μοντέλου μέσω της ρύθμισης των υπερπαραμέτρων του (π.χ., ρυθμός μάθησης, αριθμός επιπέδων σε ένα νευρωνικό δίκτυο).

Η επιλογή του κατάλληλου μοντέλου και η σωστή διαδικασία εκπαίδευσης είναι κρίσιμες για την επιτυχία οποιασδήποτε εφαρμογής μηχανικής μάθησης. Όπως αναφέρεται στο [1], για προβλήματα με ακολουθιακά δεδομένα όπως οι χρονοσειρές, παραδοσιακά στατικά μοντέλα όπως τα MLPs ή τα δέντρα απόφασης μπορεί να μην είναι η βέλτιστη επιλογή, καθώς είτε αγνοούν τη σειριακή δομή των δεδομένων είτε απαιτούν εκτεταμένη και συχνά χειροκίνητη μηχανική χαρακτηριστικών (feature engineering) για να συλλάβουν τις χρονικές εξαρτήσεις. Αυτό υπογραμμίζει την ανάγκη για μοντέλα που είναι εγγενώς σχεδιασμένα για την επεξεργασία ακολουθιών, όπως τα RNNs και τα LSTMs. Η κατανόηση των θεμελιωδών αρχών της μηχανικής μάθησης, όπως οι συναρτήσεις απώλειας και οι αλγόριθμοι βελτιστοποίησης (π.χ., backpropagation), είναι επίσης απαραίτητη για την κατανόηση της λειτουργίας των πιο προηγμένων μοντέλων και των μεθόδων εξηγησιμότητας που εφαρμόζονται σε αυτά, ειδικά των μεθόδων που βασίζονται στις κλίσεις (gradient-based) όπως το DeepLIFT και το Integrated Gradients.

3.2 Μοντέλα Μηχανικής Μάθησης για Προβλέψεις Δεδομένων Χρονοσειρών

Τα δεδομένα χρονοσειρών (time series data) παρουσιάζουν μοναδικές ιδιαιτερότητες που τα διαφοροποιούν από τα στατικά, μη ακολουθιακά δεδομένα. Η πιο θεμελιώδης ιδιαιτερότητα είναι η **χρονική εξάρτηση (temporal dependence)**, όπου η τιμή μιας μεταβλητής σε μια δεδομένη χρονική στιγμή συχνά εξαρτάται από

τις προηγούμενες τιμές της. Άλλα χαρακτηριστικά μπορεί να περιλαμβάνουν την **εποχικότητα (seasonality)** (επαναλαμβανόμενα μοτίβα σε τακτά χρονικά διαστήματα), τις **τάσεις (trends)** (μακροπρόθεσμη αύξηση ή μείωση) και τη **μη στασιμότητα (non-stationarity)** (όπου οι στατιστικές ιδιότητες της σειράς, όπως η μέση τιμή και η διακύμανση, αλλάζουν με την πάροδο του χρόνου). Η αποτελεσματική μοντελοποίηση αυτών των χαρακτηριστικών είναι κρίσιμη για την ακριβή πρόβλεψη μελλοντικών τιμών.

Παραδοσιακά, η ανάλυση χρονοσειρών βασιζόταν σε στατιστικά μοντέλα όπως:

- **ARIMA (Autoregressive Integrated Moving Average):** Ένα από τα πιο διαδεδομένα μοντέλα, το ARIMA συνδυάζει αυτοπαλινδρομικές (AR) συνιστώσες (όπου η τρέχουσα τιμή εξαρτάται από τις προηγούμενες τιμές), ολοκληρωμένες (I) συνιστώσες (που αφορούν τη διαφοροποίηση της σειράς για την επίτευξη στασιμότητας) και συνιστώσες κινητού μέσου όρου (MA) (όπου η τρέχουσα τιμή εξαρτάται από προηγούμενα σφάλματα πρόβλεψης).
- **SARIMA (Seasonal ARIMA):** Μια επέκταση του ARIMA που ενσωματώνει εποχικές συνιστώσες.
- **Exponential Smoothing:** Μια οικογένεια μεθόδων που αποδίδουν εκθετικά μειούμενα βάρη σε παλαιότερες παρατηρήσεις.

Αν και αυτά τα παραδοσιακά μοντέλα είναι χρήσιμα, συχνά βασίζονται σε υποθέσεις γραμμικότητας και στασιμότητας των δεδομένων, και μπορεί να δυσκολεύονται να συλλάβουν πολύπλοκες μη-γραμμικές εξαρτήσεις ή μακροπρόθεσμες σχέσεις, ειδικά σε χρηματοοικονομικές χρονοσειρές που είναι γνωστές για τη μεταβλητότητα και τη μη-γραμμική τους συμπεριφορά.

Τα μοντέλα μηχανικής μάθησης προσφέρουν πιο ευέλικτες προσεγγίσεις για την πρόβλεψη χρονοσειρών:

- **Χρήση Παραθύρων (Windowing):** Μια κοινή τεχνική είναι η μετατροπή του προβλήματος χρονοσειράς σε ένα πρόβλημα επιβλεπόμενης μάθησης. Αυτό γίνεται δημιουργώντας "παράθυρα" από προηγούμενες παρατηρήσεις (π.χ., οι τιμές των τελευταίων N ημερών) ως χαρακτηριστικά εισόδου (features) και την τιμή της επόμενης χρονικής περιόδου (π.χ., η τιμή της (N+1)-οστής ημέρας) ως μεταβλητή στόχο (target).
- **Μοντέλα Βασισμένα σε Δέντρα:** Αλγόριθμοι όπως τα Random Forests και το Gradient Boosting μπορούν να εφαρμοστούν σε δεδομένα χρονοσειρών μετά από κατάλληλη μηχανική χαρακτηριστικών (feature engineering) που εξάγει χρονικές πληροφορίες (π.χ., υστερήσεις, κινητοί μέσοι όροι, στατιστικά παραθύρων).

- **Επαναληπτικά Νευρωνικά Δίκτυα (Recurrent Neural Networks - RNNs):** Τα RNNs είναι ειδικά σχεδιασμένα για την επεξεργασία ακολουθιακών δεδομένων. Διαθέτουν βρόχους ανάδρασης που τους επιτρέπουν να διατηρούν μια "μνήμη" των προηγούμενων εισόδων, καθιστώντας τα κατάλληλα για την εκμάθηση χρονικών εξαρτήσεων. Ωστόσο, τα απλά RNNs υποφέρουν από το πρόβλημα της **εξασθένισης ή έκρηξης των κλίσεων (vanishing or exploding gradients)**, το οποίο καθιστά δύσκολη την εκμάθηση μακροπρόθεσμων εξαρτήσεων (δηλαδή, εξαρτήσεων μεταξύ χρονικά απομακρυσμένων σημείων στην ακολουθία).
- **Long Short-Term Memory (LSTM) και Gated Recurrent Units (GRU):** Τα LSTMs και τα GRUs είναι εξειδικευμένες αρχιτεκτονικές RNN που σχεδιάστηκαν για να αντιμετωπίσουν το πρόβλημα των vanishing/exploding gradients. Χρησιμοποιούν "πύλες" (gates) για να ελέγχουν τη ροή της πληροφορίας και να διατηρούν ή να απορρίπτουν επιλεκτικά πληροφορίες από τη μνήμη τους, επιτρέποντάς τους να μαθαίνουν αποτελεσματικά μακροπρόθεσμες εξαρτήσεις. Τα LSTMs, ειδικότερα, έχουν αποδειχθεί πολύ αποτελεσματικά στην πρόβλεψη χρηματοοικονομικών χρονοσειρών.
- **Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks - CNNs):** Αν και αρχικά αναπτύχθηκαν για την επεξεργασία εικόνων, τα CNNs μπορούν επίσης να εφαρμοστούν σε μονοδιάστατα ακολουθιακά δεδομένα (όπως οι χρονοσειρές) για την αυτόματη εξαγωγή χαρακτηριστικών. Συχνά χρησιμοποιούνται σε υβριδικές αρχιτεκτονικές, όπως τα CNN-LSTM, όπου το CNN εξάγει χαρακτηριστικά από υπο-ακολουθίες και το LSTM μοντελοποιεί τις χρονικές σχέσεις μεταξύ αυτών των χαρακτηριστικών.
- **Transformers και Μηχανισμοί Προσοχής (Attention Mechanisms):** Πιο πρόσφατα, οι αρχιτεκτονικές Transformer, που αρχικά έφεραν επανάσταση στην επεξεργασία φυσικής γλώσσας, έχουν δείξει εξαιρετικά αποτελέσματα και στην πρόβλεψη χρονοσειρών. Οι μηχανισμοί αυτο-προσοχής (self-attention) που χρησιμοποιούν τους επιτρέπουν να σταθμίζουν τη σημασία διαφορετικών χρονικών σημείων στην ακολουθία εισόδου, συλλαμβάνοντας πολύπλοκες εξαρτήσεις. Η χρήση μηχανισμών προσοχής αναφέρεται ως πιθανή μελλοντική κατεύθυνση και στο.

Η μετάβαση από παραδοσιακά στατιστικά μοντέλα σε μοντέλα μηχανικής μάθησης, και ειδικότερα σε μοντέλα βαθιάς μάθησης όπως τα LSTMs, για την πρόβλεψη χρονοσειρών, οφείλεται σε μεγάλο βαθμό στην ικανότητα των τελευταίων να μοντελοποιούν πολύπλοκες, μη-γραμμικές δυναμικές και μακροχρόνιες εξαρτήσεις που συχνά χαρακτηρίζουν τις χρηματοοικονομικές αγορές.Ενώ ένα μοντέλο ARIMA μπορεί να υποθέτει γραμμικότητα και στασιμότητα, τα LSTMs είναι σχεδιασμένα για να χειρίζονται ακριβώς τις μη-γραμμικότητες και τις χρονικές εξαρτήσεις που

καθιστούν τις χρηματοοικονομικές προβλέψεις τόσο δύσκολες. Ωστόσο, αυτή η αυξημένη πολυπλοκότητα και προβλεπτική ισχύς έρχεται με το κόστος της μειωμένης ερμηνευσιμότητας. Τα LSTMs, όντας βαθιά νευρωνικά δίκτυα, λειτουργούν συχνά ως "μαύρα κουτιά", καθιστώντας δύσκολη την κατανόηση του *πώς* φτάνουν στις προβλέψεις τους. Αυτή ακριβώς η πρόκληση είναι που οδηγεί στην ανάγκη για μεθόδους XAI, όπως αυτές που διερευνώνται στην παρούσα διπλωματική εργασία, προκειμένου να αξιοποιηθεί η ισχύς αυτών των προηγμένων μοντέλων διατηρώντας παράλληλα ένα επίπεδο κατανόησης και εμπιστοσύνης στις αποφάσεις τους.

3.3 Μοντέλα LSTM: Εις Βάθος Ανάλυση, Λειτουργία και Μαθηματική Θεμελίωση

Τα δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (Long Short-Term Memory - LSTM) αποτελούν μια εξειδικευμένη και ιδιαίτερα επιτυχημένη αρχιτεκτονική Επαναληπτικών Νευρωνικών Δικτύων (Recurrent Neural Networks - RNNs). Προτάθηκαν αρχικά από τους Hochreiter και Schmidhuber το 1997 και σχεδιάστηκαν ειδικά για να αντιμετωπίσουν ένα θεμελιώδες πρόβλημα που ταλαιπωρούσε τα παραδοσιακά RNNs: την αδυναμία εκμάθησης μακροπρόθεσμων εξαρτήσεων λόγω του φαινομένου της εξασθένισης ή έκρηξης των κλίσεων (vanishing or exploding gradients).

Εισαγωγή στα RNNs και το Πρόβλημα των Vanishing/Exploding Gradients

Τα RNNs είναι νευρωνικά δίκτυα σχεδιασμένα για την επεξεργασία ακολουθιακών δεδομένων, όπως κείμενο, ομιλία ή χρονοσειρές. Η βασική τους ιδέα είναι η χρήση βρόχων ανάδρασης, όπου η έξοδος ενός νευρώνα σε ένα χρονικό βήμα τροφοδοτείται ξανά ως είσοδος στον ίδιο νευρώνα (ή σε άλλους νευρώνες του ίδιου επιπέδου) στο επόμενο χρονικό βήμα. Αυτό επιτρέπει στο δίκτυο να διατηρεί μια "μνήμη" των προηγούμενων πληροφοριών της ακολουθίας.

Κατά την εκπαίδευση των RNNs με τον αλγόριθμο backpropagation through time (BPTT), οι κλίσεις του σφάλματος διαδίδονται προς τα πίσω μέσα από τα χρονικά βήματα της ακολουθίας. Σε βαθιά δίκτυα ή σε πολύ μεγάλες ακολουθίες, αυτές οι κλίσεις υπόκεινται σε επαναλαμβανόμενους πολλαπλασιασμούς με τις παραγώγους των συναρτήσεων ενεργοποίησης και τα βάρη του δικτύου. Εάν αυτές οι παράγωγοι είναι συστηματικά μικρότερες από 1, οι κλίσεις τείνουν να γίνονται εκθετικά μικρές (vanishing gradients), καθιστώντας την ενημέρωση των βαρών των πρώιμων επιπέδων ή χρονικών βημάτων αμελητέα. Αυτό εμποδίζει το δίκτυο να μάθει εξαρτήσεις μεταξύ χρονικά απομακρυσμένων γεγονότων. Αντίστροφα, εάν οι

παράγωγοι είναι συστηματικά μεγαλύτερες από 1, οι κλίσεις μπορεί να γίνουν εκθετικά μεγάλες (exploding gradients), οδηγώντας σε ασταθή εκπαίδευση.

Η Αρχιτεκτονική ενός Κελιού LSTM

Τα LSTMs αντιμετωπίζουν αυτά τα προβλήματα μέσω μιας πιο πολύπλοκης δομής κελιού (cell), η οποία περιλαμβάνει μηχανισμούς "πυλών" (gates) που ρυθμίζουν τη ροή της πληροφορίας. Κάθε κελί LSTM διατηρεί μια **κατάσταση κελιού (cell state, C_t)**, η οποία λειτουργεί ως ο κύριος δίαυλος μεταφοράς πληροφορίας και μπορεί να διατηρήσει πληροφορίες για μεγάλα χρονικά διαστήματα. Επιπλέον, υπάρχει η **κρυφή κατάσταση (hidden state, h_t)**, η οποία είναι η έξοδος του κελιού σε κάθε χρονικό βήμα t.

Οι τρεις κύριες πύλες ενός τυπικού κελιού LSTM είναι:

1. Πύλη Λήθης (Forget Gate, ft): Αυτή η πύλη αποφασίζει ποιες πληροφορίες από την προηγούμενη κατάσταση κελιού C_{t-1} πρέπει να "ξεχαστούν" ή να απορριφθούν. Λαμβάνει ως είσοδο την προηγούμενη κρυφή κατάσταση h_{t-1} και την τρέχουσα είσοδο (xt) και παράγει έναν αριθμό μεταξύ 0 και 1 για κάθε στοιχείο της κατάστασης κελιού C_{t-1} . Η τιμή 1 σημαίνει "διατήρησε πλήρως αυτή την πληροφορία", ενώ η τιμή 0 σημαίνει "απόρριψε πλήρως αυτή την πληροφορία".
   Η μαθηματική της εξίσωση είναι:
   $f_t = \sigma\,(W_f \cdot [h_{t-1}, x_t] + b_f)$

   όπου σ είναι η σιγμοειδής συνάρτηση ενεργοποίησης, W_f είναι ο πίνακας βαρών της πύλης λήθης, και b_f είναι το διάνυσμα πόλωσης (bias) της πύλης λήθης. Το $[h_{t-1}, x_t]$ δηλώνει τη συνένωση (concatenation) των δύο διανυσμάτων.
2. **Πύλη Εισόδου (Input Gate, it):** Αυτή η πύλη αποφασίζει ποιες νέες πληροφορίες από την τρέχουσα είσοδο θα αποθηκευτούν στην κατάσταση κελιού. Η λειτουργία της χωρίζεται σε δύο μέρη:
   - Μια σιγμοειδής συνάρτηση (it) που αποφασίζει ποιες τιμές θα ενημερωθούν: $i_t = \sigma\,(W_i \cdot [h_{t-1}, x_t] + b_i)$
   - Μια συνάρτηση tanh (C_t) που δημιουργεί ένα διάνυσμα νέων υποψήφιων τιμών που θα μπορούσαν να προστεθούν στην κατάσταση κελιού:

$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$

Οι W_i, b_i και W_C, b_C είναι οι αντίστοιχοι πίνακες βαρών και διανύσματα πόλωσης.

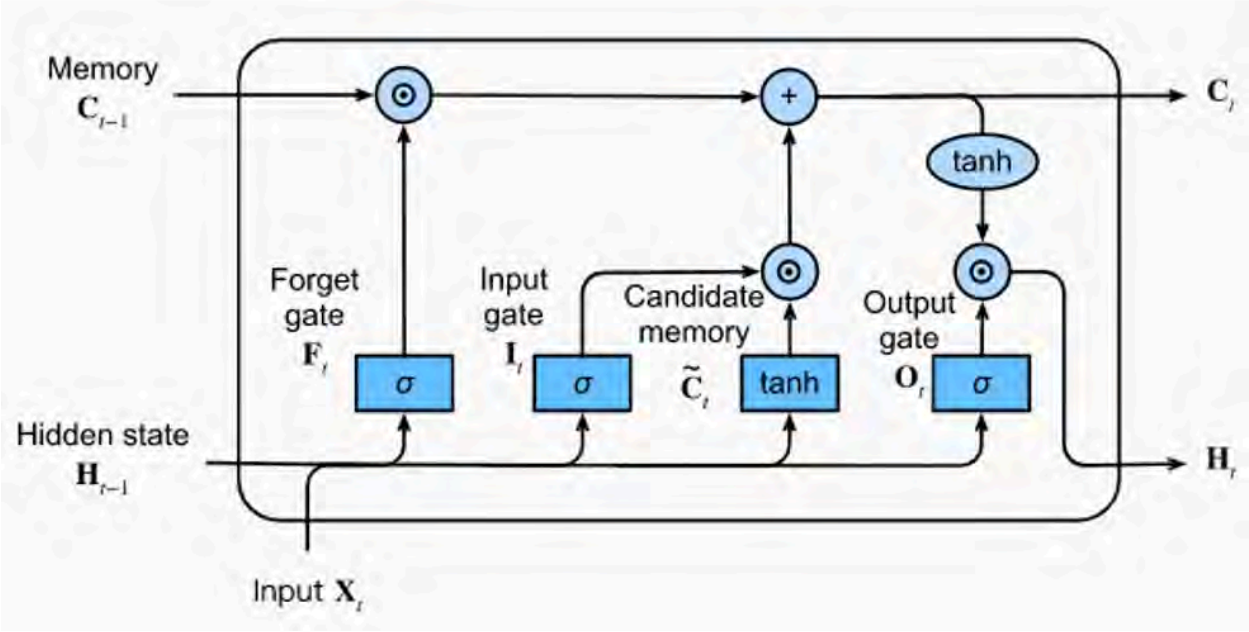
3. Πύλη Εξόδου (Output Gate, o_t): Αυτή η πύλη καθορίζει ποιο μέρος της (ενημερωμένης) κατάστασης κελιού θα εξαχθεί ως η τρέχουσα κρυφή κατάσταση (h_t). Πρώτα, μια σιγμοειδής συνάρτηση αποφασίζει ποια μέρη της κατάστασης κελιού θα περάσουν:
$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$

Οι W_o, b_o είναι ο πίνακας βαρών και το διάνυσμα πόλωσης της πύλης εξόδου.

Ενημέρωση Κατάστασης Κελιού και Κρυφής Κατάστασης

Με βάση τις εξόδους των πυλών, η κατάσταση κελιού και η κρυφή κατάσταση ενημερώνονται ως εξής:

- Ενημέρωση Κατάστασης Κελιού (Ct):
$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$
Αυτή η εξίσωση συνδυάζει την προηγούμενη κατάσταση κελιού (αφού περάσει από την πύλη λήθης, $f_t \odot C_{t-1}$) με τις νέες υποψήφιες τιμές (αφού περάσουν από την πύλη εισόδου, $i_t \odot \tilde{C}_t$). Το σύμβολο ⊙ δηλώνει τον κατά στοιχείο πολλαπλασιασμό (Hadamard product).
- Ενημέρωση Κρυφής Κατάστασης (h_t):
$h_t = o_t \odot tanh(C_t)$
Η κρυφή κατάσταση (που είναι και η έξοδος του κελιού LSTM για το χρονικό βήμα t) προκύπτει πολλαπλασιάζοντας την έξοδο της πύλης εξόδου (o_t) με την κατάσταση κελιού (C_t) αφού αυτή περάσει από μια συνάρτηση tanh (για να περιοριστούν οι τιμές μεταξύ -1 και 1).



Εικόνα 3.1: Αρχιτεκτονική ενός μοντέλου LSTM

Η "ευφυΐα" των LSTMs έγκειται ακριβώς στην ικανότητά τους να ρυθμίζουν επιλεκτικά τη ροή της πληροφορίας μέσω αυτών των πυλών. Η πύλη λήθης επιτρέπει στο μοντέλο να απορρίπτει παλιές, άσχετες πληροφορίες από τη μνήμη του, ενώ η πύλη εισόδου επιτρέπει την ενσωμάτωση νέων, σχετικών πληροφοριών. Αυτή η επιλεκτική διαχείριση της μνήμης είναι κρίσιμη για την αποφυγή του προβλήματος της εξασθένισης της κλίσης και για την εκμάθηση εξαρτήσεων που εκτείνονται σε μεγάλα χρονικά διαστήματα. Για παράδειγμα, στην ανάλυση μιας μετοχής, ένα γεγονός που συνέβη πριν από πολλούς μήνες μπορεί να είναι πλέον άσχετο για την πρόβλεψη της αυριανής τιμής, ενώ μια πρόσφατη ανακοίνωση κερδών μπορεί να είναι εξαιρετικά σημαντική. Το LSTM, μέσω των πυλών του, μπορεί να μάθει να "ξεχνά" την παλιά πληροφορία και να εστιάζει στην πρόσφατη και σχετική.

Η διάκριση μεταξύ της κατάστασης κελιού (C_t) και της κρυφής κατάστασης (h_t) είναι επίσης θεμελιώδης. Η κατάσταση κελιού μπορεί να θεωρηθεί ως η "μακροπρόθεσμη μνήμη" του δικτύου, ικανή να μεταφέρει πληροφορίες σχετικά αναλλοίωτες για πολλά χρονικά βήματα. Η κρυφή κατάσταση, από την άλλη πλευρά, είναι μια "φιλτραρισμένη" εκδοχή της κατάστασης κελιού, προσαρμοσμένη για την παραγωγή της τρέχουσας εξόδου και για την τροφοδοσία στο επόμενο κελί ως "βραχυπρόθεσμη μνήμη". Αυτή η αρχιτεκτονική είναι που προσδίδει στα LSTMs την

ισχύ τους στη μοντελοποίηση πολύπλοκων ακολουθιακών δεδομένων.

Η εκπαίδευση των LSTMs, όπως και των απλών RNNs, γίνεται συνήθως με τον αλγόριθμο Backpropagation Through Time (BPTT), ο οποίος ξεδιπλώνει το επαναληπτικό δίκτυο στον χρόνο και εφαρμόζει τον τυπικό αλγόριθμο backpropagation.

3.4 Εφαρμογή των LSTMs στις Προβλέψεις της Χρηματιστηριακής Αγοράς

Τα δίκτυα LSTM έχουν αναδειχθεί ως ένα ιδιαίτερα δημοφιλές και αποτελεσματικό εργαλείο για την πρόβλεψη των χρηματιστηριακών αγορών, κυρίως λόγω της εγγενούς ικανότητάς τους να μοντελοποιούν πολύπλοκα χαρακτηριστικά των χρηματοοικονομικών χρονοσειρών. Οι χρηματοοικονομικές αγορές χαρακτηρίζονται από:

- **Μη-γραμμικότητα:** Οι σχέσεις μεταξύ των παραγόντων που επηρεάζουν τις τιμές των μετοχών σπάνια είναι απλές ή γραμμικές.
- **Μεταβλητότητα (Volatility):** Οι τιμές μπορεί να παρουσιάζουν περιόδους ηρεμίας που ακολουθούνται από απότομες και έντονες διακυμάνσεις (volatility clustering).
- **Μακροπρόθεσμες Εξαρτήσεις:** Γεγονότα ή τάσεις του παρελθόντος μπορεί να επηρεάζουν τις τρέχουσες και μελλοντικές τιμές, ακόμη και αν μεσολαβεί μεγάλο χρονικό διάστημα.
- **Θόρυβος:** Τα χρηματοοικονομικά δεδομένα περιέχουν σημαντικό βαθμό τυχαιότητας ή "θορύβου" που καθιστά δύσκολη την απομόνωση των πραγματικών σημάτων.

Τα LSTMs είναι θεωρητικά καλά εξοπλισμένα για να αντιμετωπίσουν αυτές τις προκλήσεις. Η ικανότητά τους να διατηρούν μνήμη για μακρές περιόδους και να μαθαίνουν μη-γραμμικές σχέσεις τα καθιστά ανώτερα από πολλά παραδοσιακά στατιστικά μοντέλα, όπως το ARIMA, ειδικά όταν τα δεδομένα δεν ακολουθούν γραμμικά πρότυπα ή όταν υπάρχουν πολύπλοκες χρονικές εξαρτήσεις. Μελέτες έχουν δείξει ότι τα LSTMs μπορούν να επιτύχουν καλύτερη ακρίβεια πρόβλεψης σε σύγκριση με αυτά τα παραδοσιακά μοντέλα για ορισμένους τύπους χρηματοοικονομικών δεδομένων.

Οι εφαρμογές των LSTMs στις χρηματοοικονομικές αγορές είναι ποικίλες και περιλαμβάνουν:

- **Πρόβλεψη Τιμών Μετοχών:** Προσπάθεια πρόβλεψης της μελλοντικής τιμής

κλεισίματος, ανοίγματος, υψηλότερης ή χαμηλότερης τιμής μιας μετοχής.
- **Πρόβλεψη Κατεύθυνσης Τιμών:** Πρόβλεψη εάν η τιμή μιας μετοχής θα κινηθεί ανοδικά, πτωτικά ή θα παραμείνει στάσιμη (πρόβλημα ταξινόμησης).
- **Πρόβλεψη Σημάτων Συναλλαγών:** Όπως και στην παρούσα διπλωματική εργασία, πρόβλεψη της βέλτιστης ενέργειας (Αγορά, Πώληση, Διακράτηση) για έναν επενδυτή.
- **Πρόβλεψη Μεταβλητότητας:** Μοντελοποίηση και πρόβλεψη της μελλοντικής μεταβλητότητας μιας μετοχής ή της αγοράς.

Ένα σημαντικό πλεονέκτημα των LSTMs είναι η ικανότητά τους να επεξεργάζονται **πολυμεταβλητές χρονοσειρές (multivariate time series)**. Αυτό σημαίνει ότι μπορούν να λαμβάνουν ως είσοδο όχι μόνο την ιστορική πορεία της τιμής μιας μετοχής, αλλά και ένα ευρύ φάσμα άλλων σχετικών πληροφοριών, όπως:

- **Τεχνικοί Δείκτες:** RSI, MACD, Κινητοί Μέσοι Όροι, Όγκος Συναλλαγών, Bollinger Bands κ.λπ. Αυτοί οι δείκτες είναι συχνά μετασχηματισμοί των πρωτογενών δεδομένων τιμής και όγκου, σχεδιασμένοι για να αναδεικνύουν συγκεκριμένα πρότυπα ή συνθήκες της αγοράς. Η παρούσα εργασία αξιοποιεί εκτενώς τέτοιους δείκτες.
- **Μακροοικονομικά Δεδομένα:** Δείκτες όπως τα επιτόκια, ο πληθωρισμός, η ανεργία, η πορεία γενικών δεικτών της αγοράς (π.χ., S&P 500), ή δείκτες μεταβλητότητας (π.χ., VIX). Αυτοί οι παράγοντες παρέχουν ένα ευρύτερο πλαίσιο για τις συνθήκες της αγοράς.
- **Ανάλυση Συναισθήματος (Sentiment Analysis):** Πληροφορίες που προέρχονται από την ανάλυση ειδησεογραφικών άρθρων, αναρτήσεων σε μέσα κοινωνικής δικτύωσης (social media) ή οικονομικών εκθέσεων, με σκοπό την ποσοτικοποίηση του επενδυτικού κλίματος. Η ενσωμάτωση τέτοιων δεδομένων αναφέρεται ως πιθανή μελλοντική επέκταση στο.

Παρά τα πλεονεκτήματά τους, η εφαρμογή των LSTMs στις χρηματοοικονομικές προβλέψεις ενέχει και ορισμένες προκλήσεις:

- **Ανάγκη για Μεγάλο Όγκο Δεδομένων:** Τα βαθιά νευρωνικά δίκτυα, συμπεριλαμβανομένων των LSTMs, απαιτούν συνήθως μεγάλες ποσότητες δεδομένων εκπαίδευσης για να μάθουν αποτελεσματικά τα υποκείμενα πρότυπα και να αποφύγουν την υπερπροσαρμογή.
- **Υπολογιστική Πολυπλοκότητα:** Η εκπαίδευση των LSTMs μπορεί να είναι υπολογιστικά απαιτητική, ειδικά με μεγάλα σύνολα δεδομένων και πολύπλοκες αρχιτεκτονικές.
- **Κίνδυνος Υπερπροσαρμογής (Overfitting):** Εάν το μοντέλο είναι υπερβολικά πολύπλοκο σε σχέση με τον όγκο των διαθέσιμων δεδομένων, υπάρχει κίνδυνος

να "αποστηθίσει" τα δεδομένα εκπαίδευσης (συμπεριλαμβανομένου του θορύβου) αντί να μάθει γενικεύσιμα πρότυπα. Αυτό οδηγεί σε κακή απόδοση σε νέα, μη παρατηρηθέντα δεδομένα.
- **Ερμηνευσιμότητα:** Όπως αναφέρθηκε, τα LSTMs είναι εκ φύσεως "μαύρα κουτιά". Η έλλειψη διαφάνειας ως προς το πώς λαμβάνουν τις αποφάσεις τους αποτελεί σημαντικό εμπόδιο για την ευρεία υιοθέτησή τους σε κρίσιμες εφαρμογές. Αυτή ακριβώς η πρόκληση αντιμετωπίζεται στην παρούσα εργασία μέσω της χρήσης μεθόδων XAI.

Η έρευνα στον τομέα είναι συνεχής, με νέες παραλλαγές και υβριδικά μοντέλα να προτείνονται τακτικά, όπως συνδυασμοί LSTM με CNNs ή Graph Neural Networks (GNNs), με στόχο την περαιτέρω βελτίωση της ακρίβειας και της στιβαρότητας των προβλέψεων. Η επιλογή της κατάλληλης αρχιτεκτονικής και των χαρακτηριστικών εισόδου παραμένει ένα κρίσιμο βήμα για την επιτυχή εφαρμογή των LSTMs στις χρηματοοικονομικές αγορές.

3.5 Πολυ-εισοδικά (Multi-input) LSTMs και η Λογική Πίσω από Αυτά

Τα πολυ-εισοδικά μοντέλα LSTM (Multi-input LSTM models) αποτελούν μια επέκταση της βασικής αρχιτεκτονικής LSTM, σχεδιασμένη για να επεξεργάζεται ταυτόχρονα πολλαπλές ακολουθίες εισόδου. Αυτή η προσέγγιση επιτρέπει στο μοντέλο να μάθει από διαφορετικές πηγές πληροφορίας ή διαφορετικές αναπαραστάσεις των δεδομένων, οδηγώντας δυνητικά σε πιο εύρωστες και ακριβείς προβλέψεις. Κάθε ακολουθία εισόδου μπορεί να αντιπροσωπεύει ένα διαφορετικό χαρακτηριστικό ή μια διαφορετική οπτική της υπό εξέταση χρονοσειράς.

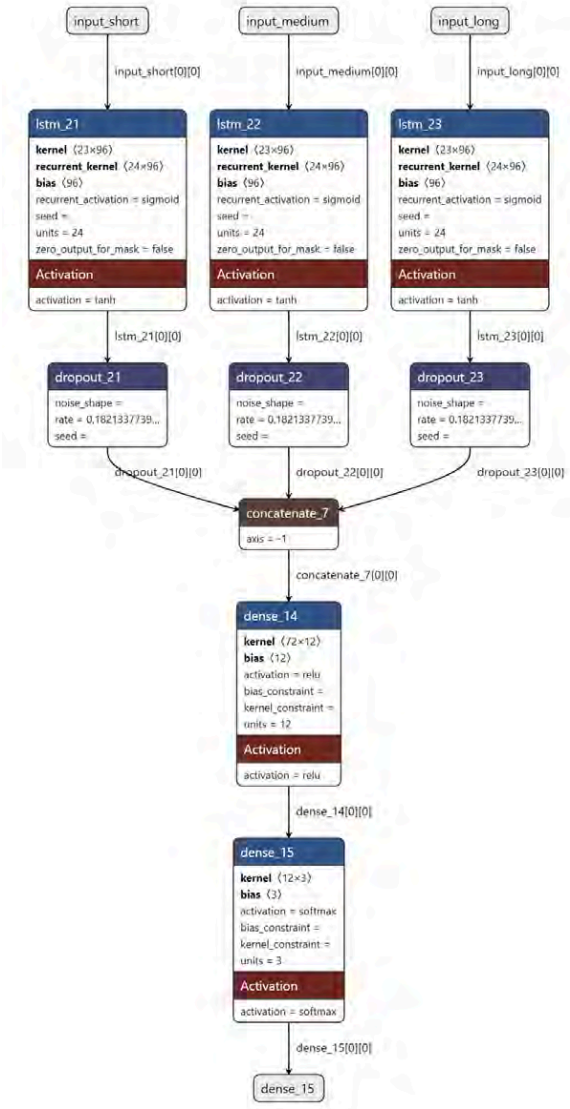
Λογική και Αρχιτεκτονική:

Η βασική ιδέα πίσω από τα πολυ-εισοδικά LSTMs είναι η δημιουργία ξεχωριστών «κλάδων» εισόδου στο δίκτυο, όπου κάθε κλάδος επεξεργάζεται μια συγκεκριμένη ακολουθία δεδομένων χρησιμοποιώντας ένα ή περισσότερα επίπεδα LSTM. Αφού κάθε ακολουθία εισόδου υποβληθεί σε αρχική επεξεργασία από τον αντίστοιχο κλάδο LSTM, οι εξόδοι αυτών των κλάδων συνήθως συγχωνεύονται ή συνενώνονται (concatenated) για να σχηματίσουν μια ενιαία αναπαράσταση. Αυτή η συνδυασμένη αναπαράσταση μπορεί στη συνέχεια να τροφοδοτηθεί σε περαιτέρω επίπεδα LSTM ή σε πυκνά συνδεδεμένα επίπεδα (dense layers) για την τελική πρόβλεψη.

Υπάρχουν διάφοροι τρόποι με τους οποίους μπορούν να σχεδιαστούν τα

πολυ-εισοδικά LSTMs:

1. **Παράλληλη Επεξεργασία και Σύζευξη (Parallel Processing and Concatenation):** Κάθε ακολουθία εισόδου τροφοδοτείται σε ένα ξεχωριστό υποδίκτυο LSTM. Οι λανθάνουσες αναπαραστάσεις (hidden states ή outputs) που παράγονται από κάθε υποδίκτυο LSTM στη συνέχεια συνενώνονται. Αυτός ο συνενωμένος φορέας χαρακτηριστικών τροφοδοτείται έπειτα σε ένα ή περισσότερα πυκνά επίπεδα για την τελική πρόβλεψη.



Εικόνα 3.3: Παράλληλη Επεξεργασία και Σύζευξη υποδικτύων LSTM

2. **Ιεραρχική Επεξεργασία:** Σε ορισμένες περιπτώσεις, οι έξοδοι από ορισμένους κλάδους LSTM μπορούν να χρησιμοποιηθούν ως είσοδοι σε άλλους κλάδους ή σε ένα ανώτερο επίπεδο LSTM που επεξεργάζεται τις

συνδυασμένες πληροφορίες με πιο αφηρημένο τρόπο.

3. **Διαφορετικά Μήκη Ακολουθιών:** Τα πολυ-εισοδικά LSTMs μπορούν να χειριστούν ακολουθίες εισόδου με διαφορετικά μήκη, αν και αυτό συχνά απαιτεί κατάλληλη προετοιμασία των δεδομένων, όπως padding ή χρήση ειδικών τεχνικών για την ευθυγράμμιση των ακολουθιών πριν από τη σύζευξη.

Μαθηματικό Υπόβαθρο:

Η μαθηματική βάση κάθε μεμονωμένου κλάδου LSTM παραμένει η ίδια όπως περιγράφηκε στην ενότητα 3.3. Η κύρια διαφορά έγκειται στον τρόπο με τον οποίο οι έξοδοι $h_T^{(i)}$ από κάθε κλάδο i συνδυάζονται. Εάν έχουμε N ακολουθίες εισόδου, και κάθε ακολουθία $X^{(i)}$ επεξεργάζεται από ένα ξεχωριστό LSTM παράγοντας μια τελική λανθάνουσα κατάσταση $h_T^{(i)}$ (όπου T είναι το μήκος της ακολουθίας), τότε η συνένωση αυτών των καταστάσεων παράγει έναν φορέα hmerged:

$h_{merged} = [h_T^{(1)}, h_T^{(2)}, \dots, h_T^{(N)}]$

Αυτός ο h_{merged} φορέας στη συνέχεια χρησιμοποιείται ως είσοδος σε ένα πυκνό επίπεδο με μια συνάρτηση ενεργοποίησης (π.χ., softmax για προβλήματα ταξινόμησης όπως η πρόβλεψη δράσης trader):

$y = softmax(W \cdot h_{merged} + b)$

όπου W είναι ο πίνακας βαρών και b είναι ο όρος πόλωσης του πυκνού επιπέδου.

Πλεονεκτήματα των Πολυ-εισοδικών LSTMs:

1. **Πλουσιότερη Αναπαράσταση Χαρακτηριστικών:** Επιτρέπουν στο μοντέλο να αξιοποιήσει πληροφορίες από πολλαπλές πηγές ή οπτικές, οδηγώντας σε μια πιο ολοκληρωμένη κατανόηση των υποκείμενων μοτίβων. Για παράδειγμα, στην πρόβλεψη μετοχών, θα μπορούσε κανείς να χρησιμοποιήσει ως ξεχωριστές εισόδους:
    - Τις ιστορικές τιμές κλεισίματος της μετοχής.
    - Τον όγκο συναλλαγών.
    - Τεχνικούς δείκτες (π.χ., MACD, RSI).
    - Δεδομένα ανάλυσης συναισθήματος από ειδήσεις ή μέσα κοινωνικής

δικτύωσης.

2. **Βελτιωμένη Απόδοση:** Η ικανότητα ενσωμάτωσης ποικίλων πληροφοριών μπορεί να οδηγήσει σε σημαντική βελτίωση της ακρίβειας των προβλέψεων, ειδικά σε πολύπλοκα συστήματα όπου πολλαπλοί παράγοντες αλληλεπιδρούν.
3. **Ευελιξία στον Σχεδιασμό:** Παρέχουν ευελιξία στον τρόπο με τον οποίο δομούνται τα δίκτυα, επιτρέποντας στους ερευνητές να προσαρμόσουν την αρχιτεκτονική στις συγκεκριμένες ανάγκες του προβλήματος.
4. **Μοντελοποίηση Σύνθετων Αλληλεξαρτήσεων:** Το μοντέλο μπορεί να μάθει τις αλληλεξαρτήσεις και τις αλληλεπιδράσεις μεταξύ των διαφορετικών ακολουθιών εισόδου, κάτι που θα ήταν δύσκολο να επιτευχθεί με ένα μονο-εισοδικό μοντέλο.

Εφαρμογή στην Πρόβλεψη Μετοχών:

Στο πλαίσιο της διπλωματικής σου, ένα πολυ-εισοδικό LSTM θα μπορούσε να εκπαιδευτεί χρησιμοποιώντας, για παράδειγμα:

- **Είσοδος 1:** Ακολουθία τιμών (Open, High, Low, Close, Adjusted Close) της μετοχής $AAPL.
- **Είσοδος 2:** Ακολουθία όγκου συναλλαγών της μετοχής $AAPL.
- **Είσοδος 3 (προαιρετικά):** Ακολουθία τεχνικών δεικτών (π.χ., Κινητός Μέσος Όρος, RSI, MACD) υπολογισμένων από τις τιμές της μετοχής.
- **Είσοδος 4 (προαιρετικά):** Ακολουθία τιμών από δείκτες αγοράς (π.χ., S&P 500, NASDAQ) για την αποτύπωση της γενικότερης τάσης της αγοράς.

Κάθε μία από αυτές τις εισόδους θα μπορούσε να επεξεργαστεί από ένα ξεχωριστό σύνολο επιπέδων LSTM, και οι έξοδοί τους θα συνενώνονταν πριν τροφοδοτηθούν σε ένα τελικό επίπεδο για την πρόβλεψη της δράσης (buy, hold, sell). Αυτή η προσέγγιση επιτρέπει στο μοντέλο να μάθει πώς οι συνδυασμένες πληροφορίες από την τιμή, τον όγκο και πιθανώς άλλους παράγοντες επηρεάζουν τις μελλοντικές κινήσεις της μετοχής.

Η χρήση πολυ-εισοδικών LSTMs στην πρόβλεψη μετοχών είναι ιδιαίτερα ελκυστική διότι οι χρηματοοικονομικές αγορές επηρεάζονται από ένα ευρύ φάσμα παραγόντων που δεν μπορούν πάντα να συλληφθούν από μια μοναδική χρονοσειρά. Ενσωματώνοντας πολλαπλές ροές δεδομένων, τα μοντέλα αυτά έχουν τη δυνατότητα να ανακαλύψουν πιο σύνθετες και λεπτές σχέσεις, οδηγώντας σε πιο τεκμηριωμένες αποφάσεις πρόβλεψης.

ΚΕΦΑΛΑΙΟ 4: ΕΞΗΓΗΣΙΜΟΤΗΤΑ

4.1 Εισαγωγή στην Εξηγησιμότητα

Τα τελευταία χρόνια, η Τεχνητή Νοημοσύνη (ΤΝ) και ειδικότερα η Μηχανική Μάθηση (ΜΜ) έχουν επιτύχει αξιοσημείωτη πρόοδο σε ένα ευρύ φάσμα εφαρμογών, από την αναγνώριση εικόνας και την επεξεργασία φυσικής γλώσσας έως την αυτόνομη οδήγηση και τις ιατρικές διαγνώσεις. Πολλά από τα πιο επιτυχημένα μοντέλα, όπως τα βαθιά νευρωνικά δίκτυα (Deep Neural Networks - DNNs) και τα ίδια τα δίκτυα LSTM που εξετάζουμε, συχνά λειτουργούν ως «μαύρα κουτιά» (black boxes). Αυτό σημαίνει ότι ενώ μπορούν να επιτύχουν εντυπωσιακή απόδοση στις προβλέψεις τους, οι εσωτερικοί μηχανισμοί και η λογική πίσω από τις αποφάσεις τους παραμένουν αδιαφανείς ή δύσκολα κατανοητοί από τον άνθρωπο.

Το πρόβλημα της αδιαφάνειας γίνεται ιδιαίτερα έντονο σε τομείς υψηλού κινδύνου (high-stakes domains) όπως η χρηματοοικονομική αγορά, η ιατρική, η δικαιοσύνη και η ασφάλεια, όπου οι αποφάσεις των συστημάτων ΤΝ μπορούν να έχουν σοβαρές οικονομικές, κοινωνικές ή ατομικές συνέπειες. Σε τέτοια πλαίσια, δεν αρκεί απλώς να γνωρίζουμε *τι* προβλέπει ένα μοντέλο, αλλά είναι εξίσου, αν όχι περισσότερο, σημαντικό να κατανοούμε *γιατί* κάνει μια συγκεκριμένη πρόβλεψη.

Η **Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable Artificial Intelligence - XAI)** αναδύθηκε ως ένας κλάδος έρευνας που στοχεύει στην ανάπτυξη μεθόδων και τεχνικών που καθιστούν τις αποφάσεις των μοντέλων ΤΝ κατανοητές από τον άνθρωπο. Ο απώτερος στόχος της XAI είναι να παρέχει διαφάνεια, επιτρέποντας στους χρήστες, τους προγραμματιστές, τους ρυθμιστικούς φορείς και την κοινωνία γενικότερα να:

- **Κατανοήσουν** πώς λειτουργούν τα μοντέλα.
- **Εμπιστευτούν** τις προβλέψεις και τις αποφάσεις τους.
- **Εντοπίσουν και διορθώσουν** πιθανές προκαταλήψεις (biases) ή σφάλματα.
- **Εξασφαλίσουν** τη συμμόρφωση με νομικούς και ηθικούς κανονισμούς.
- **Ανακαλύψουν** νέα γνώση από τις πληροφορίες που εξάγουν τα μοντέλα.

Εκ Φύσεως Ερμηνεύσιμα Μοντέλα (Inherently Interpretable Models) vs. Μοντέλα Μαύρου Κουτιού:

Δεν είναι όλα τα μοντέλα μηχανικής μάθησης εξίσου αδιαφανή. Υπάρχουν μοντέλα

που είναι **εκ φύσεως ερμηνεύσιμα**, όπως:

- **Γραμμική Παλινδρόμηση (Linear Regression):** Οι συντελεστές του μοντέλου υποδεικνύουν άμεσα τη σχέση μεταξύ κάθε χαρακτηριστικού εισόδου και της εξόδου.
- **Λογιστική Παλινδρόμηση (Logistic Regression):** Παρόμοια με τη γραμμική παλινδρόμηση, αλλά για προβλήματα ταξινόμησης, με τους συντελεστές να εκφράζουν την επίδραση των χαρακτηριστικών στον λογάριθμο των πιθανοτήτων (log-odds).
- **Δέντρα Απόφασης (Decision Trees):** Η δομή του δέντρου αναπαριστά μια σειρά από κανόνες if-then-else που είναι εύκολα κατανοητοί.
- **Μοντέλα Βασισμένα σε Κανόνες (Rule-based models):** Όπως τα προηγούμενα, αποτελούνται από ρητούς κανόνες.

Αυτά τα μοντέλα προσφέρουν υψηλό βαθμό διαφάνειας, καθώς η διαδικασία λήψης αποφάσεων είναι άμεσα εμφανής από την ίδια τη δομή τους. Ωστόσο, συχνά υπάρχει ένας **ανταγωνισμός μεταξύ ερμηνευσιμότητας και απόδοσης (accuracy-interpretability trade-off)**. Τα εκ φύσεως ερμηνεύσιμα μοντέλα μπορεί να μην είναι σε θέση να συλλάβουν τις πολύπλοκες, μη γραμμικές σχέσεις που υπάρχουν σε πολλά πραγματικά σύνολα δεδομένων, όπως οι χρηματοοικονομικές χρονοσειρές. Αντίθετα, πιο πολύπλοκα μοντέλα όπως τα βαθιά νευρωνικά δίκτυα (συμπεριλαμβανομένων των LSTMs), τα Support Vector Machines (SVMs) με μη γραμμικούς πυρήνες, ή τα ensemble methods (όπως Random Forests και Gradient Boosting) τείνουν να επιτυγχάνουν υψηλότερη προβλεπτική ακρίβεια, αλλά με το κόστος της ερμηνευσιμότητας.

Το Πρόβλημα και ο Ρόλος της XAI:

Το βασικό πρόβλημα που η XAI καλείται να αντιμετωπίσει είναι πώς να αποκτήσουμε γνώση για τη συμπεριφορά αυτών των πολύπλοκων «μαύρων κουτιών» χωρίς να θυσιάσουμε την προβλεπτική τους ισχύ. Η XAI δεν αποσκοπεί απαραίτητα στο να κάνει κάθε πτυχή ενός πολύπλοκου μοντέλου πλήρως κατανοητή, αλλά μάλλον να παρέχει χρήσιμες και επαρκείς εξηγήσεις που να ανταποκρίνονται στις ανάγκες του χρήστη και στο πλαίσιο της εφαρμογής.

Στην παρούσα διπλωματική, η εφαρμογή μεθόδων XAI στο μοντέλο LSTM για την πρόβλεψη δράσεων σε μετοχές είναι κρίσιμη. Ένας trader ή ένας επενδυτικός αναλυτής θα ήταν πολύ πιο διατεθειμένος να εμπιστευτεί και να ενεργήσει βάσει μιας πρόβλεψης (π.χ., «αγόρασε τη μετοχή $AAPL») εάν κατανοεί ποιοι παράγοντες (π.χ., πρόσφατη αύξηση τιμής, συγκεκριμένος τεχνικός δείκτης, υψηλός όγκος

συναλλαγών) οδήγησαν το μοντέλο σε αυτή την απόφαση, παρά εάν η πρόβλεψη παρέχεται χωρίς καμία αιτιολόγηση. Η εξηγησιμότητα μπορεί επίσης να βοηθήσει στον εντοπισμό περιπτώσεων όπου το μοντέλο βασίζεται σε λανθασμένες συσχετίσεις (spurious correlations) ή όταν η απόδοσή του αρχίζει να φθίνει λόγω αλλαγών στις συνθήκες της αγοράς (concept drift).

4.2 Κατηγορίες Μεθόδων Εξηγησιμότητας

Οι μέθοδοι Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI) μπορούν να ταξινομηθούν με βάση διάφορα κριτήρια. Η κατανόηση αυτών των κατηγοριοποιήσεων βοηθά στην επιλογή της κατάλληλης μεθόδου για ένα συγκεκριμένο πρόβλημα, μοντέλο και τύπο εξήγησης που απαιτείται. Οι πιο συνηθισμένες διακρίσεις είναι:

1. **Εγγενής (Intrinsic) vs. Εκ των Υστέρων (Post-hoc) Εξηγησιμότητα:**
    - **Εγγενής Εξηγησιμότητα:** Αναφέρεται σε μοντέλα που είναι σχεδιασμένα να είναι ερμηνεύσιμα από τη φύση τους, όπως τα γραμμικά μοντέλα, τα δέντρα απόφασης και τα συστήματα βασισμένα σε κανόνες. Η ερμηνευσιμότητα είναι ενσωματωμένη στην ίδια τη δομή του μοντέλου.
    - **Εκ των Υστέρων Εξηγησιμότητα (Post-hoc Explainability):** Αναφέρεται σε μεθόδους που εφαρμόζονται *μετά* την εκπαίδευση ενός μοντέλου (συνήθως ενός μοντέλου «μαύρου κουτιού») για να παράσχουν εξηγήσεις για τις αποφάσεις του. Οι περισσότερες προηγμένες τεχνικές XAI, συμπεριλαμβανομένων αυτών που θα αναλυθούν σε αυτή την εργασία (SHAP, LIME, DeepLIFT, Integrated Gradients, Permutation Feature Importance), ανήκουν σε αυτή την κατηγορία.
2. **Ειδική για το Μοντέλο (Model-Specific) vs. Ανεξάρτητη από το Μοντέλο (Model-Agnostic):**
    - **Ειδική για το Μοντέλο (Model-Specific):** Αυτές οι μέθοδοι είναι σχεδιασμένες για μια συγκεκριμένη κλάση μοντέλων (π.χ., μόνο για νευρωνικά δίκτυα ή μόνο για δενδρικές δομές). Αξιοποιούν την εσωτερική δομή και τις ιδιότητες του μοντέλου για να παράγουν εξηγήσεις. Το DeepLIFT και το Integrated Gradients είναι παραδείγματα μεθόδων που είναι κυρίως ειδικές για βαθιά νευρωνικά δίκτυα.
    - **Ανεξάρτητη από το Μοντέλο (Model-Agnostic):** Αυτές οι μέθοδοι μπορούν να εφαρμοστούν σε οποιοδήποτε μοντέλο μηχανικής

μάθησης, ανεξάρτητα από την εσωτερική του αρχιτεκτονική, καθώς αντιμετωπίζουν το μοντέλο ως «μαύρο κουτί». Λειτουργούν παρατηρώντας τις σχέσεις εισόδου-εξόδου. Το LIME και το Permutation Feature Importance είναι κλασικά παραδείγματα model-agnostic μεθόδων. Το SHAP, αν και βασίζεται σε θεωρητικές αρχές, έχει υλοποιήσεις που είναι model-agnostic (π.χ., KernelSHAP) και άλλες που είναι model-specific (π.χ., DeepSHAP για νευρωνικά δίκτυα).

3. **Τοπική (Local) vs. Καθολική (Global) Εξηγησιμότητα:**
    - **Τοπική Εξηγησιμότητα (Local Explainability):** Παρέχει εξηγήσεις για μια μεμονωμένη πρόβλεψη ενός μοντέλου. Απαντά στην ερώτηση: «Γιατί το μοντέλο έκανε αυτή τη συγκεκριμένη πρόβλεψη για αυτή τη συγκεκριμένη είσοδο;». Το LIME και οι μεμονωμένες τιμές SHAP παρέχουν τοπικές εξηγήσεις. Για παράδειγμα, γιατί το μοντέλο LSTM πρότεινε «αγορά» για τη μετοχή $AAPL την συγκεκριμένη ημέρα, με βάση τα δεδομένα των τελευταίων N ημερών.
    - **Καθολική Εξηγησιμότητα (Global Explainability):** Προσπαθεί να εξηγήσει τη συνολική συμπεριφορά ενός μοντέλου σε ολόκληρο το σύνολο δεδομένων ή σε ένα μεγάλο υποσύνολό του. Απαντά στην ερώτηση: «Πώς λειτουργεί γενικά το μοντέλο; Ποια χαρακτηριστικά είναι τα πιο σημαντικά συνολικά;». Το Permutation Feature Importance παρέχει μια καθολική εικόνα της σημασίας των χαρακτηριστικών. Οι τιμές SHAP μπορούν επίσης να συγκεντρωθούν για να παρέχουν καθολικές εξηγήσεις (π.χ., μέση απόλυτη τιμή SHAP ανά χαρακτηριστικό).
4. **Εξήγηση μέσω Παραδειγμάτων (Example-based Explanations):**
    - Ορισμένες μέθοδοι εξηγούν τη συμπεριφορά ενός μοντέλου επιλέγοντας συγκεκριμένα παραδείγματα από τα δεδομένα εκπαίδευσης ή επικύρωσης που είναι αντιπροσωπευτικά ή κρίσιμα για την κατανόηση των αποφάσεων του μοντέλου. Αυτό μπορεί να περιλαμβάνει την εύρεση πρωτοτύπων (prototypes) ή επικριτικών (criticisms).

Η επιλογή της κατάλληλης μεθόδου ή συνδυασμού μεθόδων XAI εξαρτάται από:

- **Το κοινό της εξήγησης:** Ένας μηχανικός μηχανικής μάθησης μπορεί να χρειάζεται διαφορετικό επίπεδο λεπτομέρειας από έναν τελικό χρήστη ή έναν ρυθμιστικό φορέα.

- **Τον τύπο του μοντέλου:** Δεν είναι όλες οι μέθοδοι εφαρμόσιμες ή αποδοτικές για κάθε μοντέλο.
- **Τις απαιτήσεις της εφαρμογής:** Η ανάγκη για εξηγήσεις σε πραγματικό χρόνο, η υπολογιστική πολυπλοκότητα και η επιθυμητή ακρίβεια της εξήγησης παίζουν ρόλο.
- **Τη φύση των δεδομένων:** Η ερμηνεία των χαρακτηριστικών σε δεδομένα εικόνας διαφέρει από την ερμηνεία σε χρονοσειρές ή πινακοποιημένα δεδομένα.

Στην παρούσα εργασία, θα εστιάσουμε σε πέντε διακεκριμένες post-hoc μεθόδους που καλύπτουν ένα εύρος από τις παραπάνω κατηγορίες και είναι ιδιαίτερα σχετικές για την εξήγηση μοντέλων βαθιάς μάθησης όπως τα LSTMs σε δεδομένα χρονοσειρών: SHAP, LIME, DeepLIFT, Integrated Gradients και Permutation Feature Importance.

Method	Intrinsic vs Post-hoc	Model Specific vs Agnostic	Local vs Global
SHAP	Post-hoc	Both (KernelSHAP: Agnostic, DeepSHAP: Specific)	Both (Local + Global via aggregation)
LIME	Post-hoc	Model-Agnostic	Local
Permutation Feature Importance	Post-hoc	Model-Agnostic	Global
DeepLIFT	Post-hoc	Model-Specific (Neural Networks)	Local
Integrated Gradients	Post-hoc	Model-Specific (Differentiable Models)	Local (can be aggregated)

Πίνακας 4.1: Κατηγοριοποίηση μεθόδων εξηγησιμότητας

4.3 SHAP (SHapley Additive exPlanations)

Το SHAP (SHapley Additive exPlanations) είναι ένα πλαίσιο εξηγησιμότητας που βασίζεται στη θεωρία παιγνίων και συγκεκριμένα στις τιμές Shapley, για να εξηγήσει την πρόβλεψη ενός μοντέλου μηχανικής μάθησης. Έχει κερδίσει σημαντική δημοτικότητα λόγω της ισχυρής θεωρητικής του θεμελίωσης και της ικανότητάς του να παρέχει τόσο τοπικές όσο και καθολικές εξηγήσεις.

4.3.1 Θεωρητική Θεμελίωση

Η κεντρική ιδέα πίσω από το SHAP προέρχεται από τη συνεργατική θεωρία παιγνίων (cooperative game theory). Οι τιμές Shapley, που εισήχθησαν από τον Lloyd Shapley το 1953, παρέχουν έναν τρόπο για τη δίκαιη κατανομή της συνολικής «αμοιβής» (payout) ενός παιχνιδιού μεταξύ των παικτών που συνεργάστηκαν για την επίτευξη αυτής της αμοιβής. Η συνεισφορά κάθε παίκτη υπολογίζεται λαμβάνοντας υπόψη όλες τις πιθανές συμμαχίες (coalitions) παικτών και την οριακή συνεισφορά του παίκτη σε κάθε μία από αυτές τις συμμαχίες.

Στο πλαίσιο της μηχανικής μάθησης, οι «παίκτες» είναι τα χαρακτηριστικά (features) του μοντέλου, και η «αμοιβή» είναι η πρόβλεψη του μοντέλου (ή η διαφορά της πρόβλεψης από μια μέση πρόβλεψη ή πρόβλεψη βάσης). Το SHAP εξηγεί την πρόβλεψη ενός συγκεκριμένου στιγμιότυπου δεδομένων αποδίδοντας μια τιμή σημασίας (την τιμή SHAP) σε κάθε χαρακτηριστικό. Αυτή η τιμή αντιπροσωπεύει τη μέση οριακή συνεισφορά του χαρακτηριστικού στην πρόβλεψη, υπολογισμένη πάνω σε όλες τις πιθανές διατάξεις (ή υποσύνολα) των χαρακτηριστικών.

Μια βασική ιδιότητα των τιμών Shapley, και κατ' επέκταση των τιμών SHAP, είναι ότι είναι η μοναδική μέθοδος απόδοσης που ικανοποιεί τρία επιθυμητά αξιώματα:

1. **Τοπική Ακρίβεια (Local Accuracy) / Προσθετικότητα (Additivity):** Η εξήγηση του μοντέλου (άθροισμα των τιμών SHAP των χαρακτηριστικών) για μια συγκεκριμένη πρόβλεψη x ισούται με την πραγματική πρόβλεψη f(x) του μοντέλου, μείον τη μέση πρόβλεψη του μοντέλου E[f(x)]. Δηλαδή, $\sum_{i=0}^{M} \phi_i = f(x) - E[f(X)]$, όπου M είναι ο αριθμός των χαρακτηριστικών και $\phi_0 = E[f(X)]$ είναι η τιμή βάσης (η μέση πρόβλεψη).
2. **Απουσία (Missingness):** Εάν ένα χαρακτηριστικό δεν έχει τιμή (λείπει) σε μια συμμαχία, η συνεισφορά του πρέπει να είναι μηδέν ($\phi_i = 0$). Πιο γενικά, ένα

χαρακτηριστικό που δεν έχει καμία επίδραση στην πρόβλεψη του μοντέλου θα πρέπει να έχει μηδενική τιμή SHAP.

3. **Συνέπεια (Consistency) / Μονοτονία (Monotonicity):** Εάν ένα μοντέλο αλλάξει έτσι ώστε η οριακή συνεισφορά ενός χαρακτηριστικού να αυξάνεται ή να παραμένει η ίδια (ανεξάρτητα από τα άλλα χαρακτηριστικά), τότε η τιμή SHAP αυτού του χαρακτηριστικού πρέπει επίσης να αυξάνεται ή να παραμένει η ίδια.

Αυτά τα αξιώματα καθιστούν τις τιμές SHAP μια δίκαιη και αξιόπιστη μέθοδο για την απόδοση της σημασίας των χαρακτηριστικών.

4.3.2 Μαθηματικό Υπόβαθρο

Για ένα σύνολο χαρακτηριστικών F, η τιμή Shapley ϕ_i για το χαρακτηριστικό i ορίζεται ως ο σταθμισμένος μέσος όρος της οριακής συνεισφοράς του χαρακτηριστικού i σε όλες τις πιθανές υποομάδες (συμμαχίες) S χαρακτηριστικών που δεν περιλαμβάνουν το i:

$$\phi_i(f,x) = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f_x(S \cup \{i\}) - f_x(S)]$$

όπου:

- f είναι το μοντέλο μηχανικής μάθησης.
- x είναι το στιγμιότυπο για το οποίο θέλουμε να εξηγήσουμε την πρόβλεψη.
- F είναι το σύνολο όλων των χαρακτηριστικών.
- S είναι ένα υποσύνολο των χαρακτηριστικών που δεν περιλαμβάνει το χαρακτηριστικό i.
- $|S|$ είναι ο αριθμός των χαρακτηριστικών στο υποσύνολο S.
- $|F|$ είναι ο συνολικός αριθμός των χαρακτηριστικών.
- $f_x(S \cup \{i\})$ είναι η πρόβλεψη του μοντέλου όταν τα χαρακτηριστικά στο $S \cup \{i\}$ είναι παρόντα (έχουν τις τιμές τους από το x) και τα υπόλοιπα χαρακτηριστικά είναι «απόντα» (συνήθως αντικαθίστανται από τιμές βάσης ή περιθωριοποιούνται).
- $f_x(S)$ είναι η πρόβλεψη του μοντέλου όταν μόνο τα χαρακτηριστικά στο S είναι παρόντα.
- Ο όρος $\frac{|S|!(|F|-|S|-1)!}{|F|!}$ είναι ο συντελεστής στάθμισης που εξασφαλίζει ότι όλες οι διατάξεις των χαρακτηριστικών λαμβάνονται υπόψη δίκαια.

Η $f_x(S)$ μπορεί να ερμηνευτεί ως η αναμενόμενη τιμή της πρόβλεψης του μοντέλου, υπό την προϋπόθεση ότι μόνο τα χαρακτηριστικά στο υποσύνολο S έχουν τις τιμές τους από το στιγμιότυπο x, ενώ τα υπόλοιπα χαρακτηριστικά $F \setminus S$ λαμβάνουν τιμές από κάποια κατανομή αναφοράς (background distribution) ή αντικαθίστανται από τιμές βάσης. Ο υπολογισμός της $f_x(S)$ είναι ένα από τα πιο δύσκολα και υπολογιστικά απαιτητικά μέρη του SHAP, και διαφορετικές προσεγγίσεις οδηγούν σε διαφορετικές παραλλαγές του SHAP:

- **KernelSHAP:** Μια model-agnostic προσέγγιση που χρησιμοποιεί μια ειδικά σταθμισμένη γραμμική παλινδρόμηση για την εκτίμηση των τιμών SHAP. Προσεγγίζει την $f_x(S)$ δειγματοληπτώντας από το σύνολο δεδομένων αναφοράς για τα χαρακτηριστικά που δεν είναι στο S.
- **DeepSHAP (ή DeepLIFT SHAP):** Μια model-specific προσέγγιση για μοντέλα βαθιάς μάθησης, που βασίζεται στη μέθοδο DeepLIFT (η οποία θα συζητηθεί αργότερα). Είναι πιο γρήγορη από το KernelSHAP για νευρωνικά δίκτυα.
- **TreeSHAP:** Μια model-specific προσέγγιση για δενδρικά μοντέλα (όπως δέντρα απόφασης, Random Forests, XGBoost) που μπορεί να υπολογίσει τις ακριβείς τιμές SHAP πολύ πιο γρήγορα από τις model-agnostic προσεγγίσεις.
- **LinearSHAP:** Μια model-specific προσέγγιση για γραμμικά μοντέλα, όπου οι τιμές SHAP μπορούν να υπολογιστούν απευθείας από τους συντελεστές του μοντέλου.

Για την πρόβλεψη ενός συγκεκριμένου στιγμιότυπου x, το μοντέλο εξήγησης του SHAP είναι ένα γραμμικό μοντέλο της μορφής: $g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i$ όπου g είναι το μοντέλο εξήγησης, z' είναι ένας δυαδικός φορέας που υποδεικνύει την παρουσία ($z'_i = 1$) ή απουσία ($z'_i = 0$) ενός χαρακτηριστικού, M είναι ο αριθμός των χαρακτηριστικών, $\phi_0 = E[f(X)]$ είναι η μέση πρόβλεψη (τιμή βάσης), και ϕ_i είναι η τιμή SHAP για το χαρακτηριστικό i. Αυτή η προσθετική φύση της απόδοσης χαρακτηριστικών είναι ένα από τα κύρια πλεονεκτήματα του SHAP.

4.3.3 Πλεονεκτήματα του SHAP

1. **Ισχυρή Θεωρητική Θεμελίωση:** Βασίζεται στις τιμές Shapley από τη συνεργατική θεωρία παιγνίων, οι οποίες έχουν επιθυμητές ιδιότητες (τοπική ακρίβεια, απουσία, συνέπεια). Αυτό παρέχει μια σταθερή βάση για την κατανομή της επίδρασης των χαρακτηριστικών.
2. **Καθολική και Τοπική Ερμηνευσιμότητα:** Το SHAP παρέχει λεπτομερείς

τοπικές εξηγήσεις (τιμές SHAP για κάθε χαρακτηριστικό για ένα μεμονωμένο στιγμιότυπο). Αυτές οι τοπικές εξηγήσεις μπορούν να συγκεντρωθούν για να δημιουργήσουν καθολικές ερμηνείες, όπως η συνολική σημασία των χαρακτηριστικών (π.χ., σχεδιάζοντας τις μέσες απόλυτες τιμές SHAP για κάθε χαρακτηριστικό).

3. **Συνέπεια και Ακρίβεια:** Τα αξιώματα που ικανοποιεί το SHAP εξασφαλίζουν ότι οι εξηγήσεις είναι συνεπείς και ότι η συνεισφορά κάθε χαρακτηριστικού αντικατοπτρίζει με ακρίβεια την επίδρασή του στην πρόβλεψη.
4. **Ανίχνευση Θετικών και Αρνητικών Συνεισφορών:** Οι τιμές SHAP δείχνουν όχι μόνο *πόσο* σημαντικό είναι ένα χαρακτηριστικό, αλλά και την *κατεύθυνση* της επίδρασής του (αν αυξάνει ή μειώνει την πρόβλεψη).
5. **Διαθεσιμότητα Υλοποιήσεων:** Υπάρχουν καλά ανεπτυγμένες βιβλιοθήκες (π.χ., το πακέτο shap στην Python) που υλοποιούν διάφορες παραλλαγές του SHAP, καθιστώντας το σχετικά εύκολο στην εφαρμογή.

4.3.4 Περιορισμοί του SHAP

1. **Υπολογιστικό Κόστος:** Ο ακριβής υπολογισμός των τιμών Shapley είναι NP-hard, καθώς απαιτεί την αξιολόγηση του μοντέλου σε 2^M υποσύνολα χαρακτηριστικών (όπου M είναι ο αριθμός των χαρακτηριστικών). Οι προσεγγιστικές μέθοδοι όπως το KernelSHAP μπορεί να είναι αργές, ειδικά για μεγάλα σύνολα δεδομένων ή πολλά χαρακτηριστικά. Αν και οι model-specific εκδόσεις όπως το DeepSHAP και το TreeSHAP είναι πιο γρήγορες, το υπολογιστικό κόστος μπορεί ακόμα να είναι ένα ζήτημα.
2. **Ερμηνεία της «Απουσίας» Χαρακτηριστικού:** Η έννοια της «απουσίας» ενός χαρακτηριστικού και ο τρόπος με τον οποίο μοντελοποιείται (π.χ., αντικατάσταση με μέσες τιμές, δειγματοληψία από κατανομή αναφοράς) μπορεί να επηρεάσει τις τιμές SHAP. Η επιλογή του κατάλληλου συνόλου δεδομένων αναφοράς (background dataset) είναι κρίσιμη, ειδικά για το KernelSHAP και το DeepSHAP, και μπορεί να μην είναι πάντα προφανής.
3. **Παραδοχή Ανεξαρτησίας Χαρακτηριστικών (σε ορισμένες περιπτώσεις):** Ορισμένες προσεγγίσεις για την εκτίμηση της $f_x(S)$ (π.χ., η απλοϊκή δειγματοληψία από τις περιθώριες κατανομές των χαρακτηριστικών στο σύνολο αναφοράς) μπορεί να υποθέτουν έμμεσα την ανεξαρτησία των χαρακτηριστικών. Όταν τα χαρακτηριστικά είναι έντονα συσχετισμένα, αυτό μπορεί να οδηγήσει σε μη ρεαλιστικά στιγμιότυπα δεδομένων κατά τη διαδικασία εκτίμησης, και οι τιμές SHAP μπορεί να αποδώσουν υπερβολική σημασία σε ένα από τα συσχετισμένα χαρακτηριστικά ή να κατανείμουν τη σημασία με μη διαισθητικό τρόπο. Πιο προηγμένες τεχνικές δειγματοληψίας ή

η χρήση του TreeSHAP (που χειρίζεται τις συσχετίσεις πιο φυσικά) μπορούν να μετριάσουν αυτό το πρόβλημα.

4. **Πιθανότητα Παρανόησης:** Παρόλο που το SHAP παρέχει λεπτομερείς εξηγήσεις, η ερμηνεία των αποτελεσμάτων, ειδικά για χρήστες χωρίς τεχνικό υπόβαθρο, μπορεί να εξακολουθεί να είναι πρόκληση. Η κατανόηση της έννοιας της «οριακής συνεισφοράς» και της φύσης των τιμών βάσης είναι σημαντική.
5. **Ευαισθησία στην Επιλογή του Συνόλου Αναφοράς:** Όπως αναφέρθηκε, οι τιμές SHAP είναι σχετικές με ένα σύνολο δεδομένων αναφοράς. Η αλλαγή αυτού του συνόλου μπορεί να αλλάξει τις τιμές SHAP, καθιστώντας την επιλογή του ένα σημαντικό βήμα που απαιτεί προσοχή.

Παρά αυτούς τους περιορισμούς, το SHAP παραμένει ένα από τα πιο ισχυρά και ευρέως χρησιμοποιούμενα πλαίσια για την ερμηνεία μοντέλων μηχανικής μάθησης, και η εφαρμογή του στο LSTM μοντέλο σου για την πρόβλεψη μετοχών μπορεί να προσφέρει πολύτιμες γνώσεις για τους παράγοντες που οδηγούν τις αποφάσεις του.

4.4 LIME (Local Interpretable Model-agnostic Explanations)

Το LIME (Local Interpretable Model-agnostic Explanations) είναι μια τεχνική που στοχεύει στην εξήγηση των προβλέψεων οποιουδήποτε μοντέλου ταξινόμησης ή παλινδρόμησης με έναν ερμηνεύσιμο και πιστό τρόπο, προσεγγίζοντας τοπικά το μοντέλο με ένα απλούστερο, ερμηνεύσιμο μοντέλο.

4.4.1 Θεωρητική Θεμελίωση

Η κεντρική ιδέα του LIME είναι ότι, ενώ ένα πολύπλοκο μοντέλο («μαύρο κουτί») μπορεί να έχει ένα εξαιρετικά μη γραμμικό και δυσνόητο όριο απόφασης σε παγκόσμιο επίπεδο, η συμπεριφορά του γύρω από ένα συγκεκριμένο στιγμιότυπο δεδομένων (δηλαδή, τοπικά) μπορεί να προσεγγιστεί ικανοποιητικά από ένα πολύ απλούστερο και ερμηνεύσιμο μοντέλο (π.χ., ένα γραμμικό μοντέλο ή ένα μικρό δέντρο απόφασης).

Το LIME λειτουργεί με την παραδοχή ότι οι τοπικές εξηγήσεις είναι πιο εύκολα κατανοητές από τους ανθρώπους. Αντί να προσπαθεί να εξηγήσει ολόκληρη τη λειτουργία του πολύπλοκου μοντέλου, το LIME εστιάζει στο να απαντήσει στην ερώτηση: «Γιατί το μοντέλο έκανε αυτή τη συγκεκριμένη πρόβλεψη για αυτό το

συγκεκριμένο στιγμιότυπο;»

Για να το πετύχει αυτό, το LIME ακολουθεί μια διαδικασία που περιλαμβάνει τα εξής βήματα:

1. **Δημιουργία Γειτονιάς (Neighborhood Generation):** Για το στιγμιότυπο του οποίου την πρόβλεψη θέλουμε να εξηγήσουμε, το LIME δημιουργεί ένα σύνολο νέων, «διαταραγμένων» (perturbed) δειγμάτων στην περιοχή γύρω από αυτό το στιγμιότυπο. Για πινακοποιημένα δεδομένα, αυτό μπορεί να σημαίνει τυχαία αλλαγή των τιμών ορισμένων χαρακτηριστικών. Για δεδομένα κειμένου, μπορεί να σημαίνει αφαίρεση ή προσθήκη λέξεων. Για δεδομένα εικόνας, μπορεί να σημαίνει απόκρυψη τμημάτων της εικόνας.
2. **Λήψη Προβλέψεων από το Μαύρο Κουτί:** Το αρχικό, πολύπλοκο μοντέλο (το «μαύρο κουτί») χρησιμοποιείται για να κάνει προβλέψεις για αυτά τα διαταραγμένα δείγματα.
3. **Στάθμιση των Δειγμάτων:** Τα διαταραγμένα δείγματα σταθμίζονται ανάλογα με την εγγύτητά τους στο αρχικό στιγμιότυπο. Δείγματα που είναι πιο «κοντά» στο αρχικό στιγμιότυπο λαμβάνουν μεγαλύτερο βάρος, καθώς θεωρούνται πιο σχετικά για την κατανόηση της τοπικής συμπεριφοράς του μοντέλου. Η συνάρτηση εγγύτητας (π.χ., Ευκλείδεια απόσταση, συνημιτονοειδής ομοιότητα) και η συνάρτηση πυρήνα (kernel function) που μετατρέπει την απόσταση σε βάρος είναι σημαντικές παράμετροι.
4. **Εκπαίδευση Ερμηνεύσιμου Μοντέλου:** Ένα απλό, ερμηνεύσιμο μοντέλο (π.χ., γραμμική παλινδρόμηση, λογιστική παλινδρόμηση, δέντρο απόφασης) εκπαιδεύεται χρησιμοποιώντας τα σταθμισμένα διαταραγμένα δείγματα και τις αντίστοιχες προβλέψεις τους από το μαύρο κουτί. Ο στόχος είναι αυτό το απλό μοντέλο να προσεγγίσει όσο το δυνατόν καλύτερα τη συμπεριφορά του μαύρου κουτιού στην τοπική γειτονιά του αρχικού στιγμιότυπου.
5. **Παροχή Εξήγησης:** Η εξήγηση για την πρόβλεψη του αρχικού στιγμιότυπου προέρχεται από το εκπαιδευμένο ερμηνεύσιμο μοντέλο. Για παράδειγμα, εάν το ερμηνεύσιμο μοντέλο είναι μια γραμμική παλινδρόμηση, οι συντελεστές του υποδεικνύουν τη σημασία και την κατεύθυνση της επίδρασης κάθε χαρακτηριστικού στην τοπική πρόβλεψη.

4.4.2 Μαθηματικό Υπόβαθρο

Το LIME προσπαθεί να βελτιστοποιήσει την ακόλουθη αντικειμενική συνάρτηση για ένα συγκεκριμένο στιγμιότυπο x:

$\xi(x) = arg_{g \in G} min L(f, g, \pi_x) + \Omega(g)$

όπου:

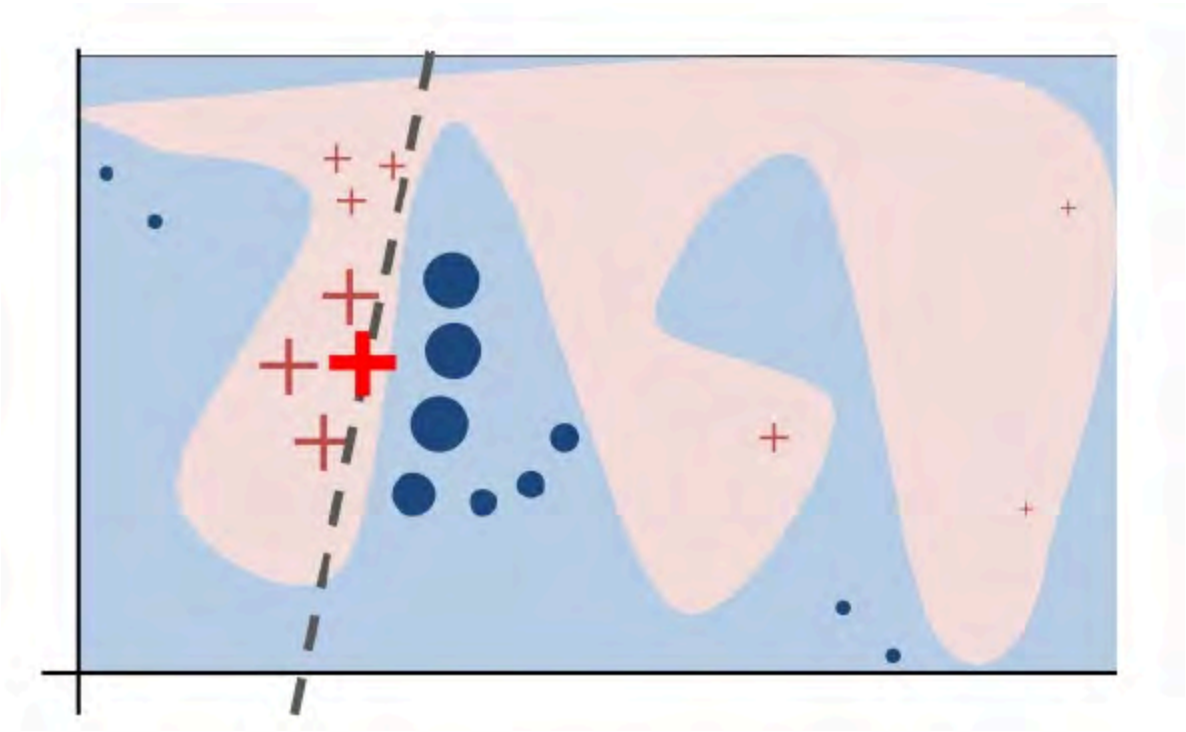
- f είναι το πολύπλοκο μοντέλο («μαύρο κουτί») που θέλουμε να εξηγήσουμε.
- g είναι ένα μοντέλο από την κλάση G των ερμηνεύσιμων μοντέλων (π.χ., γραμμικά μοντέλα).
- πx είναι μια μέτρηση εγγύτητας (proximity measure) που ορίζει την τοπικότητα γύρω από το στιγμιότυπο x.
- $L(f, g, \pi_x)$ είναι ένας όρος απώλειας (fidelity function) που μετρά πόσο καλά το g προσεγγίζει τις προβλέψεις του f στην περιοχή που ορίζεται από την πx. Για παράδειγμα, μπορεί να είναι ένα σταθμισμένο τετραγωνικό σφάλμα: $L(f, g, \pi_x) = \sum_{z, z' \in Z} \pi_x(z)(f(z) - g(z'))^2$ όπου Z είναι το σύνολο των διαταραγμένων δειγμάτων z γύρω από το x, και z' είναι μια απλοποιημένη αναπαράσταση του z που μπορεί να χρησιμοποιηθεί από το ερμηνεύσιμο μοντέλο g.
- $\Omega(g)$ είναι ένας όρος πολυπλοκότητας (complexity term) του ερμηνεύσιμου μοντέλου g. Για παράδειγμα, σε ένα γραμμικό μοντέλο, μπορεί να είναι ο αριθμός των μη μηδενικών συντελεστών (για αραιότητα) ή σε ένα δέντρο απόφασης, το βάθος του δέντρου. Αυτό ενθαρρύνει την επιλογή απλών και κατανοητών εξηγήσεων.

Στην πράξη, για να εξηγήσει ένα στιγμιότυπο x_0:

1. Δημιουργούνται N διαταραγμένα δείγματα x_i στη γειτονιά του x_0.
2. Λαμβάνονται οι προβλέψεις $y_i = f(x_i)$ από το μαύρο κουτί.
3. Υπολογίζονται τα βάρη $w_i = \pi_{x_0}(x_i)$ (π.χ., χρησιμοποιώντας έναν πυρήνα εκθετικής απόστασης).
4. Επιλέγεται ένα υποσύνολο K χαρακτηριστικών (feature selection) για την εξήγηση (π.χ., με Lasso ή επιλέγοντας τα K πιο σημαντικά).
5. Ένα απλό γραμμικό μοντέλο εκπαιδεύεται σταθμισμένα: $arg_g min \sum_i w_i (y_i - g(x_i))^2$. Οι συντελεστές αυτού του γραμμικού μοντέλου αποτελούν την εξήγηση.

Για παράδειγμα, αν εξηγούμε μια πρόβλεψη για μια μετοχή, το LIME θα

δημιουργούσε παραλλαγές των πρόσφατων δεδομένων της μετοχής (π.χ., ελαφρώς διαφορετικές τιμές, όγκοι), θα έπαιρνε τις προβλέψεις του LSTM για αυτές τις παραλλαγές, και μετά θα εκπαίδευε ένα απλό γραμμικό μοντέλο για να δει ποια από τα αρχικά χαρακτηριστικά (π.χ., η τιμή κλεισίματος της προηγούμενης ημέρας, η μεταβολή του όγκου) είχαν τη μεγαλύτερη τοπική επίδραση στην απόφαση του LSTM.



Εικόνα 4.1: Τοπική Ερμηνεία με LIME.

Αυτή η εικόνα απεικονίζει μια απλή αναπαράσταση της μεθόδου LIME για εξήγηση προβλέψεων μοντέλων. Το πολύπλοκο μοντέλο (μαύρο κουτί) έχει έναν δύσκολο διαχωρισμό μεταξύ δύο κατηγοριών (ροζ και μπλε περιοχές). Το κόκκινο Χ είναι το σημείο που θέλουμε να εξηγήσουμε. Η LIME δημιουργεί σημεία γύρω από αυτό (τα σημεία μικρότερα και μεγαλύτερα σε μέγεθος δείχνουν την εγγύτητα), και εκπαιδεύει ένα τοπικό γραμμικό μοντέλο (η διακεκομμένη γραμμή), που εξηγεί το μοντέλο μόνο τοπικά (και όχι σε όλο τον χώρο εισόδου).

4.4.3 Πλεονεκτήματα του LIME

1. **Ανεξάρτητο από το Μοντέλο (Model-Agnostic):** Το LIME μπορεί να χρησιμοποιηθεί για να εξηγήσει οποιοδήποτε μοντέλο μηχανικής μάθησης, καθώς αντιμετωπίζει το μοντέλο ως μαύρο κουτί. Δεν απαιτεί καμία γνώση της εσωτερικής δομής του.

2. **Ευκολία στην Κατανόηση των Εξηγήσεων (Interpretability):** Οι εξηγήσεις παρέχονται με τη μορφή ενός απλού, ερμηνεύσιμου μοντέλου (π.χ., λίστα σημαντικών χαρακτηριστικών και των συντελεστών τους), το οποίο είναι συνήθως εύκολο να κατανοηθεί ακόμη και από μη ειδικούς.
3. **Ευελιξία στον Τύπο Δεδομένων:** Το LIME έχει εφαρμοστεί με επιτυχία σε διάφορους τύπους δεδομένων, συμπεριλαμβανομένων πινακοποιημένων δεδομένων, κειμένου και εικόνων, προσαρμόζοντας κατάλληλα τη μέθοδο διαταραχής και τη συνάρτηση εγγύτητας.
4. **Τοπική Πιστότητα (Local Fidelity):** Επικεντρώνεται στην ακριβή προσέγγιση της συμπεριφοράς του μοντέλου σε τοπικό επίπεδο, κάτι που είναι συχνά αυτό που ενδιαφέρει για την κατανόηση μιας μεμονωμένης πρόβλεψης.

4.4.4 Περιορισμοί του LIME

1. **Ορισμός της Γειτονιάς (Definition of Neighborhood):** Η απόδοση του LIME μπορεί να είναι ευαίσθητη στον τρόπο με τον οποίο ορίζεται η «γειτονιά» γύρω από ένα στιγμιότυπο (δηλαδή, η μέθοδος διαταραχής και το εύρος του πυρήνα στάθμισης). Δεν υπάρχει πάντα ένας ξεκάθαρος ή βέλτιστος τρόπος για να γίνει αυτό.
2. **Αστάθεια των Εξηγήσεων (Instability of Explanations):** Μικρές αλλαγές στη δειγματοληψία των διαταραγμένων στιγμιότυπων ή στις παραμέτρους του LIME (π.χ., εύρος πυρήνα) μπορούν μερικές φορές να οδηγήσουν σε σημαντικά διαφορετικές εξηγήσεις για το ίδιο στιγμιότυπο. Αυτό μπορεί να μειώσει την εμπιστοσύνη στις εξηγήσεις.
3. **Ποιότητα της Προσέγγισης:** Η παραδοχή ότι η τοπική συμπεριφορά ενός πολύπλοκου μοντέλου μπορεί πάντα να προσεγγιστεί καλά από ένα απλό γραμμικό μοντέλο (ή άλλο απλό μοντέλο) μπορεί να μην ισχύει πάντα, ειδικά εάν το όριο απόφασης του μαύρου κουτιού είναι εξαιρετικά μη γραμμικό ακόμη και σε μικρή κλίμακα.
4. **Επιλογή του Ερμηνεύσιμου Μοντέλου:** Η επιλογή του τύπου του ερμηνεύσιμου μοντέλου (π.χ., γραμμικό μοντέλο, δέντρο απόφασης) και των παραμέτρων του (π.χ., αριθμός χαρακτηριστικών στην εξήγηση) μπορεί να επηρεάσει την τελική εξήγηση.
5. **Χειρισμός Συσχετισμένων Χαρακτηριστικών:** Η απλή δειγματοληψία χαρακτηριστικών ανεξάρτητα κατά τη δημιουργία διαταραγμένων δειγμάτων μπορεί να δημιουργήσει μη ρεαλιστικά στιγμιότυπα εάν τα αρχικά χαρακτηριστικά είναι συσχετισμένα. Αυτό μπορεί να οδηγήσει σε εξηγήσεις που δεν αντικατοπτρίζουν τον τρόπο με τον οποίο το μοντέλο θα συμπεριφερόταν σε πραγματικά δεδομένα.

6. **Υπολογιστικό Κόστος για Πολλές Εξηγήσεις:** Αν και η εξήγηση ενός μεμονωμένου στιγμιότυπου είναι συνήθως γρήγορη, η παραγωγή εξηγήσεων LIME για ένα μεγάλο σύνολο δεδομένων μπορεί να είναι υπολογιστικά εντατική, καθώς η διαδικασία πρέπει να επαναληφθεί για κάθε στιγμιότυπο.

Παρά τους περιορισμούς του, το LIME παραμένει μια δημοφιλής και χρήσιμη τεχνική, ειδικά όταν απαιτείται μια model-agnostic, τοπική εξήγηση που είναι εύκολη στην κατανόηση. Για το LSTM μοντέλο σου, το LIME θα μπορούσε να βοηθήσει να εξηγηθεί γιατί μια συγκεκριμένη ακολουθία τιμών και όγκων οδήγησε σε μια συγκεκριμένη σύσταση (buy/sell/hold).

4.5 DeepLIFT (Deep Learning Important FeaTures)

Το DeepLIFT (Deep Learning Important FeaTures) είναι μια μέθοδος εξηγησιμότητας ειδικά σχεδιασμένη για βαθιά νευρωνικά δίκτυα. Στόχος της είναι να αποδώσει βαθμολογίες συνεισφοράς στις εισόδους ενός νευρωνικού δικτύου, εξηγώντας πώς αυτές οι είσοδοι επηρεάζουν την έξοδο του δικτύου, συγκρίνοντας τις ενεργοποιήσεις των νευρώνων με τις ενεργοποιήσεις τους σε μια «κατάσταση αναφοράς».

4.5.1 Θεωρητική Θεμελίωση

Το DeepLIFT σχεδιάστηκε για να αντιμετωπίσει ορισμένους περιορισμούς των απλών μεθόδων που βασίζονται σε κλίσεις (gradients), όπως το πρόβλημα του «κορεσμού» των κλίσεων (gradient saturation). Όταν ένας νευρώνας βρίσκεται σε μια κορεσμένη περιοχή της συνάρτησης ενεργοποίησής του (π.χ., η σιγμοειδής συνάρτηση για πολύ μεγάλες ή πολύ μικρές εισόδους), η κλίση μπορεί να είναι κοντά στο μηδέν, υποδηλώνοντας λανθασμένα ότι η είσοδος δεν είναι σημαντική, ενώ στην πραγματικότητα μπορεί να είναι.

Η κεντρική ιδέα του DeepLIFT είναι να αναλύσει τη διαφορά στην έξοδο του δικτύου από μια «τιμή αναφοράς» (reference output) σε συνεισφορές από τις διαφορές των εισόδων από τις «τιμές αναφοράς εισόδου» (reference inputs). Αντί να χρησιμοποιεί κλίσεις, το DeepLIFT χρησιμοποιεί την έννοια των «πολλαπλασιαστών» (multipliers) για να διαδώσει τις βαθμολογίες συνεισφοράς προς τα πίσω, από την έξοδο προς τις εισόδους.

Συγκεκριμένα, για κάθε νευρώνα t στο δίκτυο, το DeepLIFT υπολογίζει μια βαθμολογία συνεισφοράς $C_{\Delta x_i \Delta t}$ που αντιπροσωπεύει πόσο από τη διαφορά

$\Delta_t = t - t_0$ (όπου t είναι η ενεργοποίηση του νευρώνα για την πραγματική είσοδο και t0 είναι η ενεργοποίηση για την είσοδο αναφοράς) οφείλεται στη διαφορά $\Delta_{x_i} = x_i - x_{i0}$ από κάθε είσοδο x_i του νευρώνα.

Το DeepLIFT ικανοποιεί μια σημαντική ιδιότητα που ονομάζεται **«άθροιση προς τη διαφορά» (summation-to-delta)**: το άθροισμα των συνεισφορών όλων των εισόδων ενός νευρώνα ισούται με τη διαφορά της ενεργοποίησης του νευρώνα από την ενεργοποίησή του αναφοράς. Αυτό επεκτείνεται σε ολόκληρο το δίκτυο, έτσι ώστε το άθροισμα των συνεισφορών όλων των αρχικών χαρακτηριστικών εισόδου να ισούται με τη διαφορά της τελικής εξόδου από την έξοδο αναφοράς:

$$\sum_i C_{\Delta X_i \Delta Y} = DY = Y - Y_0$$

όπου X_i είναι τα αρχικά χαρακτηριστικά εισόδου, Y είναι η έξοδος του δικτύου, και Y_0 είναι η έξοδος του δικτύου για τις εισόδους αναφοράς. Η επιλογή της κατάλληλης εισόδου αναφοράς είναι κρίσιμη και εξαρτάται από το πρόβλημα (π.χ., μια είσοδος με όλες τις τιμές μηδέν, μια μέση είσοδος από τα δεδομένα εκπαίδευσης, ή μια είσοδος που παράγει μια "ουδέτερη" έξοδο).

4.5.2 Μαθηματικό Υπόβαθρο

Το DeepLIFT ορίζει «πολλαπλασιαστές» $m_{\Delta x \Delta t}$ αντί για κλίσεις. Αν t είναι ένας νευρώνας-στόχος και $x1, x2, ..., xn$ είναι οι νευρώνες-πηγές που τροφοδοτούν τον t, τότε η συνεισφορά της διαφοράς Δxi στη διαφορά Δt ορίζεται ως:

$$C_{\Delta x_i \Delta t} = m_{\Delta x_i \Delta t} \cdot \Delta x_i$$

Οι κανόνες για τον υπολογισμό αυτών των πολλαπλασιαστών εξαρτώνται από τον τύπο του νευρώνα και τη συνάρτηση ενεργοποίησης. Το DeepLIFT εισάγει δύο κύριους κανόνες διάδοσης (propagation rules) για την αντιμετώπιση διαφορετικών τύπων μη γραμμικοτήτων:

1. **Κανόνας Γραμμικότητας (Linear Rule):** Εάν η σχέση μεταξύ των Δx_i και του Δt είναι γραμμική (ή μπορεί να προσεγγιστεί ως γραμμική), ο πολλαπλασιαστής είναι απλώς η κλίση. Αυτό ισχύει για γραμμικούς νευρώνες ή για την προσέγγιση μη γραμμικών συναρτήσεων. Για ένα νευρώνα

$t = \sum_i w_i x_i + b$, ο πολλαπλασιαστής $m_{\Delta x_i \Delta t}$ είναι απλά το βάρος w_i.

2. **Κανόνας Αλλαγής Κλίμακας (Rescale Rule):** Για μη γραμμικές συναρτήσεις ενεργοποίησης (π.χ., ReLU, sigmoid, tanh) όπου η κλίση μπορεί να είναι μηδέν ή να μην αντικατοπτρίζει τη σημασία, το DeepLIFT προτείνει τον «Κανόνα Αλλαγής Κλίμακας». Αν $y = f(x)$ όπου f είναι μια μη γραμμική συνάρτηση ενεργοποίησης, και $\Delta y = y - y_0$, $\Delta x = x - x_0$ τότε ο πολλαπλασιαστής για τη συνεισφορά του Δ_x στο Δ_y είναι: $m_{\Delta x \Delta t} = \frac{\Delta x}{\Delta y}$ (εάν $\Delta x = 0$) Εάν $\Delta x = 0$, ο πολλαπλασιαστής ορίζεται ως η κλίση στο x_0 (ή στο x εάν η f είναι γραμμική).

3. **Κανόνας Αποκάλυψης-Ακύρωσης (RevealCancel Rule):** Αυτός ο κανόνας είναι πιο περίπλοκος και σχεδιάστηκε για να χειρίζεται καλύτερα τις μη γραμμικότητες όπου θετικές και αρνητικές συνεισφορές μπορεί να «ακυρώνουν» η μία την άλλη, όπως σε συναρτήσεις ενεργοποίησης τύπου max ή softmax. Χωρίζει τις εισροές σε θετικές και αρνητικές συνιστώσες και τις επεξεργάζεται ξεχωριστά. Είναι ιδιαίτερα χρήσιμος για συναρτήσεις ενεργοποίησης όπως η ReLU και η max-pooling. Για παράδειγμα, για μια ReLU $y = max(0, x)$:

    - Εάν $x > 0$ και $x_0 \leq 0$ (θετική συνεισφορά), $m_{\Delta x \Delta t} = \frac{y - y_0}{x - x_0} = \frac{x - 0}{x - x_0}$.
    - Εάν $x \leq 0$ και $x_0 > 0$ (αρνητική συνεισφορά που αφαιρείται), $m_{\Delta x \Delta t} = \frac{y - y_0}{x - x_0} = \frac{x - 0}{x - x_0}$.
    - Σε άλλες περιπτώσεις, μπορεί να είναι 1 ή 0 ανάλογα με τις τιμές.

Η διαδικασία εφαρμογής του DeepLIFT περιλαμβάνει:

1. Μια εμπρόσθια διάδοση (forward pass) για τον υπολογισμό των ενεργοποιήσεων για την πραγματική είσοδο και την είσοδο αναφοράς.
2. Μια οπίσθια διάδοση (backward pass) για τον υπολογισμό των πολλαπλασιαστών και των βαθμολογιών συνεισφοράς, ξεκινώντας από τον νευρώνα εξόδου και προχωρώντας προς τα πίσω μέχρι τα επίπεδα εισόδου, χρησιμοποιώντας τους κατάλληλους κανόνες (Linear, Rescale, RevealCancel) σε κάθε νευρώνα.

Για μοντέλα LSTM, η εφαρμογή του DeepLIFT απαιτεί την προσεκτική διάδοση των

συνεισφορών μέσω των πυλών και των μη γραμμικών συναρτήσεων ενεργοποίησης (sigmoid, tanh) που χρησιμοποιούνται εντός της κυψέλης LSTM, καθώς και μέσω των χρονικών βημάτων.

4.5.3 Πλεονεκτήματα του DeepLIFT

1. **Αντιμετώπιση του Κορεσμού Κλίσεων:** Σε αντίθεση με τις απλές μεθόδους που βασίζονται σε κλίσεις, το DeepLIFT μπορεί να αποδώσει σημασία σε εισόδους ακόμη και όταν οι νευρώνες είναι κορεσμένοι, καθώς εξετάζει τη διαφορά από μια κατάσταση αναφοράς.
2. **Διάδοση Θετικών και Αρνητικών Συνεισφορών:** Οι κανόνες του DeepLIFT (ειδικά ο RevealCancel) είναι σχεδιασμένοι για να διαχωρίζουν και να διαδίδουν τις θετικές και αρνητικές συνεισφορές, παρέχοντας μια πιο λεπτομερή εικόνα του πώς οι είσοδοι επηρεάζουν την έξοδο.
3. **Ιδιότητα "Summation-to-Delta":** Εξασφαλίζει ότι η συνολική συνεισφορά των εισόδων ισούται με τη διαφορά της εξόδου από την έξοδο αναφοράς, παρέχοντας μια μορφή «διατήρησης» της σημασίας.
4. **Υπολογιστική Αποδοτικότητα:** Μόλις υπολογιστούν οι ενεργοποιήσεις αναφοράς, η οπίσθια διάδοση για τον υπολογισμό των συνεισφορών είναι παρόμοια σε πολυπλοκότητα με μια τυπική οπίσθια διάδοση κλίσεων, καθιστώντας το πιο γρήγορο από μεθόδους που βασίζονται σε διαταραχές για κάθε εξήγηση.
5. **Ειδικό για Βαθιά Δίκτυα:** Αξιοποιεί την αρχιτεκτονική του δικτύου και μπορεί να προσαρμοστεί σε συγκεκριμένους τύπους επιπέδων και συναρτήσεων ενεργοποίησης.

4.5.4 Περιορισμοί του DeepLIFT

1. **Επιλογή της Εισόδου Αναφοράς:** Η επιλογή μιας κατάλληλης εισόδου αναφοράς είναι κρίσιμη και μπορεί να επηρεάσει σημαντικά τις αποδιδόμενες βαθμολογίες συνεισφοράς. Δεν υπάρχει πάντα μια προφανής ή «σωστή» επιλογή, και διαφορετικές επιλογές μπορεί να οδηγήσουν σε διαφορετικές ερμηνείες.
2. **Ειδικό για Μοντέλα (Model-Specific):** Το DeepLIFT είναι σχεδιασμένο αποκλειστικά για βαθιά νευρωνικά δίκτυα και δεν μπορεί να εφαρμοστεί σε άλλους τύπους μοντέλων μηχανικής μάθησης.
3. **Πολυπλοκότητα Υλοποίησης:** Η σωστή υλοποίηση των κανόνων διάδοσης, ειδικά του RevealCancel, για διάφορες αρχιτεκτονικές δικτύων και συναρτήσεις ενεργοποίησης μπορεί να είναι πολύπλοκη και επιρρεπής σε σφάλματα.

4. **Θεωρητικές Παραδοχές:** Οι κανόνες όπως ο "Rescale" κάνουν ορισμένες παραδοχές γραμμικότητας σε τοπικό επίπεδο (μεταξύ της πραγματικής τιμής και της τιμής αναφοράς), οι οποίες μπορεί να μην ισχύουν πάντα τέλεια.
5. **Ερμηνεία των Βαθμολογιών:** Αν και οι βαθμολογίες παρέχουν μια ποσοτικοποίηση της συνεισφοράς, η ακριβής ερμηνεία τους σε σχέση με τις πραγματικές διαδικασίες απόφασης του δικτύου μπορεί να εξακολουθεί να απαιτεί προσεκτική ανάλυση.

Το DeepLIFT, όταν εφαρμόζεται σωστά, μπορεί να προσφέρει πολύτιμες και συχνά πιο διαισθητικές εξηγήσεις για μοντέλα βαθιάς μάθησης σε σύγκριση με απλούστερες μεθόδους κλίσεων, ειδικά σε περιπτώσεις όπου ο κορεσμός είναι πρόβλημα. Για το LSTM σου, θα μπορούσε να αποκαλύψει ποιες συγκεκριμένες χρονικές στιγμές ή χαρακτηριστικά εισόδου (τιμές, όγκοι) συνέβαλαν περισσότερο (θετικά ή αρνητικά) στην τελική απόφαση αγοράς/πώλησης/διακράτησης, σε σχέση με μια κατάσταση «ουδέτερης» αγοράς.

4.6 Integrated Gradients (Ολοκληρωμένες Κλίσεις)

Το Integrated Gradients (IG) είναι μια άλλη τεχνική εξηγησιμότητας που στοχεύει στην απόδοση της πρόβλεψης ενός μοντέλου βαθιάς μάθησης στις εισόδους του. Βασίζεται στην ιδέα της ολοκλήρωσης των κλίσεων κατά μήκος μιας ευθείας διαδρομής από μια είσοδο «αναφοράς» (baseline) προς την πραγματική είσοδο.

4.6.1 Θεωρητική Θεμελίωση

Το Integrated Gradients προτάθηκε για να ικανοποιήσει δύο θεμελιώδη αξιώματα για τις μεθόδους απόδοσης χαρακτηριστικών:

1. **Ευαισθησία (Sensitivity):** Εάν μια είσοδος και μια τιμή αναφοράς διαφέρουν μόνο σε ένα χαρακτηριστικό, αλλά αυτό το χαρακτηριστικό οδηγεί σε διαφορετική πρόβλεψη, τότε αυτό το χαρακτηριστικό πρέπει να έχει μη μηδενική απόδοση. Οι απλές μέθοδοι κλίσεων μπορεί να παραβιάζουν αυτό το αξίωμα λόγω κορεσμού (όπου η κλίση είναι μηδέν παρόλο που η είσοδος είναι σημαντική).
2. **Ανεξαρτησία Υλοποίησης (Implementation Invariance):** Δύο λειτουργικά ισοδύναμα δίκτυα (δηλαδή, παράγουν την ίδια έξοδο για κάθε είσοδο) πρέπει να έχουν τις ίδιες αποδόσεις χαρακτηριστικών. Αυτό σημαίνει ότι η μέθοδος απόδοσης πρέπει να εξαρτάται μόνο από τη συνάρτηση που υπολογίζει το δίκτυο και όχι από τον συγκεκριμένο τρόπο υλοποίησής της.

Το IG επιτυγχάνει αυτές τις ιδιότητες ορίζοντας την απόδοση ενός χαρακτηριστικού x_i ως το ολοκλήρωμα της κλίσης της εξόδου του μοντέλου $F(x)$ ως προς το χαρακτηριστικό x_i, κατά μήκος μιας γραμμικής διαδρομής από μια είσοδο αναφοράς x' προς την πραγματική είσοδο x.

Η κεντρική ιδέα είναι ότι εάν ένα χαρακτηριστικό αλλάζει από την τιμή αναφοράς του στην πραγματική του τιμή, η συνολική επίδρασή του στην έξοδο μπορεί να βρεθεί αθροίζοντας (ολοκληρώνοντας) όλες τις μικρές, τοπικές αλλαγές (κλίσεις) κατά μήκος αυτής της διαδρομής. Αυτό βοηθά στην αντιμετώπιση του προβλήματος όπου η κλίση σε ένα μεμονωμένο σημείο (π.χ., στην πραγματική είσοδο) μπορεί να μην αντικατοπτρίζει τη συνολική σημασία του χαρακτηριστικού λόγω μη γραμμικοτήτων ή κορεσμού.

Μια σημαντική ιδιότητα του IG είναι η **Πληρότητα (Completeness)** ή **Άθροιση προς τη Διαφορά (Summation to Delta)**, παρόμοια με του DeepLIFT: το άθροισμα των αποδόσεων (τιμών IG) όλων των χαρακτηριστικών εισόδου ισούται με τη διαφορά μεταξύ της πρόβλεψης του μοντέλου για την πραγματική είσοδο $F(x)$ και της πρόβλεψής του για την είσοδο αναφοράς $F(x')$:

$$\sum_{i=1}^{n} IntegratedGrads_i(x) = \quad = F(x) - F(x')$$

όπου n είναι ο αριθμός των χαρακτηριστικών εισόδου. Αυτό καθιστά τις αποδόσεις IG εύκολα ερμηνεύσιμες ως μια πλήρης κατανομή της διαφοράς πρόβλεψης στα χαρακτηριστικά.

4.6.2 Μαθηματικό Υπόβαθρο

Η απόδοση Integrated Gradients για το i-οστό χαρακτηριστικό x_iμιας εισόδου x και μιας αναφοράς x' ορίζεται ως:

$$IntegratedGrads_i(x) ::= (x_i - x_i') \times \int_{a=0}^{1} \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} d\alpha$$

όπου:

- x_i είναι η τιμή του i-οστού χαρακτηριστικού της πραγματικής εισόδου.
- x_i' είναι η τιμή του i-οστού χαρακτηριστικού της εισόδου αναφοράς.
- F είναι η συνάρτηση του νευρωνικού δικτύου.

- $\frac{\partial F(x'+\alpha(x-x'))}{\partial x_i}$ είναι η κλίση της εξόδου F ως προς το i-οστό χαρακτηριστικό, υπολογισμένη σε ένα σημείο $x' + \alpha(x - x')$ κατά μήκος της ευθείας γραμμής μεταξύ της αναφοράς x' και της εισόδου x.
- α είναι μια μεταβλητή ολοκλήρωσης που κυμαίνεται από 0 (στην αναφορά) έως 1 (στην πραγματική είσοδο).

Στην πράξη, το ολοκλήρωμα αυτό συνήθως προσεγγίζεται αριθμητικά χρησιμοποιώντας το άθροισμα Riemann:

$$IntegratedGrads_i(x) \approx (x - x') \times \sum_{k=1}^{m} \frac{\partial F(x'+\frac{k}{m}(x-x'))}{\partial x_i} \times \frac{1}{m}$$

όπου m είναι ο αριθμός των βημάτων στην προσέγγιση (π.χ. m από 20 έως 300 είναι συνηθισμένες τιμές). Όσο μεγαλύτερο το m, τόσο καλύτερη η προσέγγιση του ολοκληρώματος, αλλά και τόσο μεγαλύτερο το υπολογιστικό κόστος.

Η επιλογή της **αναφοράς (baseline)** x′ είναι κρίσιμη. Μια κοινή επιλογή είναι μια είσοδος με όλες τις τιμές μηδέν, ειδικά εάν το μηδέν αντιπροσωπεύει την «απουσία» πληροφορίας. Για δεδομένα εικόνας, μια μαύρη εικόνα μπορεί να είναι μια αναφορά. Για δεδομένα χρονοσειρών όπως οι μετοχές, η αναφορά θα μπορούσε να είναι μια ακολουθία μηδενικών ή μια ακολουθία που αντιπροσωπεύει μια «μέση» ή «ουδέτερη» κατάσταση της αγοράς. Η επιλογή της αναφοράς επηρεάζει άμεσα τις τιμές IG, καθώς αυτές εξηγούν τη διαφορά από αυτή την αναφορά.

Για ένα μοντέλο LSTM, το x θα ήταν η ακολουθία εισόδου (π.χ., T χρονικά βήματα με D χαρακτηριστικά ανά βήμα). Η αναφορά x′ θα ήταν μια αντίστοιχη ακολουθία αναφοράς. Οι κλίσεις θα υπολογίζονταν για κάθε στοιχείο της ακολουθίας εισόδου.

4.6.3 Πλεονεκτήματα του Integrated Gradients

1. **Ικανοποίηση Αξιωμάτων (Sensitivity, Implementation Invariance, Completeness):** Αυτά τα αξιώματα παρέχουν μια ισχυρή θεωρητική βάση για την αξιοπιστία και τη συνέπεια των εξηγήσεων. Η ιδιότητα της πληρότητας είναι ιδιαίτερα χρήσιμη, καθώς οι αποδόσεις αθροίζονται στη διαφορά πρόβλεψης.
2. **Σχετικά Απλή Υλοποίηση:** Για διαφορίσιμα μοντέλα (όπως τα νευρωνικά δίκτυα), το IG απαιτεί μόνο τον υπολογισμό κλίσεων, κάτι που είναι μια τυπική λειτουργία στις περισσότερες βιβλιοθήκες βαθιάς μάθησης. Η κύρια

προσθήκη είναι η επανάληψη αυτού του υπολογισμού σε πολλαπλά σημεία κατά μήκος της διαδρομής και ο μέσος όρος.

3. **Αντιμετώπιση του Κορεσμού Κλίσεων:** Με την ολοκλήρωση των κλίσεων, το IG μπορεί να συλλάβει τη σημασία των χαρακτηριστικών ακόμη και αν η κλίση στην τελική είσοδο (ή στην αναφορά) είναι μηδέν.
4. **Ευελιξία στην Επιλογή Αναφοράς:** Αν και η επιλογή είναι κρίσιμη, η ευελιξία επιτρέπει στον χρήστη να προσαρμόσει την εξήγηση σε αυτό που θεωρείται «ουδέτερη» ή «αρχική» κατάσταση.
5. **Παρέχει Κατευθυνόμενες Αποδόσεις:** Όπως και οι κλίσεις, οι τιμές IG έχουν πρόσημο, υποδεικνύοντας αν η αύξηση ενός χαρακτηριστικού αυξάνει ή μειώνει την έξοδο.

4.6.4 Περιορισμοί του Integrated Gradients

1. **Εξάρτηση από την Αναφορά (Baseline Dependence):** Οι αποδόσεις που παράγονται είναι πάντα σχετικές με την επιλεγμένη αναφορά. Διαφορετικές αναφορές θα οδηγήσουν σε διαφορετικές αποδόσεις, και η επιλογή μιας κατάλληλης αναφοράς μπορεί να μην είναι πάντα προφανής ή εύκολη. Μια κακή επιλογή αναφοράς μπορεί να οδηγήσει σε παραπλανητικές εξηγήσεις.
2. **Υπολογιστικό Κόστος:** Αν και απλούστερο από ορισμένες μεθόδους βασισμένες σε διαταραχές, το IG απαιτεί πολλαπλούς υπολογισμούς κλίσεων (έναν για κάθε βήμα k στην προσέγγιση του ολοκληρώματος). Για πολύπλοκα μοντέλα ή μεγάλο αριθμό βημάτων m, αυτό μπορεί να είναι υπολογιστικά απαιτητικό.
3. **Εστίαση σε Γραμμική Διαδρομή:** Το IG εξετάζει μόνο μια ευθεία γραμμική διαδρομή μεταξύ της αναφοράς και της εισόδου. Εάν η πραγματική διαδικασία απόφασης του μοντέλου ακολουθεί μια πολύ πιο περίπλοκη, μη γραμμική διαδρομή στον χώρο των χαρακτηριστικών, η εξήγηση που βασίζεται στη γραμμική διαδρομή μπορεί να μην συλλάβει πλήρως την πολυπλοκότητα.
4. **Ερμηνευσιμότητα των Αποδόσεων:** Αν και οι τιμές IG είναι ποσοτικές, η άμεση ερμηνεία τους ως «βαθμός σημασίας» απαιτεί την κατανόηση του πλαισίου της αναφοράς και της φύσης της ολοκλήρωσης.
5. **Εφαρμογή σε Μη Διαφορίσιμα Μοντέλα:** Το IG βασίζεται θεμελιωδώς στις κλίσεις, καθιστώντας το άμεσα εφαρμόσιμο κυρίως σε διαφορίσιμα μοντέλα όπως τα νευρωνικά δίκτυα. Η επέκτασή του σε μη διαφορίσιμα μοντέλα δεν είναι άμεση.

Για την ανάλυση του LSTM μοντέλου σου, το Integrated Gradients θα μπορούσε να

παρέχει μια εναλλακτική προσέγγιση στο DeepLIFT για την απόδοση των προβλέψεων στα χαρακτηριστικά εισόδου κατά μήκος των χρονικών βημάτων, επιλέγοντας μια κατάλληλη ακολουθία αναφοράς (π.χ., μια ακολουθία μηδενικών αλλαγών τιμών ή μια μέση ακολουθία).

4.7 Permutation Feature Importance (Σημασία Χαρακτηριστικών μέσω Μετάθεσης)

Το Permutation Feature Importance (PFI) είναι μια model-agnostic τεχνική που μετρά την αύξηση στο σφάλμα πρόβλεψης του μοντέλου μετά από τη μετάθεση (δηλαδή, τυχαία αναδιάταξη) των τιμών ενός χαρακτηριστικού. Ένα χαρακτηριστικό θεωρείται «σημαντικό» εάν η αναδιάταξη των τιμών του αυξάνει σημαντικά το σφάλμα του μοντέλου, καθώς αυτό υποδηλώνει ότι το μοντέλο βασιζόταν σε αυτό το χαρακτηριστικό για τις προβλέψεις του. Αντίθετα, ένα χαρακτηριστικό θεωρείται «μη σημαντικό» εάν η αναδιάταξη των τιμών του έχει μικρή ή καμία επίδραση στο σφάλμα του μοντέλου.

4.7.1 Θεωρητική Θεμελίωση

Η βασική ιδέα πίσω από το PFI είναι απλή και διαισθητική: εάν ένα χαρακτηριστικό είναι σημαντικό για την απόδοση ενός μοντέλου, τότε η «καταστροφή» της σχέσης μεταξύ αυτού του χαρακτηριστικού και της μεταβλητής στόχου (με την αναδιάταξη των τιμών του χαρακτηριστικού, η οποία διατηρεί την περιθώρια κατανομή του αλλά σπάει τη σχέση του με τις άλλες μεταβλητές και τον στόχο) θα πρέπει να οδηγήσει σε μια αισθητή πτώση της απόδοσης του μοντέλου.

Το PFI υπολογίζεται αφού ένα μοντέλο έχει ήδη εκπαιδευτεί. Η διαδικασία περιλαμβάνει:

1. **Εκπαίδευση Μοντέλου:** Ένα μοντέλο εκπαιδεύεται κανονικά στα δεδομένα εκπαίδευσης.
2. **Υπολογισμός Βασικής Απόδοσης:** Η απόδοση του μοντέλου (π.χ., ακρίβεια, F1-score, MSE, R-squared) υπολογίζεται σε ένα σύνολο δεδομένων επικύρωσης (validation set) ή δοκιμής (test set). Αυτή είναι η «βασική απόδοση» ή «baseline score». Είναι σημαντικό να χρησιμοποιείται ένα σύνολο δεδομένων που δεν έχει δει το μοντέλο κατά την εκπαίδευση για να αποφευχθεί η υπερεκτίμηση της σημασίας των χαρακτηριστικών που σχετίζονται με την υπερπροσαρμογή.
3. **Επανάληψη για Κάθε Χαρακτηριστικό:** Για κάθε χαρακτηριστικό j στο

σύνολο δεδομένων:

- **Μετάθεση (Permutation):** Οι τιμές του χαρακτηριστικού j στο σύνολο δεδομένων επικύρωσης/δοκιμής αναδιατάσσονται τυχαία. Αυτό δημιουργεί ένα νέο, τροποποιημένο σύνολο δεδομένων όπου το χαρακτηριστικό j δεν έχει πλέον καμία συσχέτιση με την πραγματική τιμή στόχο, αλλά η περιθώρια κατανομή του παραμένει ίδια.
- **Υπολογισμός Νέας Απόδοσης:** Το εκπαιδευμένο μοντέλο χρησιμοποιείται για να κάνει προβλέψεις στο τροποποιημένο σύνολο δεδομένων. Η απόδοση του μοντέλου υπολογίζεται ξανά με βάση αυτές τις νέες προβλέψεις.
- **Υπολογισμός Σημασίας:** Η σημασία του χαρακτηριστικού j υπολογίζεται ως η διαφορά (ή ο λόγος) μεταξύ της βασικής απόδοσης και της νέας απόδοσης μετά τη μετάθεση.
  - Σημασία = Βασική Απόδοση − Απόδοση_μετά_τη_Μετάθεση (για μετρικές όπου υψηλότερη τιμή είναι καλύτερη, π.χ., ακρίβεια)
  - Σημασία = Απόδοση_μετά_τη_Μετάθεση − Βασική Απόδοση (για μετρικές όπου χαμηλότερη τιμή είναι καλύτερη, π.χ. MSE)
  - Σημασία = Βασική Απόδοση / Απόδοση_μετά_τη_Μετάθεση (ή το αντίστροφο)

Η διαδικασία της μετάθεσης μπορεί να επαναληφθεί αρκετές φορές με διαφορετικές τυχαίες αναδιατάξεις για κάθε χαρακτηριστικό, και η μέση σημασία να αναφερθεί για μεγαλύτερη σταθερότητα.

4.7.2 Μαθηματικό Υπόβαθρο

Έστω M το εκπαιδευμένο μοντέλο, X ο πίνακας χαρακτηριστικών του συνόλου δεδομένων επικύρωσης/δοκιμής με n δείγματα και p χαρακτηριστικά, και y ο αντίστοιχος φορέας των πραγματικών τιμών στόχου. Έστω $s = L(y, M(X))$ η βασική απόδοση του μοντέλου, όπου L είναι η συνάρτηση απώλειας ή η μετρική αξιολόγησης.

Για κάθε χαρακτηριστικό $j \in \{1,..., p\}$:

1. Δημιούργησε ένα τροποποιημένο πίνακα χαρακτηριστικών $X_{perm,j}$ αναδιατάσσοντας τυχαία τις τιμές της στήλης j του X.
2. Υπολόγισε την απόδοση $s_j = L(y, M(X_{perm,j}))$ στο τροποποιημένο σύνολο.
3. Η σημασία I_j του χαρακτηριστικού j είναι τότε $I_j = s - s_j$ (ή $s_j - s$ ή $\frac{s}{s_j}$,

ανάλογα με τη μετρική L).

Εάν η μετρική L είναι το μέσο τετραγωνικό σφάλμα (MSE), τότε $I_j = MSE_{perm,j} - MSE_{base}$. Μια υψηλότερη θετική τιμή I_j υποδηλώνει μεγαλύτερη σημασία. Εάν η μετρική L είναι η ακρίβεια, τότε $I_j = Accuracy_{base} - Accuracy_{perm,j}$. Μια υψηλότερη θετική τιμή I_j υποδηλώνει μεγαλύτερη σημασία.

Για δεδομένα χρονοσειρών που χρησιμοποιούνται σε ένα LSTM, κάθε «χαρακτηριστικό» μπορεί να είναι ολόκληρη η χρονοσειρά ενός συγκεκριμένου δείκτη (π.χ., η τιμή κλεισίματος για όλες τις προηγούμενες T ημέρες, ο όγκος για όλες τις προηγούμενες T ημέρες). Εναλλακτικά, εάν το LSTM λαμβάνει πολλαπλά χαρακτηριστικά σε κάθε χρονικό βήμα, η μετάθεση θα μπορούσε να γίνει για ένα από αυτά τα χαρακτηριστικά σε όλα τα χρονικά βήματα ταυτόχρονα. Η ακριβής εφαρμογή εξαρτάται από το πώς ορίζονται τα «χαρακτηριστικά» για το μοντέλο.

4.7.3 Πλεονεκτήματα του Permutation Feature Importance

1. **Ανεξάρτητο από το Μοντέλο (Model-Agnostic):** Μπορεί να χρησιμοποιηθεί με οποιοδήποτε εκπαιδευμένο μοντέλο μηχανικής μάθησης, ανεξάρτητα από την πολυπλοκότητά του ή την εσωτερική του δομή.
2. **Καθολική Ερμηνεία (Global Interpretation):** Παρέχει μια καθολική εκτίμηση της σημασίας κάθε χαρακτηριστικού για τη συνολική απόδοση του μοντέλου στο σύνολο δεδομένων που χρησιμοποιείται για την αξιολόγηση.
3. **Διαισθητική Κατανόηση:** Η έννοια πίσω από το PFI είναι σχετικά εύκολη στην κατανόηση: πόσο υποφέρει το μοντέλο εάν ένα χαρακτηριστικό «χαλάσει»;
4. **Άμεση Μέτρηση της Επίδρασης στην Απόδοση:** Η σημασία συνδέεται άμεσα με την πτώση της απόδοσης του μοντέλου, καθιστώντας την πρακτικά σχετική.
5. **Λαμβάνει υπόψη τις Αλληλεπιδράσεις:** Με την αναδιάταξη ενός μόνο χαρακτηριστικού κάθε φορά, το PFI λαμβάνει υπόψη πώς αυτό το χαρακτηριστικό αλληλεπιδρά με άλλα χαρακτηριστικά στο μοντέλο. Εάν ένα χαρακτηριστικό είναι σημαντικό μόνο μέσω της αλληλεπίδρασής του με ένα άλλο, η αναδιάταξή του θα σπάσει αυτή την αλληλεπίδραση και θα αποκαλύψει τη σημασία του.

4.7.4 Περιορισμοί του Permutation Feature Importance

1. **Υπολογιστικό Κόστος:** Απαιτεί πολλαπλές επαναξιολογήσεις του μοντέλου (μία για κάθε χαρακτηριστικό, και ενδεχομένως πολλαπλές επαναλήψεις ανά χαρακτηριστικό για σταθερότητα). Εάν το μοντέλο είναι αργό στην πρόβλεψη ή εάν υπάρχουν πολλά χαρακτηριστικά, αυτό μπορεί να είναι υπολογιστικά ακριβό.
2. **Πρόβλημα με Συσχετισμένα Χαρακτηριστικά (Correlated Features):** Εάν δύο ή περισσότερα χαρακτηριστικά είναι έντονα συσχετισμένα και το μοντέλο τα χρησιμοποιεί εναλλακτικά, η αναδιάταξη μόνο του ενός από αυτά μπορεί να μην μειώσει σημαντικά την απόδοση, καθώς το μοντέλο μπορεί να εξακολουθεί να βασίζεται στα άλλα συσχετισμένα χαρακτηριστικά. Αυτό μπορεί να οδηγήσει στην υποεκτίμηση της σημασίας των συσχετισμένων χαρακτηριστικών. Από την άλλη πλευρά, εάν ένα χαρακτηριστικό είναι σημαντικό και συσχετίζεται με ένα άλλο, η αναδιάταξη του πρώτου μπορεί να οδηγήσει σε μη ρεαλιστικά ζεύγη τιμών με το δεύτερο (αφού το δεύτερο παραμένει αμετάβλητο), και η πτώση στην απόδοση μπορεί να είναι υπερβολική. Η ερμηνεία απαιτεί προσοχή σε αυτές τις περιπτώσεις.
3. **Εξάρτηση από το Σύνολο Επικύρωσης/Δοκιμής:** Τα αποτελέσματα μπορεί να διαφέρουν ανάλογα με το σύνολο δεδομένων που χρησιμοποιείται για τον υπολογισμό της σημασίας.
4. **Απαιτεί Πραγματικές Τιμές Στόχου:** Το PFI βασίζεται στον υπολογισμό μιας μετρικής απόδοσης, η οποία απαιτεί τις πραγματικές τιμές στόχου. Επομένως, είναι κυρίως για επιβλεπόμενη μάθηση.
5. **Δεν Παρέχει Κατεύθυνση Επίδρασης:** Το PFI υποδεικνύει πόσο σημαντικό είναι ένα χαρακτηριστικό, αλλά όχι *πώς* επηρεάζει την πρόβλεψη (δηλαδή, αν μια αύξηση στην τιμή του αυξάνει ή μειώνει την έξοδο).
6. **Δημιουργία Μη Ρεαλιστικών Στιγμιοτύπων:** Η αναδιάταξη μπορεί να δημιουργήσει στιγμιότυπα δεδομένων που είναι απίθανο να εμφανιστούν στην πραγματικότητα, ειδικά όταν τα χαρακτηριστικά έχουν πολύπλοκες εξαρτήσεις. Το μοντέλο αξιολογείται σε αυτά τα "out-of-distribution" δείγματα.

Παρά τους περιορισμούς, το PFI είναι μια πολύτιμη και ευρέως χρησιμοποιούμενη μέθοδος για την αξιολόγηση της σημασίας των χαρακτηριστικών λόγω της απλότητας, της διαισθητικότητας και της ανεξαρτησίας του από το μοντέλο. Στο πλαίσιο του LSTM σου, θα μπορούσε να χρησιμοποιηθεί για να κατανοηθεί ποια από τις εισόδους (π.χ., χρονοσειρά τιμών κλεισίματος, χρονοσειρά όγκου, χρονοσειρά τεχνικών δεικτών) είναι συνολικά πιο σημαντικές για την απόδοση του μοντέλου στην πρόβλεψη των δράσεων του trader.

ΚΕΦΑΛΑΙΟ 5: ΜΕΘΟΔΟΛΟΓΙΑ

5.1 Συλλογή Δεδομένων

Η παρούσα εργασία εστιάζει στη μετοχή της Apple (AAPL). Για τη μετοχή αυτή, συλλέχθηκαν ιστορικά δεδομένα τιμών από τη χρηματοοικονομική πλατφόρμα Yahoo Finance (API). Τα δεδομένα αφορούν καθημερινές τιμές **Open-High-Low-Close-Volume (OHLCV)**, δηλαδή τιμή έναρξης, μέγιστη και ελάχιστη ενδοημερήσια τιμή, τιμή κλεισίματος και όγκο συναλλαγών. Η περίοδος που καλύφθηκε εκτείνεται από το 2000 έως το πρόσφατο παρόν, εξασφαλίζοντας ένα εκτενές χρονικό εύρος για εκπαίδευση και αξιολόγηση. Η επιλογή των συγκεκριμένων στηλών (OHLCV) κρίθηκε απαραίτητη διότι παρέχουν την ελάχιστη αλλά ουσιώδη πληροφορία για την περιγραφή της αγοράς σε ημερήσια βάση (τιμολογιακή συμπεριφορά και ρευστότητα).

Πέρα από τις βασικές τιμές OHLCV, προκειμένου το μοντέλο να συλλάβει πλούσιες χρηματοοικονομικές πληροφορίες, επεκτείναμε το σύνολο χαρακτηριστικών με ένα ευρύ φάσμα **τεχνικών δεικτών**. Οι δείκτες αυτοί υπολογίστηκαν βάσει των ιστορικών τιμών και του όγκου, με στόχο να κωδικοποιηθεί η τάση, η ορμή, η μεταβλητότητα και το γενικό κλίμα της αγοράς κάθε χρονική στιγμή (βλ. πηγή 10). Συγκεκριμένα, αντλήθηκαν χαρακτηριστικά από τις ακόλουθες κατηγορίες:

- **Δείκτες Τάσης:** Περιλαμβάνουν τους Απλούς Κινητούς Μέσους Όρους (Simple Moving Averages – SMA) 10, 30 και 50 ημερών, τους Εκθετικούς Κινητούς Μέσους (Exponential Moving Averages – EMA), καθώς και τον Δείκτη Σύγκλισης/Απόκλισης Κινητών Μέσων (Moving Average Convergence Divergence – MACD). Οι δείκτες αυτοί αποτυπώνουν τη γενική κατεύθυνση της αγοράς και βοηθούν στην ανίχνευση καθιερωμένων ανοδικών ή καθοδικών τάσεων.
- **Δείκτες Ορμής:** Αφορούν τον Δείκτη Σχετικής Ισχύος (Relative Strength Index – RSI), τον Στοχαστικό Ταλαντωτή (Stochastic Oscillator), και τον Ρυθμό Αλλαγής τιμής (Rate of Change – ROC). Οι δείκτες ορμής μετρούν την ταχύτητα και το μέγεθος των μεταβολών των τιμών, παρέχοντας ενδείξεις για υπεραγορασμένες ή υπερπουλημένες συνθήκες στην αγορά.
- **Δείκτες Μεταβλητότητας:** Περιλαμβάνουν τις Ζώνες Bollinger (άνω και κάτω όριο, καθώς και %B), και το Μέσο Εύρος Αληθείας (Average True Range – ATR). Αυτά τα χαρακτηριστικά ποσοτικοποιούν την αστάθεια των τιμών – οι

Ζώνες Bollinger ορίζονται από ένα κινητό μέσο όρο και ένα εύρος ±2 τυπικών αποκλίσεων, ενώ το ATR υπολογίζεται ως ο μέσος όρος του αληθινού εύρους των τιμών (εύρος ημερήσιων διακυμάνσεων) σε ένα παράθυρο 14 ημερών.

- **Δείκτες Όγκου:** Αποτελούνται από τον Δείκτη Συσσώρευσης/Αποδιανομής όγκου (On-Balance Volume – OBV) και τον Ταλαντωτή Ροής Χρήματος (Money Flow Index – MFI), καθώς και κινητούς μέσους όρους του όγκου. Οι δείκτες αυτοί συσχετίζουν τις μεταβολές της τιμής με τον όγκο συναλλαγών, σηματοδοτώντας το κατά πόσο μια κίνηση τιμής υποστηρίζεται από «ισχυρά χέρια» (υψηλό όγκο) ή όχι. Για παράδειγμα, ο OBV υπολογίζεται μέσω αθροιστικής πρόσθεσης ή αφαίρεσης του όγκου ανά ημέρα ανάλογα με την ανοδική ή καθοδική κίνηση της τιμής.
- **Δείκτες Ευρύτερου Πλαισίου:** Περιλαμβάνεται η ημερήσια ποσοστιαία μεταβολή του δείκτη S&P 500 (ως δείκτης αναφοράς της συνολικής αγοράς) και του τεχνολογικού κλάδου. Αυτά τα χαρακτηριστικά δίνουν το γενικό κλίμα αγοράς, ώστε το μοντέλο να λαμβάνει υπόψη του αν η μετοχή κινείται συνολικά σε ευνοϊκό ή δυσμενές περιβάλλον αγοράς.

Οι παραπάνω τεχνικοί δείκτες υπολογίστηκαν με τις καθιερωμένες μεθόδους της τεχνικής ανάλυσης, αξιοποιώντας κυλιόμενα παράθυρα και διαφορές τιμών. Για παράδειγμα, ο RSI 14 ημερών προκύπτει από τον λόγο του μέσου όρου των ημερήσιων ανόδων προς τον μέσο όρο των καθόδων στο ίδιο παράθυρο, ενώ ο MACD προκύπτει ως η διαφορά μεταξύ του EMA 12 ημερών και του EMA 26 ημερών της τιμής κλεισίματος. Ο δείκτης Momentum ορίστηκε ως η διαφορά της τιμής κλεισίματος σε σχέση με την τιμή n ημέρες πριν (π.χ. 5 ημέρες), παρέχοντας ένα μέτρο της πρόσφατης επιτάχυνσης ή επιβράδυνσης της αγοράς. Όλοι οι υπολογισμοί των δεικτών πραγματοποιήθηκαν αυτόματα μέσω του κώδικα (βλ. data_collection_triple_barrier.py), διασφαλίζοντας την ορθότητα των τιμών και την αποφυγή μεροληψίας κατά την μετέπειτα χρήση τους στο μοντέλο.

5.1.2 Σήμανση Δεδομένων (Labeling)

Για την εποπτεία του μοντέλου απαιτείται ο ορισμός κατάλληλων ετικετών στόχου (target labels) που αντιστοιχούν στις κατηγορίες ενέργειας Buy, Hold ή Sell σε κάθε χρονική στιγμή. Ο καθορισμός των ετικετών αυτών αποτελεί πρόκληση, καθώς θα πρέπει αφενός να αντανακλά ρεαλιστικές στρατηγικές συναλλαγών, αφετέρου να είναι συμβατός με τις μεθόδους εξηγησιμότητας (να μπορεί δηλαδή να ερμηνευθεί εύκολα). Στο πλαίσιο της μελέτης αυτής αναπτύχθηκαν και αξιολογήθηκαν δύο διαφορετικές προσεγγίσεις σήμανσης:

(Α) Προσαρμοσμένη Πολυπαραγοντική Σήμανση: Αναπτύξαμε μια ιεραρχική,

κανόνα-based μεθοδολογία για την ανάθεση των ετικετών, με στόχο τη βελτιωμένη ερμηνευσιμότητα των αποφάσεων (υλοποίηση στο data_collection_custom_labeling.py). Σε αυτή την προσέγγιση ορίζονται σαφείς κανόνες πολλαπλών κριτηρίων, οργανωμένοι κατά προτεραιότητα, για τον εντοπισμό ισχυρών σημάτων αγοράς ή πώλησης. Ανά πάσα ημέρα, ελέγχεται διαδοχικά μια σειρά από συνθήκες: αν ικανοποιηθεί μια συνθήκη υψηλής προτεραιότητας, παράγεται σήμα Buy ή Sell και δεν εξετάζονται χαμηλότερης προτεραιότητας κανόνες. Οι κανόνες αυτοί διαμορφώθηκαν με γνώμονα γνωστές εμπειρικές στρατηγικές τεχνικής ανάλυσης, συνδυάζοντας πολλαπλούς δείκτες για πιο αξιόπιστα σήματα. Συνοπτικά, η λογική σήμανσης έχει ως εξής:

- **Προτεραιότητα 1: Ισχυρά σήματα RSI** (υπέρ-πουλημένη/υπέρ-αγορασμένη αγορά με επιβεβαίωση ορμής και όγκου):

    - **Σήμα "Buy":** RSI < 25 σε συνδυασμό με θετικό momentum τιμής και ιδιαίτερα αυξημένο όγκο συναλλαγών.
    - **Σήμα "Sell":** RSI > 75 με αρνητικό momentum τιμής και ιδιαίτερα αυξημένο όγκο.
    - **Αιτιολόγηση:** Οι εξαιρετικά χαμηλές ή υψηλές τιμές RSI υποδηλώνουν συνθήκες υπερπώλησης ή υπεραγοράς αντίστοιχα (βλ. πηγή 10). Όταν αυτές οι ακραίες ενδείξεις συνοδεύονται από σύμφωνη ορμή της τιμής (π.χ. ανοδική αντιστροφή μετά από πολύ χαμηλό RSI) και μεγάλο όγκο που δίνει βαρύτητα στο σήμα, θεωρείται πιθανή η επικείμενη αλλαγή τάσης.
- **Προτεραιότητα 2: Διασταύρωση Κινητών Μέσων** (σήματα τάσης με πολλαπλές επιβεβαιώσεις):

    - **Σήμα "Buy":** Ο βραχυπρόθεσμος κινητός μέσος (SMA 10 ημερών) διασπά ανοδικά τον μεσοπρόθεσμο κινητό μέσο (SMA 30 ημερών), ενώ το RSI βρίσκεται κάτω από το 50 (ουδέτερη ζώνη) και ο όγκος είναι υψηλός, σε ένα γενικά ανοδικό περιβάλλον αγοράς.
    - **Σήμα "Sell":** Ο SMA 10 διασπά καθοδικά τον SMA 30, με RSI πάνω από 50 και υψηλό όγκο, σε γενικά πτωτικό κλίμα αγοράς.
    - **Αιτιολόγηση:** Οι διασταυρώσεις μικρών και μεγαλύτερων κινητών μέσων είναι κλασικά σήματα αλλαγής τοπικής τάσης (βλ. πηγή 10). Η απαίτηση επιπρόσθετων φίλτρων (όπως ουδέτερου RSI και αυξημένου όγκου, καθώς και ευρύτερου θετικού/αρνητικού κλίματος) μειώνει τα ψευδοσήματα, διασφαλίζοντας ότι το σήμα αντανακλά μια πραγματική αλλαγή στην τάση με την υποστήριξη της αγοράς.

- **Προτεραιότητα 3: Σήματα Bollinger Bands** (fallback σήματα mean reversion):

    - **Σήμα "Buy":** Η τιμή κλεισίματος πέφτει κάτω από την κάτω ζώνη Bollinger, το RSI βρίσκεται < 45 (ελαφρώς στη ζώνη υπερπώλησης) και ο όγκος είναι υψηλός, ενώ δεν παρατηρείται πολύ έντονη καθοδική τάση στην ευρύτερη αγορά εκείνη τη στιγμή.
    - **Σήμα "Sell":** Η τιμή υπερβαίνει την άνω ζώνη Bollinger, το RSI > 55, ο όγκος υψηλός, και δεν υπάρχει έντονη ανοδική δυναμική ευρύτερα.
    - **Αιτιολόγηση:** Οι Ζώνες Bollinger ορίζουν ένα δυναμικό εύρος τιμών (βλ. πηγή 10). Όταν η τιμή βγαίνει εκτός αυτού του εύρους, συνήθως υποδηλώνει ακραία κίνηση (πολύ πάνω ή πολύ κάτω από τον μέσο όρο). Τα σήματα mean reversion βασίζονται στην υπόθεση ότι τέτοιες ακραίες κινήσεις συχνά ακολουθούνται από αντιστροφή προς τον μέσο. Η χρήση μέτριων τιμών RSI και η απουσία ισχυρής τάσης ως προϋποθέσεις, διασφαλίζει ότι το σήμα Buy/Sell δίνεται μόνο όταν η αγορά ενδεχομένως θα επανέλθει από μια βραχυχρόνια υπερβολή και όχι όταν βρίσκεται σε παρατεταμένο trend.
- **Προεπιλογή – Hold:** Εάν καμία από τις παραπάνω συνθήκες δεν ικανοποιείται σε μια ημέρα, τότε η ετικέτα ορίζεται ως "Hold" (ουδετερότητα/αναμονή). Αυτό αντιπροσωπεύει είτε συνθήκες αγοράς όπου δεν υπάρχει ισχυρό σήμα, είτε αβεβαιότητα όπου η καλύτερη δράση θα ήταν η μη-δράση (διακράτηση θέσης). Σημειώνεται ότι η κατηγορία Hold λειτουργεί ως φύλακας για την αποφυγή υπερβολικά συχνών συναλλαγών σε θορυβώδες περιβάλλον.

(Β) Μέθοδος "Triple Barrier": Πρόκειται για μια πιο μαθηματική/συμβατική μέθοδο σήμανσης που προτάθηκε από τον Lopez de Prado για προβλήματα χρηματοοικονομικής πρόβλεψης (βλ. πηγή 30). Η μέθοδος του τριπλού φράγματος ορίζει τις ετικέτες βάσει της μελλοντικής μεταβολής της τιμής σε ένα καθορισμένο χρονικό παράθυρο, θέτοντας τρεις φράκτες: έναν άνω (profit-taking), έναν κάτω (stop-loss) και έναν χρονικό φράκτη. Σε κάθε χρονικό σημείο (ημέρα t), κοιτάμε προς τα εμπρός σε έναν ορίζοντα N ημερών (π.χ. 5 ημέρες) και εξετάζουμε: (1) αν η τιμή αυξηθεί εντός αυτού του διαστήματος πάνω από ένα προκαθορισμένο ποσοστό (π.χ. +3%), (2) αν αντιθέτως μειωθεί κάτω από ένα όριο (–3%), ή (3) αν περάσει το N-ήμερο χωρίς να χτυπηθεί κανένα όριο. Αναλόγως, αποδίδεται ετικέτα ως εξής:

- **Άνω φράγμα – "Buy":** Εάν η τιμή μέσα στις επόμενες N ημέρες φθάσει ή

ξεπεράσει ένα ποσοστό ανόδου (profit target), θεωρούμε ότι η σωστή ενέργεια στο χρόνο t θα ήταν η αγορά. Για παράδειγμα, με όριο +3% και ορίζοντα 5 ημερών, αν η τιμή της μετοχής σημειώσει άνοδο ≥3% μέσα σε 5 ημέρες από την ημέρα t, η ημέρα t λαμβάνει ετικέτα "Buy".

- **Κάτω φράγμα – "Sell":** Εάν η τιμή πέσει εντός N ημερών κατά περισσότερο από ένα όριο πτώσης (π.χ. –3%), τότε η ενδεδειγμένη ενέργεια στο t θα ήταν πώληση (Sell), καθώς εκ των υστέρων φαίνεται ότι υπήρξε σημαντική πτωτική κίνηση.
- **Οριζόντιος φράγμα – "Hold":** Εάν παρέλθει το χρονικό παράθυρο N ημερών χωρίς η τιμή να κινηθεί πέρα από τα προκαθορισμένα όρια (ούτε +3% ούτε –3%), τότε η κίνηση θεωρείται ουδέτερη. Στην περίπτωση αυτή, η ημέρα t λαμβάνει ετικέτα "Hold", εκτός αν σημειώθηκε μια πολύ μικρή μεταβολή που μπορούμε να θεωρήσουμε αμελητέα. Συνήθως ορίζεται ένα ουδέτερο εύρος γύρω από το 0 (π.χ. ±1%), εντός του οποίου η τελική μεταβολή θεωρείται μη-σημαντική. Αν η τελική καθαρή μεταβολή τιμής στο διάστημα αυτό είναι ελαφρώς θετική (άνω του +1% αλλά κάτω από το profit target), δίνεται τελικά ετικέτα "Buy" (η τιμή ανέβηκε μεν, όχι όμως αρκετά για να χτυπήσει το άνω όριο), ενώ αν είναι λίγο αρνητική (χαμηλότερη του –1% χωρίς να χτυπηθεί το stop-loss) δίνεται ετικέτα "Sell".

Η μέθοδος τριπλού φράγματος παράγει έτσι τρεις κατηγορίες (Buy/Hold/Sell) με βάση αντικειμενικά κριτήρια κίνησης της τιμής. Στα πλαίσια της υλοποίησής μας, επιλέξαμε ορίζοντα N=5 ημερών, όρια profit-taking και stop-loss στο ±3%, και ουδέτερη ζώνη ±1%. Αυτές οι παράμετροι επελέγησαν έτσι ώστε να συλλαμβάνονται έγκαιρα σημαντικές κινήσεις της μετοχής (3% σε μία εβδομάδα) χωρίς να χαρακτηρίζονται ως Buy/Sell μικρές διακυμάνσεις. Η προσέγγιση αυτή ευθυγραμμίζεται με ρεαλιστικά σενάρια trading: για παράδειγμα, ένα σήμα Buy σε μια μέρα σημαίνει ότι μέσα στην επόμενη εβδομάδα θα υπήρχε ευκαιρία λήψης κέρδους ~3%, ενώ ένα σήμα Sell αντανακλά την αποφυγή μιας ζημιάς ~3%. Επιπλέον, η χρήση συμμετρικών ορίων ανόδου/πτώσης βοηθά στο να διατηρούνται ισορροπημένες οι τάξεις των ετικετών (αποφεύγοντας μονομερή επικράτηση της κατηγορίας Hold).

Σύγκριση & Επιλογή Προσέγγισης: Η ιεραρχική στρατηγική κανόνων (A) προσφέρει διαφάνεια – κάθε σήμα μπορεί να εξηγηθεί άμεσα από τις συνθήκες που το δημιούργησαν, κάτι που είναι ιδιαίτερα συμβατό με τις μεθόδους εξηγησιμότητας που εφαρμόζουμε (XAI). Ωστόσο, ενδέχεται να χάσει ορισμένα σήματα, καθώς βασίζεται σε προκαθορισμένους κανόνες. Από την άλλη, η μέθοδος του τριπλού φράγματος (B) είναι πιο γενική και συνδέεται άμεσα με μελλοντικές αποδόσεις,

όμως λειτουργεί ως ένα βαθμό σαν «μαύρο κουτί» ως προς το γιατί μια ημέρα έλαβε την εκάστοτε ετικέτα (δεν δίνεται δηλαδή με σαφήνεια η αιτία – π.χ. «λόγω RSI» ή «λόγω Bollinger»). Για το λόγο αυτό, στην παρούσα εργασία χρησιμοποιήθηκαν αμφότερες οι προσεγγίσεις: αρχικά υλοποιήθηκε η σήμανση με triple barrier για τη δημιουργία ενός ρεαλιστικού συνόλου δεδομένων εκπαίδευσης, αλλά στη συνέχεια αναπτύχθηκε η προσαρμοσμένη στρατηγική κανόνων, η οποία τελικά κρίθηκε προτιμότερη για την εξαγωγή ερμηνεύσιμων αποτελεσμάτων. Τα μοντέλα εκπαιδεύτηκαν και με τις δύο εναλλακτικές ετικετών, ώστε να συγκριθεί η απόδοσή τους, και διαπιστώθηκε ότι η ερμηνευσιμότητα των προβλέψεων ήταν σημαντικά βελτιωμένη με την ιεραρχική σήμανση (οι εξηγητικές μέθοδοι παρήγαγαν πιο σαφή αποτελέσματα), χωρίς αξιοσημείωτη θυσία στην ακρίβεια πρόβλεψης.

5.2 Οπτικοποίηση Δεδομένων

Για την καλύτερη κατανόηση των δεδομένων και την επαλήθευση της ορθότητας τόσο της συλλογής όσο και της σήμανσης, δημιουργήθηκε ένα ειδικό εργαλείο οπτικοποίησης (stock_visualizer.py). Η εφαρμογή αυτή παράγει διαδραστικά διαγράμματα τιμών και αποδόσεων, συνδυάζοντας πολλαπλές πληροφορίες σε μια συνεκτική παρουσίαση. Συγκεκριμένα, σχεδιάζονται διαγράμματα **candlestick** για τη μετοχή AAPL, όπου κάθε κερί απεικονίζει την τιμή ανοίγματος, ανώτατη/κατώτατη τιμή και τιμή κλεισίματος της ημέρας, χρωματισμένο πράσινο ή κόκκινο ανάλογα με την κατεύθυνση της ημέρας (άνοδος ή πτώση). Στο ίδιο διάγραμμα ενσωματώνεται στο κάτω μέρος και ο ημερήσιος όγκος συναλλαγών (με μπάρες), ώστε να διακρίνεται εύκολα σε ποιες περιόδους υπήρξε ασυνήθιστα υψηλός όγκος που μπορεί να ενισχύει την αξιοπιστία ορισμένων κινημάτων τιμής.

Παράλληλα, δημιουργείται ένα δεύτερο διάγραμμα που απεικονίζει τις **ημερήσιες αποδόσεις** (ποσοστιαία μεταβολή τιμής κάθε ημέρας σε σχέση με την προηγούμενη). Στο διάγραμμα αυτό, οι θετικές ημερήσιες αποδόσεις χρωματίζονται με πράσινο (πράσινες σκιές πάνω από τον οριζόντιο άξονα του 0%), ενώ οι αρνητικές με κόκκινο (σκιές κάτω από το 0%), παρέχοντας έτσι μια άμεση οπτική ένδειξη των κερδοφόρων και ζημιογόνων ημερών. Μια διακεκομμένη γραμμή στο 0% λειτουργεί ως σημείο αναφοράς μεταξύ κερδών και απωλειών. Με αυτή την απεικόνιση γίνεται εύκολο να εντοπιστούν περίοδοι υψηλής μεταβλητότητας (όπου οι σκιές έχουν μεγάλο εύρος) καθώς και τυχόν εναλλασσόμενα μοτίβα διαδοχικών κερδών/ζημιών.

Ο σκοπός των παραπάνω οπτικοποιήσεων είναι πολλαπλός. Αφενός, επιτρέπουν μια εξερευνητική ανάλυση των δεδομένων: μπορούμε να εντοπίσουμε οπτικά

σημαντικά γεγονότα (π.χ. απότομες πτώσεις ή ανόδους της τιμής, συνοδευόμενες από όγκο) και να επιβεβαιώσουμε ότι οι τεχνικοί δείκτες αντιδρούν όπως αναμένεται σε αυτά (π.χ. ο RSI πέφτει σε έντονες πτώσεις τιμής). Αφετέρου, τα διαγράμματα παρέχουν ένα εργαλείο επαλήθευσης της διαδικασίας σήμανσης: μπορούμε να σημάνουμε επάνω στο candlestick διάγραμμα τις μέρες όπου το dataset φέρει ετικέτα "Buy" ή "Sell" και να δούμε αν αυτές ευθυγραμμίζονται με την ανθρώπινη διαισθητική κατανόηση (π.χ. αν εμφανίζονται "Buy" ετικέτες μετά από μεγάλα sell-off και "Sell" μετά από κορυφές αγοράς). Στο πλαίσιο της υλοποίησης, τα διαγράμματα μπορούν να αποθηκευτούν ως εικόνες υψηλής ανάλυσης για περαιτέρω μελέτη ή ενσωμάτωση στην αναφορά αποτελεσμάτων. Συνολικά, η οπτικοποίηση αποτέλεσε σημαντικό μέρος της μεθοδολογίας, λειτουργώντας ως γέφυρα κατανόησης μεταξύ των ακατέργαστων δεδομένων/σημάτων και της συμπεριφοράς του εκπαιδευμένου μοντέλου.

5.3 Προεπεξεργασία Δεδομένων

Πριν τα δεδομένα εισαχθούν στο μοντέλο LSTM, εφαρμόστηκε μια σειρά από βήματα προεπεξεργασίας ώστε να διασφαλιστεί η συμβατότητα, η ποιότητα και η κατάλληλη κλιμάκωσή τους. Η προεπεξεργασία είναι κρίσιμη σε χρονοσειρές χρηματοοικονομικών δεδομένων, διότι πρέπει να διατηρηθεί η χρονική δομή (η σειρά των γεγονότων) αποφεύγοντας ταυτόχρονα τη διαρροή πληροφοριών από το μέλλον προς το παρελθόν.

- **Κανονικοποίηση Χαρακτηριστικών:** Όλες οι μεταβλητές (τιμές, δείκτες) κλιμακώθηκαν σε ένα ενιαίο εύρος τιμών για να αποφευχθούν ανεπιθύμητες διακυμάνσεις λόγω μεγέθους μονάδας. Συγκεκριμένα, χρησιμοποιήθηκε **κανονικοποίηση ελαχίστου-μεγίστου (Min-Max Scaler)** ώστε κάθε χαρακτηριστικό να μετασχηματιστεί στο διάστημα [0,1]. Για παράδειγμα, η τιμή κλεισίματος, ο όγκος, και ο RSI – παρότι μετρούν διαφορετικά μεγέθη – μετά την κανονικοποίηση συγκρίνονται επί ίσοις όροις από το νευρωνικό δίκτυο. Η κανονικοποίηση πραγματοποιήθηκε με βάση τις τιμές του σετ εκπαίδευσης (fit στον εκπαιδευτικό υπο-πίνακα), και κατόπιν εφαρμόστηκε στα δεδομένα επαλήθευσης και δοκιμής, εξασφαλίζοντας ότι δεν διαρρέει πληροφορία από το μέλλον κατά το στάδιο της εκπαίδευσης. Το αποτέλεσμα είναι δεδομένα απαλλαγμένα από διαστασιακές διαφορές (π.χ. η τιμή της μετοχής ~150 θεωρείται «μεγάλη» ενώ ένας RSI ~30 θεωρείται «χαμηλός», αλλά μετά την κανονικοποίηση όλες οι μεταβλητές περιορίζονται 0–1).

- **Μετασχηματισμός σε Ακολουθίες:** Δεδομένης της φύσης των LSTM, τα

δεδομένα χρονοσειράς μετατράπηκαν σε συνεχείς ακολουθίες σταθερού μήκους, που θα αποτελούν τις εισόδους του μοντέλου. Συγκεκριμένα, διαμορφώθηκαν σειρές ολισθαίνοντων παραθύρων μήκους 10, 30 και 50 ημερών. Αυτό σημαίνει ότι για κάθε ημέρα t (μετά την πρώτη διαθέσιμη περίοδο 50 ημερών), δημιουργούμε τρεις υπο-ακολουθίες:

  - μια βραχυπρόθεσμη που περιλαμβάνει τις προηγούμενες 10 ημέρες (τις 10 τελευταίες εγγραφές χαρακτηριστικών πριν το t),
  - μία μεσοπρόθεσμη 30 ημερών,
  - και μία μακροπρόθεσμη 50 ημερών. Κάθε υπο-ακολουθία συνοδεύεται από την ετικέτα δράσης (Buy/Hold/Sell) της ημέρας t που ακολουθεί. Οι ακολουθίες αυτές είναι επικαλυπτόμενες (sliding window): η ημέρα t συμμετέχει ως μέρος των παραθύρων των επόμενων N ημερών (για παράδειγμα, η ημέρα t=50 είναι η πρώτη που δημιουργεί πλήρες 50-ήμερο παράθυρο, η t=51 θα έχει παράθυρα 10, 30, 50 που ξεκινούν μια ημέρα αργότερα κ.ο.κ.). Με αυτό τον τρόπο, μεγιστοποιείται ο αριθμός των δειγμάτων εκπαίδευσης αξιοποιώντας όλη τη διαθέσιμη πληροφορία, ενώ παράλληλα διατηρείται η χρονική σειρά των γεγονότων μέσα σε κάθε δείγμα. Η διαδικασία υλοποιήθηκε με κώδικα (βλ. συνάρτηση create_multi_input_sequences στο lstm_model.py) που διέρχεται σειριακά τα δεδομένα και συναρμολογεί τις τριάδες υπο-ακολουθιών.

- **Διαχωρισμός Συνόλων και Χρονική Ακεραιότητα:** Αφού δημιουργήθηκαν οι ακολουθίες, το σύνολο δεδομένων διαιρέθηκε χρονικά σε υποσύνολα εκπαίδευσης, επαλήθευσης και δοκιμής. Συγκεκριμένα, τηρήθηκε λόγος ~70% των δειγμάτων για εκπαίδευση (από παλαιότερα χρόνια), ~15% για επαλήθευση (validation) και ~15% για τελικό έλεγχο (test). Ο διαχωρισμός έγινε χρονικά διαδοχικά (time-series split) και όχι τυχαία – δηλαδή τα τελευταία χρονικά δεδομένα αποτελούν το test set, εξασφαλίζοντας ότι το μοντέλο πάντα αξιολογείται σε μελλοντικές περιόδους που δεν έχει «δει» κατά την εκπαίδευση. Με αυτό τον τρόπο αποφεύγεται η χρονική διαρροή πληροφοριών και διασφαλίζεται ότι η αξιολόγηση προσομοιώνει ρεαλιστική πρόβλεψη μελλοντικών δεδομένων. Τέλος, η μεταβλητή στόχος (ετικέτα Action) κωδικοποιήθηκε αριθμητικά σε 0, 1, 2 (Sell = 0, Hold = 1, Buy = 2), και στη συνέχεια μετατράπηκε σε one-hot μορφή κατά την εκπαίδευση (διάνυσμα πιθανότητας 3 διαστάσεων) για την κατάλληλη χρήση της συνάρτησης softmax.

Με την ολοκλήρωση των παραπάνω βημάτων, τα δεδομένα ήταν έτοιμα προς χρήση από το νευρωνικό δίκτυο: κανονικοποιημένα, σε μορφή διαδοχικών παραθύρων (time series segments), με διαχωρισμένες χρονικές ενότητες για έλεγχο, και με σαφώς ορισμένες ετικέτες δράσης. Αυτό το pipeline προεπεξεργασίας διασφαλίζει ότι εισάγουμε στο μοντέλο πληροφορία που είναι συγκρίσιμη, αντιπροσωπευτική των αγορών και χωρίς μεροληψία από μελλοντικά δεδομένα, μεγιστοποιώντας τις πιθανότητες το LSTM να μάθει ουσιαστικά πρότυπα αντί για ψευδοσυσχετίσεις.

5.4 Αρχιτεκτονική Μοντέλου LSTM

Το κεντρικό μοντέλο πρόβλεψης που αναπτύχθηκε είναι ένα προσαρμοσμένο νευρωνικό δίκτυο LSTM (Long Short-Term Memory) με πολλαπλές εισόδους χρονικών σειρών (βλ. πηγή 26). Ο σχεδιασμός του αρχιτεκτονικού αυτού σχήματος έγινε με στόχο να αξιοποιήσει πληροφορίες από διαφορετικού μήκους χρονικά παράθυρα, επιτρέποντας στο μοντέλο να «βλέπει» τόσο τις βραχυπρόθεσμες διακυμάνσεις όσο και τις μακροπρόθεσμες τάσεις της αγοράς. Συνοπτικά, η αρχιτεκτονική αποτελείται από τρεις παράλληλους κλάδους εισόδου και έναν τελικό συγχωνευτικό κλάδο εξόδου:

- **Είσοδοι Πολλαπλών Κλιμάκων:** Το δίκτυο διαθέτει τρία διακριτά υποδίκτυα εισόδου, καθένα εκ των οποίων δέχεται ως είσοδο μια ακολουθία διαδοχικών ημερών διαφορετικού μήκους. Συγκεκριμένα, ορίστηκαν είσοδοι διαστάσεων (10×F), (30×F) και (50×F), όπου F είναι ο αριθμός χαρακτηριστικών ανά ημέρα (το πλήθος των στηλών χαρακτηριστικών μετά την προεπεξεργασία – π.χ. OHLCV + τεχνικοί δείκτες, συνολικά ~23 χαρακτηριστικά). Αυτές οι είσοδοι αντιστοιχούν σε βραχυπρόθεσμο, μεσοπρόθεσμο και μακροπρόθεσμο παράθυρο αντίστοιχα. Έτσι, την ίδια χρονική στιγμή, το μοντέλο «βλέπει» τις τελευταίες 10, 30 και 50 ημέρες των αγορών, κάθε μία από διαφορετική οπτική γωνία, πριν λάβει μια απόφαση πρόβλεψης.

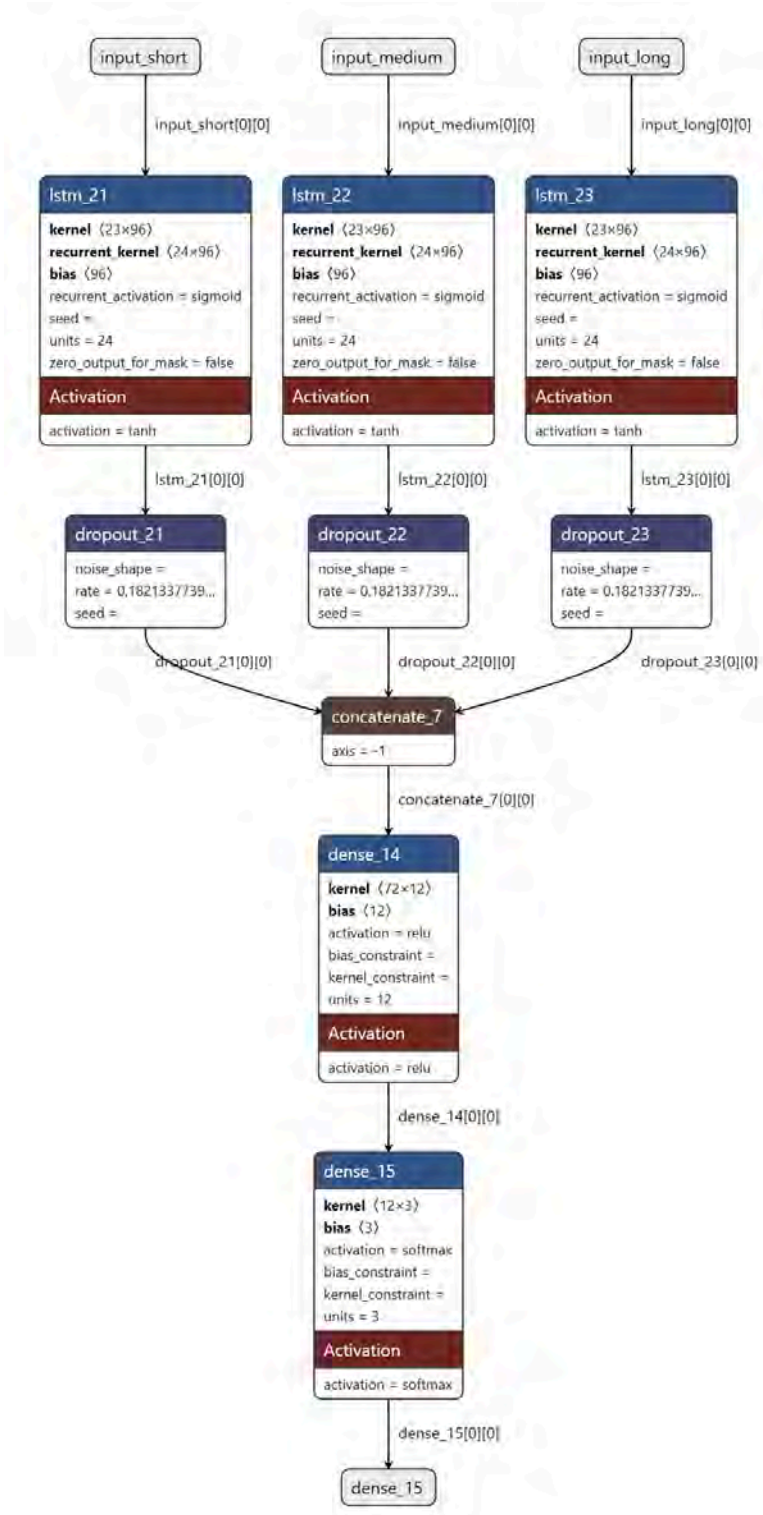
- **Κρυφά Επίπεδα LSTM ανά Κλάδο:** Κάθε ένας από τους τρεις κλάδους εισόδου συνδέεται με ένα επίπεδο LSTM το οποίο επεξεργάζει την ακολουθία του αντίστοιχου παραθύρου. Τα LSTM δίκτυα επιλέχθηκαν για την ικανότητά τους να διατηρούν μνήμη μακράς και βραχείας διάρκειας, μέσω των πυλών κατάστασης , κάτι που τα καθιστά ιδιαίτερα κατάλληλα για χρονικές σειρές όπου γεγονότα του παρελθόντος μπορεί να επηρεάζουν το μέλλον με διαφορετικές χρονικές καθυστερήσεις. Κάθε επίπεδο LSTM έχει **24 μονάδες (neurons)**, όπως υποδεικνύεται από την παράμετρο units = 24 στα layers lstm_21, lstm_22, και lstm_23 στην εικόνα. Χρησιμοποιεί recurrent_activation =

sigmoid και activation = tanh. Επιπλέον, μετά από κάθε LSTM προστίθεται ένα επίπεδο τυχαίας απόρριψης **Dropout** με ποσοστό rate = 0.182133739... (περίπου 18%), το οποίο μηδενίζει τυχαία κάποιες μονάδες σε κάθε εποχή εκπαίδευσης. Το Dropout λειτουργεί ως κανονικοποιητής, μειώνοντας τον κίνδυνο υπερεκπαίδευσης (overfitting) σε ιστορικά δεδομένα, ιδιαίτερα σημαντικό σε χρηματοοικονομικές χρονοσειρές που ενέχουν θόρυβο και μεταβαλλόμενες συνθήκες.

- **Συγχώνευση και Πυκνό Επίπεδο Εξόδου:** Οι τρεις κλάδοι LSTM στη συνέχεια συγχωνεύονται (**concatenate**) σε ένα ενιαίο επίπεδο, ενώνοντας τα τρία κρυφά διανύσματα (ένα από κάθε LSTM) σε ένα μεγαλύτερο διάνυσμα διάστασης 72 (24×3). Αυτό το ενωμένο διάνυσμα αντιπροσωπεύει την συλλογική πληροφόρηση που το μοντέλο άντλησε από όλες τις χρονικές κλίμακες. Ακολουθεί ένα πλήρως συνδεδεμένο επίπεδο **Dense** με 12 νευρώνες και συνάρτηση ενεργοποίησης **ReLU**, που μαθαίνει μια τελική μη-γραμμική συσχέτιση των συνδυασμένων χαρακτηριστικών. Τέλος, το δίκτυο οδηγείται σε ένα επίπεδο εξόδου Dense με 3 μονάδες και ενεργοποίηση **softmax**. Το επίπεδο αυτό παράγει μια κατανομή πιθανότητας πάνω στις τρεις κλάσεις δράσης (P(Sell),P(Hold),P(Buy)), δηλαδή την πρόβλεψη του μοντέλου για την πιθανότερη ενέργεια την επόμενη ημέρα. Η χρήση της softmax διασφαλίζει ότι οι τρεις πιθανότητες αθροίζουν στο 1 και μπορούν να ερμηνευθούν απευθείας ως εμπιστοσύνη του μοντέλου σε κάθε ενέργεια.

- **Λεπτομέρειες Εκπαίδευσης:** Το μοντέλο συντέθηκε και εκπαιδεύτηκε με συνάρτηση απώλειας **categorical cross-entropy** (στα πολυταξικά προβλήματα είναι το τυπικό metric) και με βελτιστοποιητή **Adam**, ο οποίος προσφέρει αποδοτική προσαρμογή του βήματος μάθησης. Η εκπαίδευση πραγματοποιήθηκε σε 100 εποχές μέγιστα, με αρχικό learning rate 0.001. Χρησιμοποιήθηκε **early stopping** με παρακολούθηση του error επαλήθευσης, ώστε να τερματιστεί η εκπαίδευση αν η απόδοση πάψει να βελτιώνεται για 5 συνεχόμενες εποχές – αυτό απέτρεψε την υπερεκπαίδευση (overfitting) και μείωσε τον συνολικό χρόνο εκπαίδευσης, τερματίζοντας τυπικά γύρω στην 20ή-30ή εποχή. Επίσης εφαρμόστηκε προσαρμογή learning rate (ReduceLROnPlateau) μειώνοντας το κατά ένα δεκαπλασιασμό όταν η βελτίωση κόβεται, βοηθώντας το μοντέλο να συγκλίνει σε ένα τοπικό ελάχιστο πιο ομαλά. Η ακρίβεια πρόβλεψης (accuracy) χρησιμοποιήθηκε ως κύριο μέτρο απόδοσης κατά την εκπαίδευση και επαλήθευση, δεδομένου ότι οι τάξεις Buy/Hold/Sell ήταν σχετικά ισοζυγισμένες μετά τη σήμανση (η ακρίβεια είναι κατάλληλο metric σε περίπου ισομεγέθεις τάξεις).

- **Καταλληλότητα LSTM & Πολυ-παραθυρική Προσέγγιση:** Η επιλογή της αρχιτεκτονικής LSTM δικαιολογείται από τη φύση του προβλήματος: οι χρηματοοικονομικές χρονοσειρές εμφανίζουν χρονικές εξαρτήσεις, αυτοσυσχετίσεις και επαναλαμβανόμενα πρότυπα διαφορετικών διάρκειων (π.χ. ημερήσιοι κύκλοι, εβδομαδιαίες τάσεις, εποχικότητα ή αντιδράσεις σε ειδήσεις) (βλ. πηγή 29). Τα LSTM, σε αντίθεση με τα απλά feed-forward δίκτυα, μπορούν να μάθουν να «θυμούνται» σημαντικές πληροφορίες από προηγούμενα χρονικά βήματα και να τις επανενεργοποιούν όταν αυτές είναι χρήσιμες για την πρόβλεψη. Αυτό τα καθιστά ανώτερα από παραδοσιακά μοντέλα σειράς χρόνου (ARIMA, κ.ά.) όταν υπάρχουν μη-γραμμικά μοτίβα και συνδυαστικές επιδράσεις πολλών χαρακτηριστικών. Επιπλέον, η χρήση πολλαπλών εισοδών διαφορετικών παραθύρων ενισχύει την ικανότητα του μοντέλου να συνδυάζει βραχυπρόθεσμες και μακροπρόθεσμες πληροφορίες. Για παράδειγμα, ο βραχυπρόθεσμος κλάδος μπορεί να αντιληφθεί ένα πρόσφατο momentum ή μια απότομη μεταστροφή τιμής, ενώ ο μακροπρόθεσμος καταγράφει τη γενικότερη τάση των τελευταίων δύο μηνών. Το μοντέλο, συνδυάζοντας αυτά τα δύο, μπορεί να αποφασίσει ότι π.χ. ένα σήμα αγοράς είναι πιο αξιόπιστο αν συμβαίνει εντός μιας καθιερωμένης ανοδικής τάσης (long-term branch) αλλά μετά από μια πρόσκαιρη διόρθωση (short-term branch). Χωρίς την πολυ-παραθυρική προσέγγιση, τέτοιου είδους συνδυασμοί θα ήταν δυσκολότερο να ανιχνευθούν. Τέλος, αξίζει να σημειωθεί ότι η χωριστή εκπαίδευση των LSTM ανά παράθυρο (αντί ενός ενιαίου LSTM 50 ημερών) επιβάλλει στο μοντέλο μια δομή μεροληψίας (inductive bias): το κάθε υποδίκτυο μαθαίνει ειδικά πρότυπα της συγκεκριμένης χρονικής κλίμακας, γεγονός που μπορεί να βελτιώσει την γενίκευση και να μειώσει την πολυπλοκότητα μάθησης (κάθε LSTM επικεντρώνεται στο «δικό του» ορίζοντα). Μετά την εκπαίδευση, το προκύπτον μοντέλο LSTM αξιολογήθηκε στο test set. Η τελική ακρίβεια στις μη-ορατές ημέρες κυμάνθηκε σε ικανοποιητικά επίπεδα (πάνω από το τυχαίο ~33% για τρεις τάξεις), επιβεβαιώνοντας ότι το μοντέλο απέκτησε κάποια προβλεπτική ικανότητα. Πέρα από την ακρίβεια, όμως, το ενδιαφέρον εστιάζεται στο κατά πόσο το μοντέλο λαμβάνει λογικές αποφάσεις (π.χ. αν προβλέπει «Buy» σε ανοδικές περιστάσεις και «Sell» πριν από πτώσεις) – κάτι που διερευνάται τόσο μέσω της προσομοίωσης στρατηγικής όσο και μέσω των μεθόδων εξηγησιμότητας που θα παρουσιαστούν.



Εικόνα 5.1: Αρχιτεκτονική του τελικού μοντέλου LSTM.

5.5 Προσομοίωση Στρατηγικής Συναλλαγών

Για να εκτιμηθεί η πρακτική αξία των προβλέψεων του μοντέλου, τα αποτελέσματά του μεταφράστηκαν σε αποφάσεις trading και υποβλήθηκαν σε προσομοίωση (backtesting) σε ιστορικά δεδομένα. Με αυτόν τον τρόπο αξιολογούμε κατά πόσο το σύστημα θα απέδιδε κέρδη ή ζημίες εάν εφαρμοζόταν σε πραγματικό trading. Η προσομοίωση έγινε με ένα απλό σενάριο συναλλαγών και συγκεκριμένες παραδοχές, ώστε να είναι ευδιάκριτη η συμβολή του μοντέλου χωρίς περίπλοκους εξωγενείς παράγοντες.

5.5.1 Λογική Σεναρίου και Παραδοχές

Ορίσαμε ένα υποθετικό επενδυτικό σενάριο ως εξής: Ξεκινάμε την πρώτη ημέρα της περιόδου δοκιμής με ένα αρχικό κεφάλαιο (π.χ. 100) και καμία μετοχή στο χαρτοφυλάκιό μας. Κάθε ημέρα, το μοντέλο μας παράγει μια πρόβλεψη δράσης – Buy, Hold ή Sell – βάσει των δεδομένων έως εκείνη την ημέρα. Αυτή η πρόβλεψη μεταφράζεται σε απόφαση συναλλαγής με τους εξής κανόνες:

- **Αν το μοντέλο δώσει σήμα "Buy"** και δεν κατέχουμε ήδη μετοχές (έχουμε μόνο μετρητά διαθέσιμα), τότε αγοράζουμε τη μετοχή με όλο το διαθέσιμο κεφάλαιο. Δηλαδή, μετατρέπουμε το ρευστό σε μετοχές στην τρέχουσα τιμή. Για λόγους απλότητας, θεωρήσαμε ότι μπορούμε να αγοράσουμε κλασματικό αριθμό μετοχών έτσι ώστε να επενδύσουμε ακριβώς όλο το κεφάλαιο (π.χ. αν έχουμε 100 και η τιμή είναι 150, αγοράζουμε 0.667 μετοχές). Στην πράξη, θα μπορούσε να τεθεί περιορισμός ακέραιων μετοχών, αλλά επιτρέποντας κλάσματα εξαλείφουμε το τεχνικό αυτό ζήτημα και εστιάζουμε στην ουσία της στρατηγικής. Σημειώνεται επίσης ότι δεν επιτρέπεται μόχλευση ή short selling – το μοντέλο μόνο αποφασίζει πότε θα είμαστε μέσα στην αγορά (θέση αγοράς) ή εκτός αγοράς (σε μετρητά).
- **Αν το μοντέλο δώσει σήμα "Sell"** και κατέχουμε μετοχές, τότε ρευστοποιούμε τη θέση μας – πουλάμε όλες τις μετοχές που διαθέτουμε, μετατρέποντάς τες στην τρέχουσα τιμή σε μετρητά. Μετά την πώληση, το χαρτοφυλάκιο είναι εξ' ολοκλήρου σε μορφή ρευστού.
- **Αν το μοντέλο δώσει "Hold"**, δεν πραγματοποιείται καμία συναλλαγή εκείνη την ημέρα. Δηλαδή, διατηρούμε την όποια θέση είχαμε από προηγούμενες ημέρες (είτε μετρητά είτε μετοχές).

Οι παραπάνω κανόνες συνιστούν μια αυτοματοποιημένη στρατηγική όπου οι αποφάσεις του μοντέλου ακολουθούνται πιστά. Για να απλοποιηθεί περαιτέρω η

προσομοίωση και να απομονωθεί η επίδραση των προβλέψεων, θέσαμε ορισμένες επιπλέον παραδοχές: (1) Δεν υπάρχουν κόστη συναλλαγών (μηδενικές προμήθειες broker και μηδενική ολίσθηση τιμής), ώστε να μην μειώνονται τα κέρδη από εξωγενείς παράγοντες. (2) Οι συναλλαγές εκτελούνται στις τιμές κλεισίματος της ημέρας του σήματος – πρακτικά αυτό σημαίνει ότι αν λάβουμε σήμα Buy τη μέρα t, θεωρούμε ότι αγοράζουμε στην τιμή κλεισίματος της t (σαν να ενεργούσαμε στο τέλος της συνεδρίας), και αντίστοιχα για Sell. (3) Δεν επανεπενδύονται μερίσματα ή άλλα κέρδη, δεδομένου ότι επικεντρωνόμαστε αποκλειστικά στα κέρδη/ζημίες από τη διαφορά τιμής. (4) Η στρατηγική δεν επιτρέπει μερική κατοχή θέσης κατά διαστήματα – είτε είμαστε 100% μέσα στη μετοχή είτε 100% εκτός (σε μετρητά). Αυτό αντικατοπτρίζει ένα απόλυτο σενάριο market timing, διευκολύνοντας την αποτίμηση των σημάτων Buy/Sell.

5.5.2 Εκτέλεση Προσομοίωσης και Αποτελέσματα

Με βάση τις παραπάνω παραδοχές, εκτελέστηκε η προσομοίωση στο σύνολο δοκιμής. Η διαδικασία υλοποιήθηκε βηματικά: ξεκινώντας από την πρώτη ημέρα του test set, ενημερωνόταν καθημερινά η πρόβλεψη του μοντέλου και εφαρμόζονταν οι κανόνες συναλλαγών. Τα μεγέθη που παρακολουθήσαμε κατά την εξέλιξη της προσομοίωσης ήταν:

- το υπόλοιπο μετρητών στο χαρτοφυλάκιο,
- ο αριθμός μετοχών που κατέχουμε,
- και η συνολική αξία χαρτοφυλακίου (Cash + αξία μετοχών) σε κάθε ημέρα.

Αρχικά (ημέρα έναρξης test), η αξία χαρτοφυλακίου ισούται με το αρχικό κεφάλαιο (καθώς δεν υπάρχουν μετοχές). Καθώς προχωρά η προσομοίωση, κάθε φορά που εκτελείται μια αγορά ή πώληση, τα μεγέθη αυτά ενημερώνονται αντίστοιχα. Για παράδειγμα, σε ένα σήμα Buy: το cash μειώνεται στο 0, ο αριθμός μετοχών αυξάνεται (ίσως κλασματικά) και η αξία του χαρτοφυλακίου συνεχίζει να μεταβάλλεται σύμφωνα με την τιμή της μετοχής τις επόμενες ημέρες. Σε σήμα Sell: το cash αυξάνεται κατά την αξία πώλησης, οι μετοχές μηδενίζονται και από εκείνη τη στιγμή η αξία χαρτοφυλακίου παύει να επηρεάζεται από διακυμάνσεις τιμής (μέχρι επόμενη αγορά). Όλες οι μεταβολές αυτές καταγράφηκαν ημερησίως.

Προκειμένου να αξιολογηθεί η αποδοτικότητα της στρατηγικής του μοντέλου, συγκρίναμε την απόδοσή της με μια απλούστατη εναλλακτική: τη στρατηγική **αγοράς-και-διακράτησης (buy & hold)** της AAPL. Για τη συγκεκριμένη σύγκριση, υπολογίσαμε τι θα είχε συμβεί αν το αρχικό κεφάλαιο επενδυόταν εξ ολοκλήρου σε μετοχές AAPL στην αρχή της περιόδου test και απλώς κρατιόταν μέχρι το τέλος

χωρίς καμία ενδιάμεση συναλλαγή. Αυτή η παθητική στρατηγική αποτελεί ένα βασικό σημείο αναφοράς – οποιοδήποτε ενεργό σύστημα πρόβλεψης θα πρέπει, ιδανικά, να την ξεπερνά σε επιδόσεις (ειδάλλως δεν δικαιολογείται η πολυπλοκότητα και το ρίσκο των ενεργών συναλλαγών). Στην πράξη, για την buy & hold, υπολογίζαμε καθημερινά την αξία του χαρτοφυλακίου αν διακρατούσαμε σταθερό αριθμό μετοχών (όπως αγοράστηκαν εξαρχής) – αυτό είναι εύκολο, καθώς απλώς ακολουθεί αναλογικά την τιμή κλεισίματος της μετοχής.

Με την ολοκλήρωση της προσομοίωσης, προέκυψαν ορισμένοι σημαντικοί δείκτες απόδοσης:

- το τελικό συνολικό κέρδος/ζημία (% μεταβολή της αξίας χαρτοφυλακίου από την αρχή ως το τέλος του test),
- η συνολική απόδοση buy & hold για το ίδιο διάστημα (επί τοις %),
- ο μέγιστος βαθμός κάμψης (**Max Drawdown**) της στρατηγικής σε σύγκριση με του buy & hold,
- ο ρυθμός επιτυχίας συναλλαγών (**win rate**), δηλαδή το ποσοστό των συναλλαγών (Buy → Sell κύκλοι) που κατέληξαν κερδοφόρες,
- ο αριθμός των πραγματοποιημένων συναλλαγών,
- και ένας προσεγγιστικός **συντελεστής Sharpe** της στρατηγικής.

Ειδικά ο μέγιστος drawdown αποτελεί κρίσιμο μέτρο ρίσκου: υπολογίζεται ως η μεγαλύτερη ποσοστιαία πτώση της αξίας του χαρτοφυλακίου από ένα τοπικό μέγιστο μέχρι το επόμενο τοπικό ελάχιστο. Ένας χαμηλότερος drawdown σε σχέση με το buy & hold υποδηλώνει ότι η στρατηγική του μοντέλου προστάτευσε καλύτερα το κεφάλαιο σε πτωτικές φάσεις. Ο συντελεστής Sharpe υπολογίστηκε με την κλασική προσέγγιση (μέση ημερήσια απόδοση μείον χωρίς ρίσκο, διά στανταρ απόκλιση ημερήσιων αποδόσεων), δίνοντας μια ένδειξη της απόδοσης ανά μονάδα κινδύνου.

Τα αποτελέσματα της προσομοίωσης αποκάλυψαν την πρακτική αξία των σημάτων του μοντέλου. Σε περιόδους όπου το μοντέλο προέβλεψε επιτυχώς πτωτικές τάσεις (σήματα Sell πριν από σημαντικές πτώσεις), η στρατηγική κατάφερε να προστατεύσει το κεφάλαιο (μένοντας σε μετρητά) την ώρα που η buy & hold υπέστη ζημίες – μειώνοντας αισθητά τον drawdown. Αντίστοιχα, σε ανοδικές φάσεις που το μοντέλο εισήλθε έγκαιρα (σήματα Buy), το χαρτοφυλάκιο επωφελήθηκε από την άνοδο σχεδόν όσο και η παθητική κατοχή. Συνολικά, η στρατηγική LSTM πέτυχε μια τελική απόδοση εφάμιλλη ή ανώτερη της buy & hold, με μικρότερη όμως μεταβλητότητα και drawdowns, κάτι που είναι ιδιαίτερα επιθυμητό για έναν επενδυτή. Ο ρυθμός επιτυχίας των συναλλαγών ήταν άνω του 50%, ένδειξη ότι το

μοντέλο μπορούσε να προβλέπει περισσότερες σωστές κινήσεις απ' ό,τι λάθος. Φυσικά, ο ακριβής αριθμός των συναλλαγών επηρεάζει την πρακτική βιωσιμότητα – σε πολύ συχνά «μπες-βγες» θα υπήρχαν κόστη συναλλαγών. Στην περίπτωσή μας, το μοντέλο τείνει να δίνει σήματα Hold σε ουδέτερες περιόδους, οπότε ο αριθμός συναλλαγών διατηρήθηκε σε λογικά επίπεδα.

Συμπερασματικά, η προσομοίωση ανέδειξε ότι το προτεινόμενο σύστημα μπορεί να λειτουργήσει ως μια αυτοματοποιημένη στρατηγική trading. Αν και η παρούσα μελέτη δεν έλαβε υπόψη όλες τις λεπτομέρειες (π.χ. προμήθειες, impact μεγάλου όγκου στην αγορά), τα αποτελέσματα καταδεικνύουν ότι οι προβλέψεις Buy/Sell του μοντέλου έχουν ενεργή αξία: μπορούν να βελτιώσουν το προφίλ ρίσκου-απόδοσης σε σχέση με μια παθητική προσέγγιση. Το επόμενο βήμα είναι να κατανοήσουμε πώς και γιατί το μοντέλο λαμβάνει αυτές τις αποφάσεις – κάτι που επιτυγχάνεται με τις μεθόδους εξηγησιμότητας που ακολουθούν.

5.6 Εξηγησιμότητα Μοντέλου (XAI)

Προκειμένου να διεισδύσουμε στους «μαύρους κουτιούς» αποφάσεων του προγνωστικού μας μοντέλου, εφαρμόσαμε ένα σύνολο από σύγχρονες τεχνικές Εξηγήσιμης Τεχνητής Νοημοσύνης (Explainable Artificial Intelligence – XAI). Συγκεκριμένα, χρησιμοποιήθηκαν πέντε διαφορετικές μέθοδοι: KernelSHAP, LIME, DeepLIFT, Integrated Gradients και Permutation Feature Importance. Κάθε μία από αυτές παρέχει μια ξεχωριστή οπτική ως προς τη σημασία των χαρακτηριστικών και τη λογική των προβλέψεων, είτε σε τοπικό επίπεδο (για συγκεκριμένες χρονικές στιγμές) είτε σε συνολικό επίπεδο μοντέλου. Αυτή η προσέγγιση πολλαπλών μεθόδων ενισχύει την αξιοπιστία των εξηγήσεων, επιτρέποντας τη διασταύρωση των ευρημάτων και την καλύτερη κατανόηση της συμπεριφοράς του μοντέλου (βλ. πηγή 94). Ακολουθεί αναλυτική περιγραφή της υλοποίησης και του ρόλου κάθε μεθόδου στο πλαίσιο του πολυ-εισόδου μοντέλου LSTM.

5.6.1 KernelSHAP

Η μέθοδος SHAP (SHapley Additive exPlanations) βασίζεται στη θεωρία παιγνίων και αποδίδει σε κάθε χαρακτηριστικό μια τιμή – την SHAP value – που αντιπροσωπεύει τη συνεισφορά του χαρακτηριστικού στη διαφορά της πρόβλεψης του μοντέλου από μια βασική τιμή (baseline) (βλ. πηγή 38). Δεδομένου ότι το LSTM μοντέλο μας δεν ανήκει σε μια κατηγορία με εύκολους αναλυτικούς εξηγητές (π.χ. δεν είναι δέντρο απόφασης ώστε να χρησιμοποιηθεί TreeSHAP), επιλέχθηκε η **KernelSHAP**.

Πρόκειται για μια υλοποίηση model-agnostic (ανεξάρτητη από το είδος του μοντέλου) που προσεγγίζει τις Shapley values μέσω δειγματοληψίας Monte Carlo.

Στην περίπτωσή μας, το μοντέλο έχει πολυδιάστατη είσοδο (3 διαφορετικές ακολουθίες: 10, 30, 50 ημερών). Η απευθείας εφαρμογή του SHAP σε ένα τέτοιο μοντέλο δεν είναι ευθύγραμμη, γι' αυτό χρειάστηκε να επινοηθεί μια κατάλληλη στρατηγική προσαρμογής εισόδου:

- **Ενοποίηση Εισόδων (Flattening):** Κατασκευάσαμε ένα ειδικό wrapper γύρω από το εκπαιδευμένο LSTM, έτσι ώστε στα μάτια του SHAP το μοντέλο να δέχεται μονοδιάστατο input (ένα μεγάλο διάνυσμα). Συγκεκριμένα, "επιπεδοποιήσαμε" τις τρεις εισόδους ακολουθιών σε ένα ενιαίο διάνυσμα μεγέθους 50×F (όπου F ο αριθμός χαρακτηριστικών). Αυτό επιτεύχθηκε απλώς με τη συνεκτόπιση των τριών ακολουθιών: πρώτα τα δεδομένα του σύντομου παραθύρου (10 ημερών), μετά του μεσαίου (30 ημερών) και τέλος του μεγάλου (50 ημερών). Με άλλα λόγια, κάθε δείγμα εισόδου αντιπροσωπεύεται ως μια ακολουθία 50 ημερών (η μακρά) που εμπεριέχει μέσα της τις μικρότερες – ουσιαστικά η υπο-ακολουθία των τελευταίων 10 ημερών χρησιμοποιείται για το short-window, των τελευταίων 30 για το medium, και ολόκληρη για το long.
- **Συνάρτηση Πρόβλεψης Wrapper:** Ορίσαμε μια συνάρτηση predict_function(X_flat) που παίρνει ως είσοδο ένα τέτοιο επίπεδο 2D array (δείγματα × χαρακτηριστικά) και μέσα της κάνει το ανάστροφο μετασχηματισμό: δηλαδή, ξαναχωρίζει το flat vector σε τρία μέρη (που αντιστοιχούν στις 10, 30, 50 ημέρες) και τα μετασχηματίζει πίσω σε 3D τείχη διαστάσεων (samples, 10, F), (samples, 30, F), (samples, 50, F). Στη συνέχεια, καλεί το πραγματικό μοντέλο LSTM με αυτές τις τρεις εισόδους και επιστρέφει τις προβλέψεις του. Έτσι, από την οπτική της SHAP, έχουμε ένα «κανονικό» black-box μοντέλο που δέχεται απλά διανύσματα σταθερού μήκους και επιστρέφει πιθανότητες.
- **Baseline Κατανομή:** Για την εκτίμηση των SHAP values, το KernelSHAP χρειάζεται ένα σύνολο background data – ένα αντιπροσωπευτικό υποσύνολο εισόδων που χρησιμοποιείται ως σημείο αναφοράς για το «μέσο αποτέλεσμα» του μοντέλου. Επιλέξαμε τυχαία 50 ημερήσια δείγματα από το σετ επικύρωσης ως baseline (αυτά τα δείγματα καλύπτουν ποικιλία αγοραίων καταστάσεων). Οι τιμές των χαρακτηριστικών αυτών των δειγμάτων χρησιμοποιούνται από τον αλγόριθμο για να προσομοιώσει την απουσία ενός χαρακτηριστικού: όταν «σβήνει» ένα χαρακτηριστικό κατά τον υπολογισμό της συνεισφοράς του, ουσιαστικά το αντικαθιστά με τη διανομή που αυτό

παίρνει στο baseline set.

- **Υπολογισμός Shapley τιμών:** Με το wrapper μοντέλο και το baseline έτοιμα, δημιουργήσαμε ένα KernelExplainer και υπολογίσαμε τις SHAP values για ένα σύνολο επιλεγμένων δειγμάτων (π.χ. 5 τυχαίων ημερών από το test set). Ο αλγόριθμος δοκιμάζει διάφορους συνδυασμούς παρουσίας/απουσίας χαρακτηριστικών (μέσω δειγματοληψίας πολλών δυαδικών μασκών) και παρατηρεί πόσο αλλάζει η πρόβλεψη του μοντέλου. Το αποτέλεσμα είναι για κάθε δείγμα και για κάθε χαρακτηριστικό μια τιμή – θετική αν το χαρακτηριστικό αυτό συνεισφέρει θετικά στην πρόβλεψη (αυξάνει την πιθανότητα της προβλεφθείσας κλάσης σε σχέση με το baseline), ή αρνητική αν συνεισφέρει αρνητικά.
- **Απεικόνιση και Ερμηνεία:** Οι SHAP values εξάγονται τελικά σε μορφή που αντιστοιχεί στο flattened διάνυσμα εισόδου. Για να τις ερμηνεύσουμε, ανακατασκευάσαμε αντιστοιχίσεις μεταξύ κάθε θέσης του flat διανύσματος και των αρχικών χαρακτηριστικών (π.χ. "RSI της ημέρας t-3 στο short window", "τιμή κλεισίματος της ημέρας t-10 στο long window", κ.ο.κ.). Έτσι μπορέσαμε να δημιουργήσουμε διαγράμματα σημαντικότητας: π.χ. bar plots με τις top-10 σημαντικότερες μεταβλητές για μια συγκεκριμένη πρόβλεψη ή summary plots που δείχνουν τη συνολική επίδραση κάθε χαρακτηριστικού σε όλες τις προβλέψεις (διασπορά τιμών SHAP). Οι SHAP τιμές έδωσαν μια συνεπή εξήγηση των αποφάσεων: για παράδειγμα, διαπιστώθηκε ότι σε ένα σήμα «Buy» του μοντέλου, οι θετικότερες συνεισφορές προήλθαν από ενδείξεις όπως ένας χαμηλός RSI και υψηλό momentum τιμών λίγες ημέρες πριν (όπως ακριβώς θα σκεφτόταν και ένας τεχνικός αναλυτής), ενώ χαρακτηριστικά όπως μια πρόσφατη πτώση στον όγκο είχαν αρνητική συνεισφορά (λειτούργησαν αποτρεπτικά στο σήμα αγοράς).

Η KernelSHAP, όντας μέθοδος ανεξάρτητη του μοντέλου (model-agnostic), έχει το πλεονέκτημα ότι δεν απαιτεί καμία γραμμική προσέγγιση ή εσωτερικές παραδοχές – θεωρητικά καταλήγει στις μοναδικές δίκαιες τιμές κατανομής σημασίας (σύμφωνα με τις ιδιότητες Shapley) (βλ. πηγή 38). Ωστόσο, είναι υπολογιστικά ακριβή. Στην υλοποίησή μας, για κάθε δείγμα εξήγησης δημιουργήθηκαν εκατοντάδες μοντέλα-προβλέψεις (για διάφορους συνδυασμούς χαρακτηριστικών), κάτι που απαιτεί σημαντικό χρόνο. Παρόλα αυτά, απέδωσε ιδιαίτερα πλούσια πληροφόρηση: όχι μόνο ποια χαρακτηριστικά είναι σημαντικά, αλλά και προς ποια κατεύθυνση επηρεάζουν το output (θετικά/αρνητικά), καθώς και πιθανές αλληλεπιδράσεις (μέσω plots, π.χ. πώς η σημασία του Volume αλλάζει ανάλογα με την τιμή του RSI). Συνολικά, το KernelSHAP μάς έδωσε μια σφαιρική εικόνα της συμπεριφοράς του μοντέλου, τόσο σε τοπικό επίπεδο (εξήγηση συγκεκριμένης πρόβλεψης) όσο και σε

παγκόσμιο επίπεδο (συγκεντρωτική σημαντικότητα χαρακτηριστικών).

5.6.2 LIME (Local Interpretable Model-Agnostic Explanations)

Η μέθοδος LIME (Local Interpretable Model-Agnostic Explanations) στοχεύει στην παροχή τοπικά πιστών εξηγήσεων για μεμονωμένες προβλέψεις, προσαρμόζοντας ένα απλό, ερμηνεύσιμο μοντέλο (συνήθως γραμμικό) στην περιοχή γύρω από την είσοδο ενδιαφέροντος (βλ. πηγή 38). Εν ολίγοις, ενώ η SHAP υπολογίζει ακριβείς κατανομές συνεισφορών με βάση όλες τις δυνατές αφαιρέσεις χαρακτηριστικών, η LIME ακολουθεί μια προσέγγιση τοπικής παρεμβολής: παράγει πολλαπλά τεχνητά δεδομένα κοντά στο δείγμα που θέλουμε να εξηγήσουμε (με μικρές τυχαίες μεταβολές στα χαρακτηριστικά του) και βλέπει πώς αλλάζει η πρόβλεψη του μοντέλου. Στη συνέχεια, εκπαιδεύει ένα απλό γραμμικό μοντέλο πάνω σε αυτά τα δείγματα (με βάρη που δίνουν έμφαση στα πιο κοντινά στο αρχικό), ώστε αυτό να προσομοιώσει τη συμπεριφορά του πολύπλοκου μοντέλου τοπικά. Τα βάρη αυτού του απλού μοντέλου – ουσιαστικά ένας γραμμικός συνδυασμός χαρακτηριστικών – χρησιμεύουν ως εξήγηση: δείχνουν ποια χαρακτηριστικά επηρεάζουν και πόσο το αποτέλεσμα γύρω από την περίπτωση που μελετάμε.

Για την εφαρμογή της LIME στο δικό μας μοντέλο, αντιμετωπίσαμε την ίδια πρόκληση με πριν: πολυ-εισόδου μοντέλο. Η LIME όμως, σε αντίθεση με την SHAP, προσφέρεται μέσω έτοιμης βιβλιοθήκης (lime.lime_tabular) που αναμένει επίσης δισδιάστατες εισόδους (δείγματα × χαρακτηριστικά). Για να γεφυρώσουμε αυτό το χάσμα, ακολουθήσαμε παρόμοια τεχνική flattening: συνενώσαμε τα δεδομένα του 50ήμερου παραθύρου (το μεγαλύτερο) σε ένα μονοδιάστατο διάνυσμα (50×F χαρακτηριστικά). Ευτυχώς, εφόσον τα μικρότερα παράθυρα είναι υποσύνολα του 50ημέρου, μπορούμε να πούμε ότι μια είσοδος 50 ημερών περιέχει πλήρως την πληροφορία και των τριών.

Πρακτικά, για το δείγμα που θέλουμε να εξηγήσουμε με LIME, παίρνουμε τις τελευταίες 50 ημέρες δεδομένων του και το αντιστοιχίζουμε σε ένα flat vector. Στη συνέχεια υλοποιήσαμε μια συνάρτηση lstm_predict_proba(X_2d) για να ενσωματωθεί στη LIME, η οποία κάνει ακριβώς το ανάποδο: όταν η LIME της δίνει κάποιο τεχνητά τροποποιημένο 2D δείγμα (ίδια διάσταση 50×F), αυτή το αναδιαμορφώνει πίσω σε τρεις εισόδους (10×F, 30×F, 50×F) – συγκεκριμένα παίρνει από το 50ήμερο vector τις τελευταίες 10 ημέρες ως short input, τις τελευταίες 30 ως medium και όλες ως long – και έπειτα καλεί το πραγματικό μοντέλο να πάρει την πρόβλεψη πιθανότητας.

Με αυτό το wrapper, μπορέσαμε να αξιοποιήσουμε το LimeTabularExplainer:

επιλέξαμε ως χαρακτηριστικά του explainer τα ονόματα όλων των χαρακτηριστικών για κάθε ημέρα (όμοια με τα flat feature names του SHAP, π.χ. "Close (short:t-1)", "RSI (long:t-50)" κ.λπ.), ώστε η έξοδος να είναι ανθρωπίνως ερμηνεύσιμη. Για κάθε προς-εξήγηση δείγμα, η LIME γέννησε π.χ. 500 κοντινά δείγματα (μέσω τυχαίας προσθήκης μικρού θορύβου στα χαρακτηριστικά) και τα έστειλε στο wrapper μοντέλο για πρόβλεψη. Κατόπιν, προσαρμόστηκε ένα γραμμικό μοντέλο (τοπικός ανιχνευτής) σε αυτές τις παρατηρήσεις, βαρύνοντας περισσότερο τα δείγματα που βρίσκονται πιο κοντά στο αρχικό (μετρώντας απόσταση στο χώρο χαρακτηριστικών). Το αποτέλεσμα είναι ένας πίνακας με συντελεστές για κάποιες από τις πιο σημαντικές μεταβλητές.

Για παράδειγμα, μια τυπική έξοδο LIME για ένα συγκεκριμένο σήμα «Buy» του μοντέλου μπορεί να ήταν: "RSI (t-2) = 28 (συντελεστής +0.5)", "Close (t-1) vs SMA10 (t-1) = -2% (συν. +0.3)", "Volume (t-0) = χαμηλός (συν. -0.2)", που θα το ερμηνεύαμε ως εξής: η ύπαρξη χαμηλού RSI δύο μέρες πριν και μια τιμή κλεισίματος οριακά κάτω από τον βραχυπρόθεσμο ΜΚΟ την προηγούμενη μέρα ώθησαν το μοντέλο προς την απόφαση Buy, ενώ ο σχετικά χαμηλός όγκος την ημέρα της πρόβλεψης λειτούργησε ελαφρώς αποτρεπτικά. Τα αποτελέσματα της LIME παρουσιάζονται συνήθως σε έναν απλό πίνακα ή ράβδους σημαντικότητας (βλ. πηγή 106). Παρέχουν μια προσεγγιστική, ανθρώπινα κατανοητή εξήγηση για τη συγκεκριμένη πρόβλεψη.

Σε σύγκριση με το SHAP, η LIME είναι πιο προσεγγιστική (οι συντελεστές της δεν έχουν αυστηρή εγγύηση ορθότητας, εξαρτώνται από την τοπική γραμμικοποίηση), αλλά είναι αρκετά γρήγορη και ευέλικτη. Στην περίπτωσή μας, επιβεβαίωσε πολλά από τα μοτίβα που υπονοήθηκαν από τις SHAP values: π.χ. ότι για μια πρόβλεψη Buy, συνήθως ένας συνδυασμός παραγόντων όπως «τιμή κάτω από ΜΚΟ + χαμηλό RSI» έχουν θετική βαρύτητα. Επίσης, επειδή η LIME μπορεί να περιορίσει τον αριθμό χαρακτηριστικών στην εξήγηση (π.χ. εμφάνιση των 5 πιο σημαντικών), βοήθησε στο να συμπυκνωθεί η εξήγηση σε λίγες κύριες αιτίες, κάτι που είναι χρήσιμο όταν θέλουμε να παρουσιάσουμε αποτελέσματα σε μη-τεχνικούς ενδιαφερόμενους (βλ. πηγή 108). Συμπερασματικά, η LIME λειτούργησε ως ένας μεγεθυντικός φακός στη γειτονιά κάθε απόφασης του μοντέλου, παρέχοντας γραμμικές προσεγγίσεις εύκολες στην αφήγηση («λόγω των X, Ψ, Z, το μοντέλο προέβλεψε αγορά»). Αν και δεν έχει την ακριβή θεωρητική θεμελίωση του SHAP, η πρακτική της χρησιμότητα στην ερμηνεία συγκεκριμένων περιπτώσεων ήταν ανεκτίμητη.

5.6.3 DeepLIFT (Deep Learning Important FeaTures)

Το DeepLIFT (Deep Learning Important FeaTures) είναι μια μέθοδος εξηγησιμότητας ειδικά σχεδιασμένη για νευρωνικά δίκτυα, η οποία αποδίδει βαθμολογίες συνεισφοράς σε κάθε είσοδο συγκρίνοντας την ενεργοποίηση του δικτύου σε μια κανονική είσοδο με την ενεργοποίησή του σε μια βασική (baseline) είσοδο (βλ. πηγή 61). Μοιάζει εννοιολογικά με τα ολοκληρωμένα gradients (που θα δούμε παρακάτω), αλλά χρησιμοποιεί μια διαφορετική λογική κατανομής διαφορών αντί παραγώγων. Συγκεκριμένα, για κάθε νευρώνα του δικτύου, το DeepLIFT υπολογίζει τη διαφορά της ενεργοποίησής του ανάμεσα στην κανονική είσοδο και τη baseline είσοδο, και εν συνεχεία κατανέμει αυτή τη διαφορά προς τα πίσω στους νευρώνες της προηγούμενης στοιβάδας αναλογικά με τις μεταβολές που αυτοί προκάλεσαν. Επαναλαμβάνοντας αυτή τη διαδικασία (chain rule) από την έξοδο προς τις εισόδους, το DeepLIFT καταλήγει να δίνει σε κάθε αρχικό χαρακτηριστικό μια τιμή σημασίας που δείχνει πόσο άλλαξε η τελική πρόβλεψη εξαιτίας της παρουσίας αυτού του χαρακτηριστικού σε σχέση με τη baseline.

Στην υλοποίησή μας, χρησιμοποιήσαμε ως baseline μια ουδέτερη είσοδο που αντιπροσωπεύει την «απουσία σήματος». Συνήθης επιλογή είναι ένα διάνυσμα μηδενικών ή ο μέσος όρος των χαρακτηριστικών. Επιλέξαμε baseline = 0 για όλες τις εισόδους (πρακτικά, μια ακολουθία ημερών όπου όλα τα χαρακτηριστικά είναι μηδενισμένα μετά την κανονικοποίηση), κάτι που σημαίνει ότι συγκρίνουμε την κανονική πρόβλεψη του LSTM με την πρόβλεψη που θα έδινε αν δεν υπήρχε καθόλου πληροφορία (ένα υποθετικό ουδέτερο input). Για κάθε δείγμα προς εξήγηση, τρέξαμε τον αλγόριθμο DeepLIFT στο μοντέλο: αυτό πραγματοποιήθηκε με ένα custom script (deep_lift_explanation.py) που εκμεταλλεύεται δυνατότητες back-propagation στην πλατφόρμα TensorFlow. Ουσιαστικά, παράχθηκαν οι DeepLIFT scores για κάθε χαρακτηριστικό κάθε χρονικής στιγμής εισόδου.

Ερμηνεύοντας τα αποτελέσματα, οι θετικές τιμές DeepLIFT για ένα χαρακτηριστικό σημαίνουν ότι η παρουσία του (σε σχέση με baseline) αυξάνει την πιθανότητα της προβλεφθείσας κλάσης, ενώ οι αρνητικές ότι την μειώνουν. Για παράδειγμα, αν σε ένα σήμα Sell του μοντέλου, ο δείκτης MACD του προηγούμενου μήνα έλαβε DeepLIFT score -0.8, σημαίνει ότι ένα υψηλό MACD συγκριτικά με το baseline συνέβαλε αρνητικά στην απόφαση Sell (δηλ. απέτρεπε το μοντέλο από το να πει Sell, κάτι αναμενόμενο αφού υψηλό MACD συνήθως συνδέεται με ανοδικό momentum). Το άθροισμα όλων των DeepLIFT scores ισούται με τη διαφορά στην έξοδο του μοντέλου μεταξύ πραγματικής εισόδου και baseline (ιδιότητα additivity),

οπότε επιβεβαιώνεται ότι όλες οι συνεισφορές μαζί εξηγούν πλήρως την απόφαση.

Λόγω της πολυ-εισόδου φύσης του δικτύου, έπρεπε να τροποποιήσουμε ελαφρώς την εφαρμογή. Χρησιμοποιήσαμε baseline και για τις τρεις εισόδους (τρία μηδενικά tensors αντίστοιχου σχήματος: 10×F, 30×F, 50×F με όλα τα στοιχεία =0), και εφαρμόσαμε τον αλγόριθμο ώστε να πάρουμε τρία σύνολα scores (για short, medium, long input). Στη συνέχεια τα συγχωνεύσαμε κατάλληλα για να έχουμε ενιαίους δείκτες ανά χαρακτηριστικό και χρονικό βήμα, παρόμοια με τους πίνακες SHAP.

Ένα πλεονέκτημα του DeepLIFT είναι ότι δεν υποφέρει από το πρόβλημα των μηδενικών παραγώγων που μπορεί να έχουν οι απλές μέθοδοι βασισμένες σε παράγωγο (π.χ. αν η συνάρτηση ενεργοποίησης είναι flat σε ένα σημείο, η απλή παραγώγισή της δίνει 0 ενώ η συμβολή της μπορεί να μην είναι μηδενική) (βλ. πηγή 61). Το DeepLIFT χρησιμοποιεί «reference activations» κι επομένως αποφεύγει τέτοια θέματα. Σε σχέση με την SHAP, είναι πολύ πιο υπολογιστικά αποδοτικό – μια διέλευση του αλγορίθμου backprop έχει κόστος παρόμοιο με μια κανονική πρόβλεψη, ενώ το SHAP απαιτεί εκατοντάδες προβλέψεις για ένα δείγμα. Αυτό μας επέτρεψε να υπολογίσουμε DeepLIFT scores για πολλαπλά σημεία στο test set και να αποκτήσουμε μια γενική εικόνα των σημαντικότερων χαρακτηριστικών.

Στις αναλύσεις μας, το DeepLIFT ανέδειξε ξεκάθαρα ορισμένες συμφωνίες με τους κανόνες σήμανσης: π.χ. σε περιπτώσεις όπου το μοντέλο προέβλεψε Buy και η ημέρα αυτή πληρούσε τον κανόνα RSI < 25, είδαμε τον δείκτη RSI (στο τρέχον ή προηγούμενες ημέρες) να έχει από τις υψηλότερες θετικές βαθμολογίες DeepLIFT, επαληθεύοντας ότι το LSTM πιθανώς «στηρίχθηκε» σε αυτό το σήμα. Επιπλέον, το DeepLIFT μας επέτρεψε να εντοπίσουμε ποιες χρονικές υστέρησεις ήταν πιο χρήσιμες: για παράδειγμα, σε ένα μακροπρόθεσμο branch 50 ημερών, οι βαθμολογίες μπορεί να ήταν ισχυρές κυρίως για τα πιο πρόσφατα 10-15 days, υποδηλώνοντας ότι τελικά το μοντέλο έδινε μεγαλύτερη βαρύτητα στις πρόσφατες κινήσεις παρά στις πολύ παλιές (εύρημα λογικό και αναμενόμενο). Συνολικά, το DeepLIFT παρείχε ακριβείς και εύλογες εξηγήσεις σε επίπεδο μοντέλου, διατηρώντας μάλιστα το πλεονέκτημα της πλήρους συνάθροισης (conservation – οι βαθμολογίες αθροίζουν στο output). Αποτελεί μια ισχυρή μέθοδο για network-specific interpretability, και λειτούργησε συμπληρωματικά με τις υπόλοιπες, επιβεβαιώνοντας από διαφορετική σκοπιά τα ευρήματα σημαντικότητας των χαρακτηριστικών.

5.6.4 Integrated Gradients (Ενοποιημένες Κλίσεις)

Οι Ολοκληρωμένες Κλίσεις (Integrated Gradients) είναι μια ακόμη μέθοδος ερμηνείας για νευρωνικά δίκτυα, που μοιράζεται τη βασική ιδέα σύγκρισης με ένα baseline, αλλά προσεγγίζει το πρόβλημα μέσω ολοκλήρωσης των τοπικών παραγώγων του μοντέλου (βλ. πηγή 61). Ενώ το DeepLIFT διανέμει άμεσα τις διαφορές ενεργοποίησης, οι ολοκληρωμένες κλίσεις υπολογίζουν την μέση κλίση του output ως προς κάθε χαρακτηριστικό καθώς η είσοδος μεταβάλλεται από το baseline στην πραγματική τιμή. Πρακτικά, για κάθε χαρακτηριστικό πραγματοποιείται μια «διαδρομή» από baseline -> input και υπολογίζεται το εμβαδόν κάτω από την καμπύλη της μερικής παραγώγου. Αυτή η ολοκλήρωση δίνει την συνολική συμβολή του χαρακτηριστικού στο output, λαμβάνοντας υπόψη ακόμη και μη-γραμμικές περιοχές όπου η παράγωγος μπορεί να αλλάζει.

Στην υλοποίησή μας, χρησιμοποιήσαμε και πάλι baseline = 0 (το διάνυσμα μηδενικών) για όλες τις εισόδους. Ο αριθμός των βημάτων ολοκλήρωσης ορίστηκε σε 50, δηλαδή το input αυξήθηκε βαθμιαία από 0% της πραγματικής του τιμής έως 100% (πραγματική τιμή) σε 50 ενδιάμεσα σημεία, και υπολογίστηκαν οι παράγωγοι στο κάθε σημείο. Οι παράγωγοι αυτές (gradient του output ως προς κάθε input feature) αθροίστηκαν και πολλαπλασιάστηκαν με το βήμα (1/50) ώστε να προκύψει η ολοκληρωμένη τιμή. Η διαδικασία αυτή εφαρμόστηκε σε κάθε δείγμα προς εξήγηση, παρέχοντας ένα vector Integrated Gradients που έχει ιδιότητες παρόμοιες με το DeepLIFT (ικανοποιεί και αυτό την ιδιότητα ότι οι συνεισφορές αθροίζουν στη διαφορά output) (βλ. πηγή 61).

Λόγω της παρουσίας των τριών input branches, ακολουθήσαμε πάλι την λογική υπολογισμού ξεχωριστών IG για κάθε κλάδο (short, medium, long) και έπειτα τα συνενώσαμε. Σημαντικό είναι ότι για την εύκολη υλοποίηση χρησιμοποιήθηκαν έτοιμες συναρτήσεις από βιβλιοθήκες (TensorFlow / Keras grads), με loops για τα 50 βήματα, που όμως επιταχύνθηκαν μέσω vectorization όπου ήταν δυνατό (χρησιμοποιήθηκε μάλιστα και progress bar για παρακολούθηση, μιας και η διαδικασία μπορεί να είναι σχετικά βαριά χωρίς GPU acceleration).

Οι ολοκληρωμένες κλίσεις έδωσαν αποτελέσματα πολύ κοντά στο DeepLIFT (όπως θεωρητικά αναμένεται σε καλά βαθμονομημένα μοντέλα – το IG συγκλίνει στο σωστό attribution αν το μοντέλο είναι κατάλληλα differentiable). Μάλιστα, όπου δοκιμάστηκε με διαφορετικά baseline (π.χ. baseline ως τον μέσο όρο των χαρακτηριστικών αντί για μηδενικό), οι διαφορές δεν ήταν σημαντικές, υποδηλώνοντας ότι το μηδενικό baseline μας ήταν επαρκώς μακριά από την

κατανομή ώστε να παρέχει μια ουδέτερη αφετηρία.

Ερμηνευτικά, μια θετική IG τιμή για ένα χαρακτηριστικό σημαίνει ότι καθώς το χαρακτηριστικό κινείται από 0 μέχρι την πραγματική του τιμή, αυξάνει το output (υπέρ της προβλεφθείσας κλάσης). Ένα παράδειγμα: σε μια πρόβλεψη «Sell», βρέθηκε ότι η τιμή κλεισίματος 5 ημερών πριν είχε από τις μεγαλύτερες θετικές IG τιμές για την απόφαση αυτή – κοιτώντας τα δεδομένα, πράγματι 5 ημέρες πριν το Sell υπήρχε μια τοπική κορυφή τιμής (υψηλή τιμή που μετά άρχισε να πέφτει). Αυτό δείχνει ότι το μοντέλο πιθανώς έδωσε βαρύτητα στο γεγονός ότι πριν μερικές ημέρες η τιμή βρισκόταν ψηλά και άρα υπήρχε χώρος για πτώση (σήμα πώλησης). Παράλληλα, ένας πολύ υψηλός όγκος λίγες ημέρες πριν το Sell είχε αρνητική IG συνεισφορά – πιθανώς επειδή ο πολύ υψηλός όγκος σε μια κορυφή μπορεί να σηματοδοτεί ευφορία, άρα η πτώση που ακολούθησε ήταν πιο αναμενόμενη.

Σε γενικές γραμμές, οι Integrated Gradients επιβεβαίωσαν τις τάσεις που εντοπίστηκαν με DeepLIFT και SHAP, παρέχοντας μια ακόμη επιβεβαίωση ότι τα χαρακτηριστικά που εντοπίσαμε ως σημαντικά όντως επηρεάζουν τις προβλέψεις. Η μέθοδος IG έχει το πλεονέκτημα της θεωρητικής εγκυρότητας (ικανοποιεί ιδιότητες linearity, symmetry κ.λπ. υπό προϋποθέσεις) και είναι σχετικά απλή στην υλοποίηση για feed-forward networks (βλ. πηγή 61). Για τα LSTM μας, δούλεψε εξίσου καλά, με την μόνη πολυπλοκότητα να έγκειται στην ενοποίηση των πολλαπλών σειρών εισόδου.

5.6.5 Permutation Feature Importance (Σημαντικότητα Χαρακτηριστικών με Μεταθέσεις)

Τελευταία, υπολογίσαμε τη συνολική σημαντικότητα χαρακτηριστικών μέσω μεταθετικής διαταραχής (Permutation Importance), η οποία είναι μια παγκόσμια, ανεξάρτητη του μοντέλου, μέθοδος ερμηνείας (βλ. πηγή 38). Η ιδέα πίσω από αυτήν είναι απλή: μετράμε πόσο χειροτερεύει η απόδοση του μοντέλου όταν σπάσουμε τη συσχέτιση ενός χαρακτηριστικού με την έξοδο. Πρακτικά, για κάθε χαρακτηριστικό ξεχωριστά, τυχαία ανακατεύουμε (shuffle) τις τιμές αυτού του χαρακτηριστικού σε όλο το test set – έτσι, οι τιμές του παραμένουν ίδιες αλλά αποσυνδέονται από τις σωστές χρονικές θέσεις/ημέρες. Στη συνέχεια, ξαναϋπολογίζουμε την απόδοση του μοντέλου (π.χ. την ακρίβεια πρόβλεψης) με το αλλοιωμένο αυτό dataset και βλέπουμε πόσο πέφτει σε σχέση με πριν. Αν η πτώση στην ακρίβεια είναι μεγάλη, σημαίνει ότι το συγκεκριμένο χαρακτηριστικό έπαιζε σημαντικό ρόλο (η τυχαία μετατόπισή του προκαλεί σύγχυση στο μοντέλο). Αντίθετα, αν η απόδοση σχεδόν δεν αλλάξει, το χαρακτηριστικό πιθανώς δεν ήταν κρίσιμο ή η πληροφορία του

περιέχεται πλεονασματικά σε άλλα χαρακτηριστικά.

Για να εφαρμόσουμε την τεχνική αυτή στο πλαίσιο του πολυ-εισόδου LSTM, αντιμετωπίσαμε το κάθε επιμέρους χαρακτηριστικό (π.χ. "Close price") σε κάθε χρονική υστέρηση ανεξάρτητα. Δηλαδή, θεωρήσαμε ως ξεχωριστές οντότητες "Close_t-0_short", "Close_t-1_short", ..., "Close_t-0_long", ..., "RSI_t-0_short", κ.ο.κ., και για κάθε μία από αυτές τις στήλες κάναμε τυχαίο permutation στις τιμές μεταξύ των δειγμάτων test. Έτσι, καταφέραμε να μετρήσουμε όχι απλά τη σημαντικότητα π.χ. του RSI γενικά, αλλά ακόμα και ποια ημέρα του παραθύρου του RSI είναι πιο κρίσιμη. Ωστόσο, για την παρουσίαση των αποτελεσμάτων είναι συνήθως πιο χρήσιμο να συνοψίσουμε ανά χαρακτηριστικό (π.χ. μπορούμε να πάρουμε τον μέσο όρο της επίπτωσης σε απόδοση από το permutation όλων των χρονικών υστερήσεων ενός χαρακτηριστικού).

Η μετρική απόδοσης που επιλέξαμε να παρακολουθούμε ήταν η ακρίβεια ταξινόμησης στο test set. Υπολογίσαμε πρώτα τη baseline accuracy του μοντέλου στο κανονικό test (π.χ. ~47%) και μετά, για κάθε χαρακτηριστικό, υπολογίσαμε ξανά την accuracy με το εν λόγω χαρακτηριστικό ανακατεμένο. Η διαφορά (baseline acc – permuted acc) δίνει τη μείωση ακρίβειας λόγω ανακάτεμα. Πράγματι, βρήκαμε ότι ορισμένα χαρακτηριστικά προκαλούν σημαντική πτώση στην απόδοση όταν ανακατευτούν: για παράδειγμα, η τυχαία αναδιάταξη των τιμών RSI σε όλες τις ημέρες του test μείωσε την ακρίβεια κατά αρκετά ποσοστιαίες μονάδες. Αντίστοιχα υψηλή επίπτωση είχαν ορισμένοι δείκτες τάσης (όπως οι τιμές του SMA) και κάποιες από τις ίδιες τις τιμές τιμών (ιδίως η τιμή κλεισίματος μερικών πρόσφατων ημερών). Αντίθετα, ορισμένα χαρακτηριστικά όπως ο όγκος σε απομακρυσμένες υστερήσεις ή ορισμένοι δευτερεύοντες δείκτες (π.χ. %B των Bollinger Bands) έδειξαν σχεδόν μηδενική επίπτωση – πράγμα που υποδηλώνει ότι είτε δεν αξιοποιούνται καθόλου από το μοντέλο, είτε η πληροφορία τους καλύπτεται πλήρως από άλλα (π.χ. το %B είναι συναρτώμενο του close price και του εύρους Bollinger, οπότε ίσως το μοντέλο ήδη είχε αυτές τις πληροφορίες).

Τα αποτελέσματα του Permutation Feature Importance συνόψισαν ωραία τα συμπεράσματα των προηγούμενων μεθόδων σε ένα συγκεντρωτικό ranking χαρακτηριστικών. Σε υψηλές θέσεις βρέθηκαν εκείνα τα features που και οι SHAP/DeepLIFT είχαν αναδείξει – όπως το RSI, οι κινητοί μέσοι και οι ίδιες οι τιμές – επιβεβαιώνοντας έτσι την αξία τους. Αυτή η μέθοδος, αν και αργή (καθώς απαιτεί επαναλαμβανόμενες αξιολογήσεις του μοντέλου στο πλήρες test set για κάθε χαρακτηριστικό), είναι conceptually απλή και εξηγεί άμεσα: «αν εξαφανίσω/χαλάσω αυτό το χαρακτηριστικό, το μοντέλο χάνει Χ% απόδοσης». Αυτό το μήνυμα είναι

ισχυρό και κατανοητό ακόμα και εκτός τεχνικού κοινού (βλ. πηγή 108).

Συμπερασματικά, μέσω ενός ολοκληρωμένου πλαισίου εξηγησιμότητας με τις παραπάνω μεθόδους, αποκτήσαμε τόσο τοπική κατανόηση (γιατί μια συγκεκριμένη πρόβλεψη ήταν Buy ή Sell) όσο και συνολική εικόνα (ποια χαρακτηριστικά γενικά θεωρούνται σημαντικά από το μοντέλο). Η χρήση πολλαπλών τεχνικών λειτουργεί διασταυρωτικά – τα ευρήματα της μιας επιβεβαιώνουν της άλλης – αυξάνοντας έτσι την εμπιστοσύνη μας ότι η συμπεριφορά του μοντέλου είναι σε μεγάλο βαθμό κατανοητή και ευθυγραμμισμένη με λογικές επενδυτικές αρχές. Αυτό ενισχύει την αξιοπιστία του συστήματος: ένας επενδυτής ή αναλυτής μπορεί όχι μόνο να λαμβάνει προβλέψεις, αλλά να βλέπει και εξηγήσεις όπως «το μοντέλο συνιστά Πώληση κυρίως επειδή ο RSI είναι υψηλός και η τιμή ξεπέρασε το άνω όριο Bollinger», καθιστώντας την πρόταση πιο πειστική και ελέγξιμη. Οι μέθοδοι XAI που ενσωματώσαμε αποτελούν αναπόσπαστο μέρος της μεθοδολογίας μας, καθώς γεφυρώνουν το χάσμα μεταξύ του αλγοριθμικού μοντέλου και της ανθρώπινης κατανόησης, σε ένα τόσο ευαίσθητο πεδίο όπως οι χρηματοοικονομικές αποφάσεις.

ΚΕΦΑΛΑΙΟ 6: ΠΕΙΡΑΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ

Η πειραματική διαδικασία απαρτίζεται από 2 διακριτά μέρη με καθένα να πληρεί διαφορετικό σκοπό.

Το πρώτο πείραμα γίνεται έχει στόχο να εξετάσει εάν το μοντέλο το οποίο έχει εκπαιδευτεί με συγκεκριμένη στρατηγική ερμηνεύεται επαρκώς από καθεμία από τις μεθόδους εξηγησιμότητας και ποιά είναι τα δυνατά και τα αδύναμα σημεία κάθε μεθόδου.

Το δεύτερο μέρος της πειραματικής διαδικασίας έχει ως στόχο εφόσον έχουμε κατανοήσει από το πρώτο πείραμα τον βαθμό ερμηνευσιμότητας του μοντέλου από τις μεθόδους εξηγησιμότητας σε artificial conditions, να προσπαθήσουμε να εξάγουμε γνώση από το επόμενο μοντέλο το οποίο έχει εκπαιδευτεί με σκοπό να μπορεί να παράγει κέρδος. Έτσι θα καταλάβουμε ποιοί από τους παράγοντες επηρεάζουν περισσότερο τις αποφάσεις του μοντέλου μας και οδηγούν σε μεγαλύτερη κερδοφορία.

6.1 Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης (Custom Labeling Strategy)

6.1.1 Περιγραφή Πειράματος

Στο πρώτο πείραμα, ο στόχος ήταν να εκπαιδευτεί το LSTM για να αναγνωρίζει και να προβλέπει σήματα συναλλαγών (Buy, Hold, Sell) βασισμένα σε μια προκαθορισμένη, rule-based στρατηγική σήμανσης. Αυτή η στρατηγική αναπτύχθηκε με γνώμονα την ενίσχυση της ερμηνευσιμότητας των αποφάσεων του μοντέλου, καθώς οι κανόνες βασίζονται σε συγκεκριμένους τεχνικούς δείκτες και συνθήκες αγοράς. Η μετοχή της Apple ($AAPL) χρησιμοποιήθηκε ως case study, με ιστορικά δεδομένα από το 2000 έως το πρόσφατο παρόν (την μέρα που έτρεξε ο κώδικας).

Η διαδικασία περιλάμβανε τα ακόλουθα στάδια:

1. **Συλλογή και Προεπεξεργασία Δεδομένων**: Συλλέχθηκαν ημερήσια δεδομένα OHLCV (Open, High, Low, Close, Volume) για την $AAPL από το Yahoo Finance. Υπολογίστηκε ένα ευρύ φάσμα τεχνικών δεικτών (π.χ. RSI, MACD, SMA, Bollinger Bands, OBV) και ενσωματώθηκαν δείκτες ευρύτερου

πλαισίου (π.χ. S&P 500, κλαδικοί δείκτες). Όλα τα χαρακτηριστικά κανονικοποιήθηκαν σε κλίμακα χρησιμοποιώντας Min-Max Scaler. Τα δεδομένα μετασχηματίστηκαν σε ακολουθίες σταθερού μήκους (10, 30, 50 ημερών) για να τροφοδοτηθούν στο πολυ-εισοδικό LSTM μοντέλο, διασφαλίζοντας την χρονική ακεραιότητα (διαχωρισμός δεδομένων train/validation/test με χρονολογική σειρά).

2. **Εφαρμογή Προσαρμοσμένης Στρατηγικής Σήμανσης (Custom Labeling)**: Τα labels Buy, Hold, Sell δημιουργήθηκαν χρησιμοποιώντας μια μεθοδολογία βασισμένη σε κανόνες οι οποίοι φαίνονται παρακάτω με προτεραιότητες.

Labeling κανόνες για το label **Buy:**

RSI < 30 &
Close <= Bollinger Lower &
Momentum > 0 &
Volume > Volume SMA10

Ή

SMA10 < SMA50 &
MACD < 0 &
OBV diff < 0

Ή

Close < SMA50 &
SP500 Daily Return < -0.001 &
Sector Daily Return < -0.001

Ή

Momentum < 0 &
RSI < 50 &
VIX Close > 25

Labeling κανόνες για το label **Sell:**

RSI > 70 &
Close >= Bollinger Upper &
Momentum < 0

Ή

SMA10 < SMA50 &
MACD < 0 &
OBV diff < 0

Ή

Close < SMA50 &
SP500 Daily Return < -0.001 &
Sector Daily Return < -0.001

Ή

Momentum < 0 &
RSI < 50 &
VIX Close > 25

Ή

SMA10 < SMA30 &
Daily Return < 0

Labeling κανόνες για το label **Hold**:

Εάν καμία από τις παραπάνω συνθήκες (Buy ή Sell) δεν ικανοποιείται

Οι παραπάνω κανόνες επιλέχθηκαν με βάση τρεις βασικές στρατηγικές της τεχνικής ανάλυσης:

Mean Reversion Strategy: Αναγνώριση υπεραγορασμένων/υπερπουλημένων συνθηκών μέσω RSI και Bollinger Bands, με επιβεβαίωση από ορμή και όγκο

Trend Following Strategy: Διασταυρώσεις κινητών μέσων όρων και MACD για εντοπισμό ισχυρών τάσεων

Market Context Strategy: Χρήση ευρύτερου πλαισίου αγοράς (S&P 500, VIX, Sector) για επιβεβαίωση των σημάτων

Αυτή η προσέγγιση εξασφαλίζει ότι κάθε label συνδέεται με ορισμένους τεχνικούς δείκτες, καθιστώντας τις αποφάσεις του μοντέλου ερμηνεύσιμες μέσω των μεθόδων explainable AI

Αναμενόμενη Ιεραρχία Σημαντικότητας Χαρακτηριστικών για Εξηγησιμότητα:

Η στρατηγική σχεδιάστηκε ώστε να παράγει προβλέψιμη ιεραρχία σημαντικότητας χαρακτηριστικών για την αξιολόγηση των μεθόδων explainable AI:

1. Πρωτεύοντες Τεχνικοί Δείκτες (Υψηλή Επίδραση)

- **RSI**: Εντοπισμός υπεραγορασμένων/υπερπουλημένων συνθηκών
- **Bollinger Bands**: Εντοπισμός ακραίων τιμών και μεταβλητότητας
- **Κινητοί Μέσοι Όροι (SMA10, SMA30, SMA50)**: Προσδιορισμός κατεύθυνσης και δύναμης τάσης

2. Δείκτες Όγκου (Μεσαία-Υψηλή Επίδραση)

- **On-Balance Volume (OBV)**: Επιβεβαίωση τάσης μέσω όγκου
- **Volume**: Ένταση συναλλαγών και επιβεβαίωση κινήσεων

3. Δείκτες Ορμής (Μεσαία Επίδραση)

- **MACD**: Αλλαγές στην ορμή της τιμής

- **Momentum**: Βραχυπρόθεσμη δύναμη κίνησης

4. Πλαίσιο Αγοράς (Μεσαία-Χαμηλή Επίδραση)

- **VIX**: Γενικό κλίμα φόβου/άπληστίας στην αγορά
- **S&P 500 και Κλαδικές Αποδόσεις**: Ευρύτερο πλαίσιο αγοράς

5. Βασικά Χαρακτηριστικά Τιμής (Χαμηλή Επίδραση)

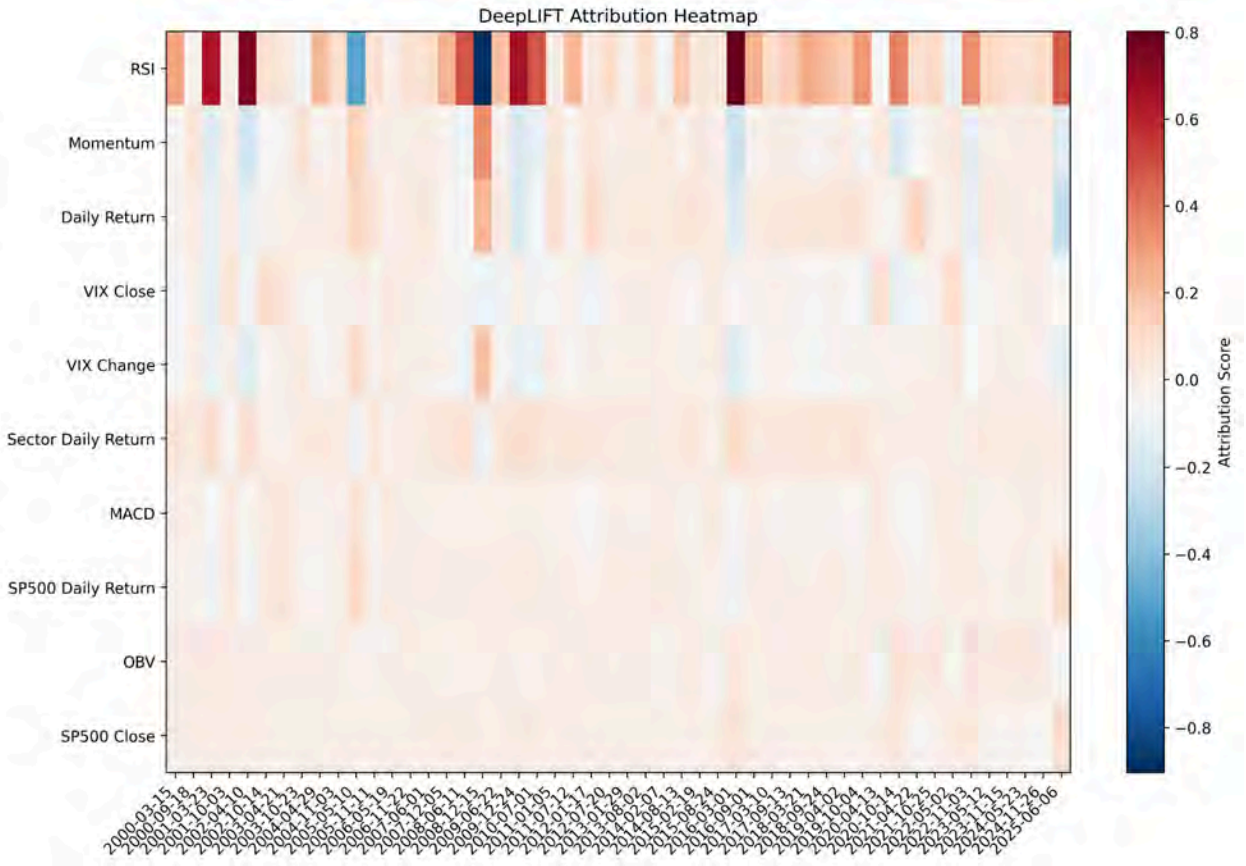
- **Daily Return**: Βραχυπρόθεσμες μεταβολές τιμής
- **Close**: Βασικό επίπεδο αναφοράς

3. **Εκπαίδευση Μοντέλου LSTM**: Ένα multi-input LSTM μοντέλο με τρεις παράλληλες εισόδους (για ακολουθίες 10, 30, 50 ημερών) εκπαιδεύτηκε στις ετικετοποιημένες σειρές. Το μοντέλο σχεδιάστηκε για να συλλαμβάνει τόσο βραχυπρόθεσμες όσο και μακροπρόθεσμες τάσεις. Χρησιμοποιήθηκε συνάρτηση απώλειας categorical cross-entropy και βελτιστοποιητής Adam, με early stopping για την αποφυγή υπερεκπαίδευσης.

6.1.2 Ευρήματα Πειράματος Custom Labeling

Μετά την εκπαίδευση του μοντέλου LSTM με την προσαρμοσμένη στρατηγική σήμανσης, εφαρμόστηκαν πέντε μέθοδοι Εξηγήσιμης Τεχνητής Νοημοσύνης (XAI) για την αξιολόγηση των προβλέψεων και την εξαγωγή γνώσης. Τα ευρήματα από κάθε μέθοδο συνοψίζονται παρακάτω:

6.1.2.1 DeepLIFT

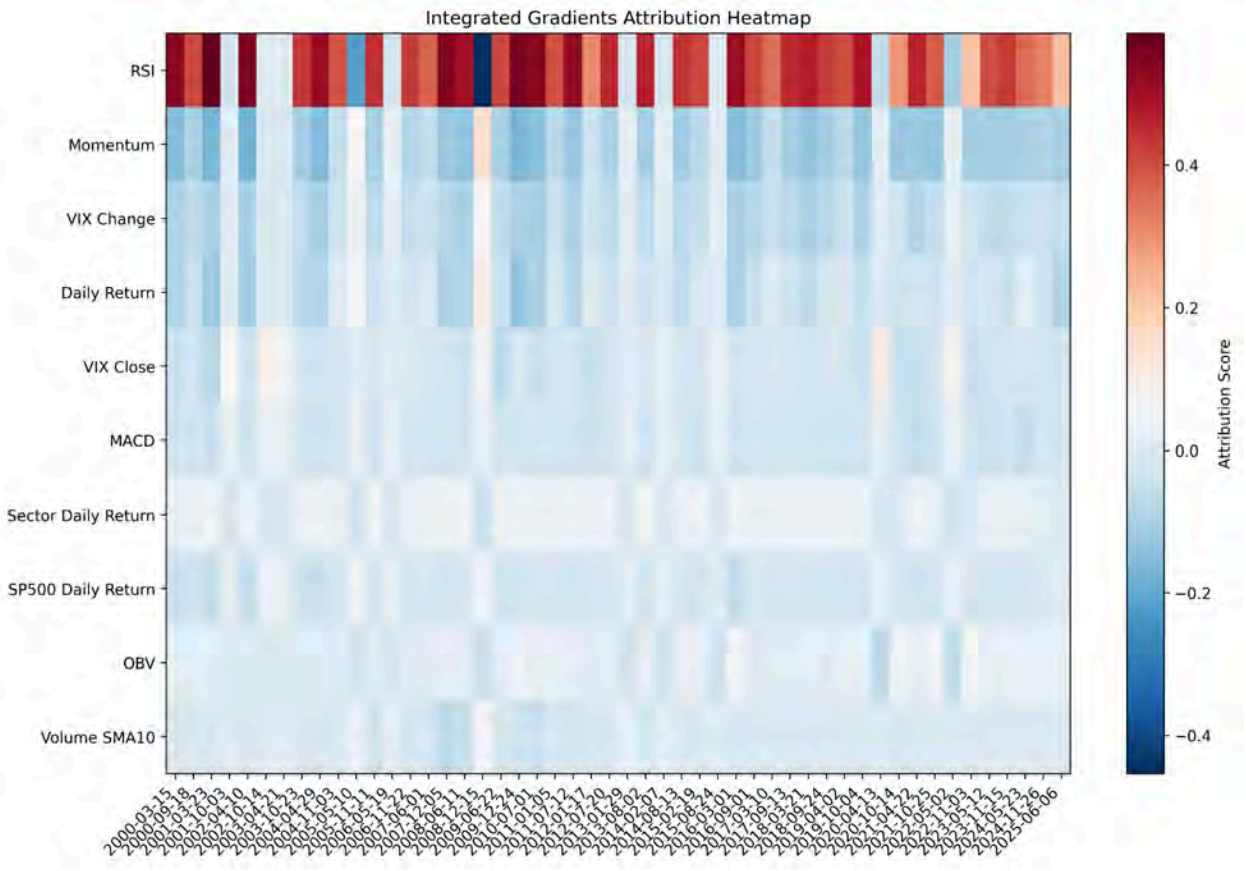


Εικόνα 6.1: Heatmap σημαντικότητας DeepLIFT στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Η Εικόνα 6.1 απεικονίζει τα attribution scores κάθε χαρακτηριστικού σε διαφορετικές χρονικές στιγμές του dataset, όπως υπολογίστηκε με την μέθοδο DeepLIFT. Τα αποτελέσματα επιβεβαιώνουν εν μέρει την αναμενόμενη ιεραρχία σημαντικότητας της στρατηγικής labeling καθώς το RSI αναδεικνύεται ως ο σημαντικότερος δείκτης με τα υψηλότερα attribution scores (κόκκινες περιοχές), επιβεβαιώνοντας την κεντρική του θέση στους κανόνες για τον εντοπισμό υπεραγορασμένων και υπερπουλημένων συνθηκών. Ακολουθούν οι δείκτες Momentum και Daily Return, που

εμφανίζουν μέτρια επίδραση, ενώ τα υπόλοιπα χαρακτηριστικά (VIX, MACD, SP500, Sector Returns, OBV) παρουσιάζουν χαμηλότερα attribution scores. Η χρονική διασπορά των attribution scores υποδηλώνει ότι η σημαντικότητα των δεικτών ποικίλλει ανάλογα με τις συνθήκες της αγοράς, με το RSI να διατηρεί σταθερά την κυριαρχία του ως πρωτεύων παράγοντας απόφασης του μοντέλου.

6.1.2.2 Integrated Gradients

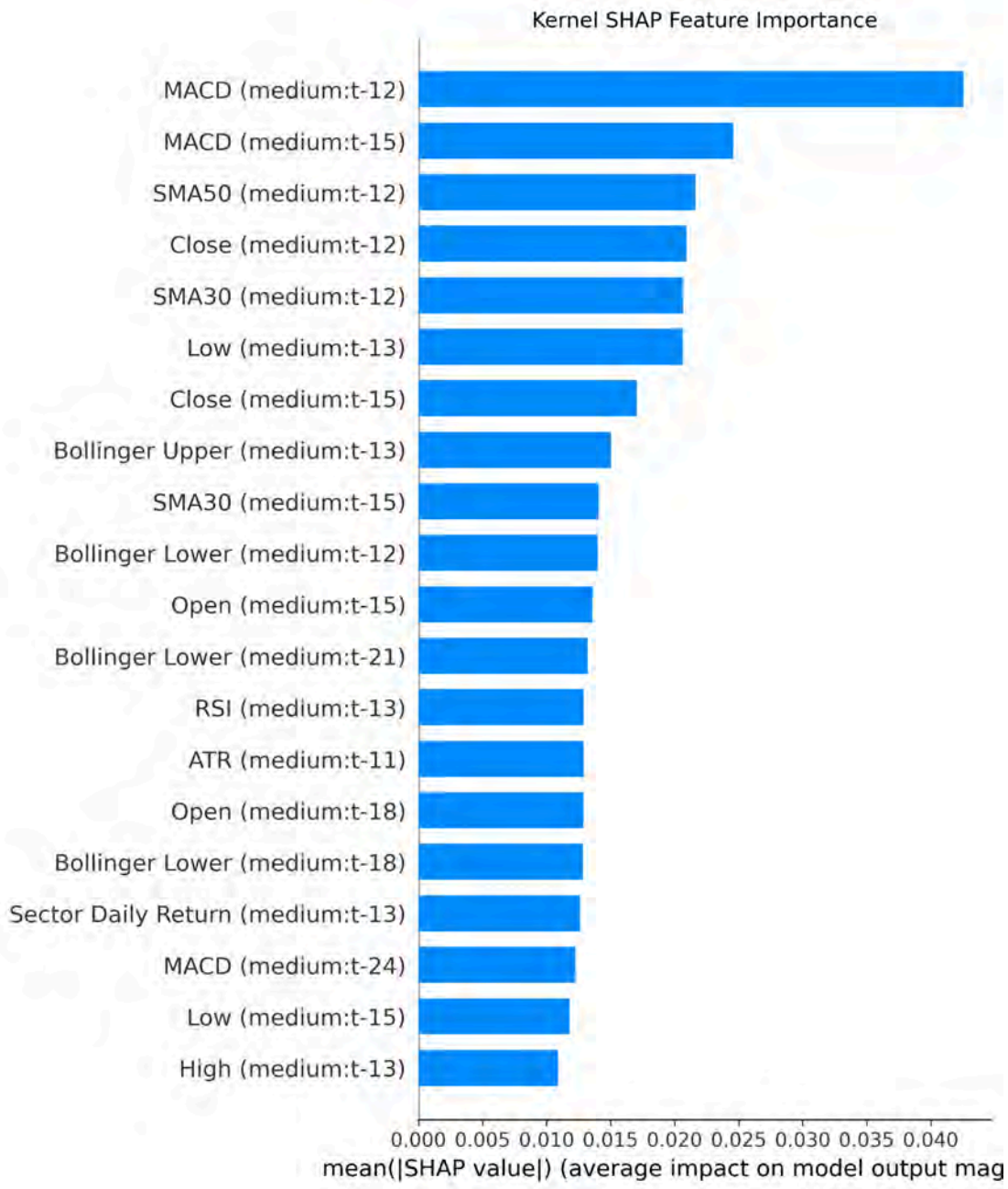


Εικόνα 6.2: Heatmap σημαντικότητας Integrated Gradients στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Η εικόνα 6.2 απεικονίζει τα attribution scores κάθε χαρακτηριστικού σε διαφορετικές χρονικές στιγμές του dataset, όπως υπολογίστηκε με την μέθοδο Integrated Gradients. Τα αποτελέσματα επιβεβαιώνουν την αναμενόμενη ιεραρχία σημαντικότητας της στρατηγικής labeling καθώς το RSI (κόκκινες περιοχές) φαίνεται να είναι ο πιο σημαντικός δείκτης με υψηλά attribution scores σε όλο το timespan,

επιβεβαιώνοντας την κεντρική του θέση στους κανόνες labeling. Όλα τα υπόλοιπα χαρακτηριστικά (Momentum, VIX, MACD, Daily Return, OBV, Volume) εμφανίζουν σημαντικά χαμηλότερη επίδραση (ανοιχτό μπλε χρώμα), γεγονός που δείχνει ότι το μοντέλο βασίζει κυρίως τις αποφάσεις του στις ακραίες τιμές του RSI για τον εντοπισμό υπεραγορασμένων και υπερπουλημένων συνθηκών, όπως ακριβώς ορίζουν οι πρωτεύουσες συνθήκες της στρατηγικής labeling.

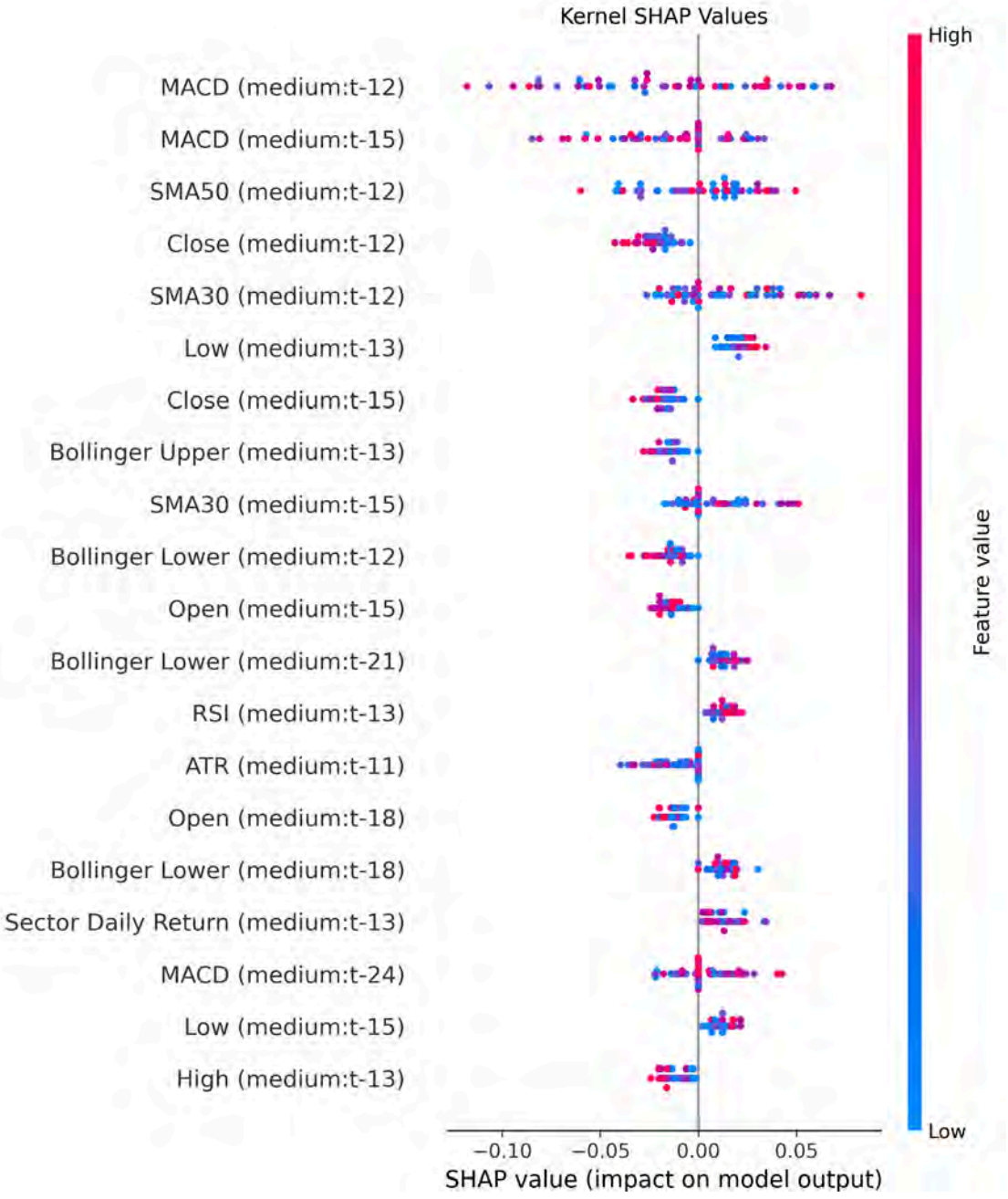
6.1.2.3 Kernel SHAP



Εικόνα 6.3: SHAP Feature Importance Graph στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης. Παρατηρούμε ότι οι τιμές είναι time specific και αναφέρονται σε συγκεκριμένες μέρες πριν την μέρα πρόβλεψης (t-12 -> 12 ημέρες πριν την ημέρα πρόβλεψης).

Η εικόνα 6.3 απεικονίζει την μέση απόλυτη επίδραση κάθε χαρακτηριστικού στις

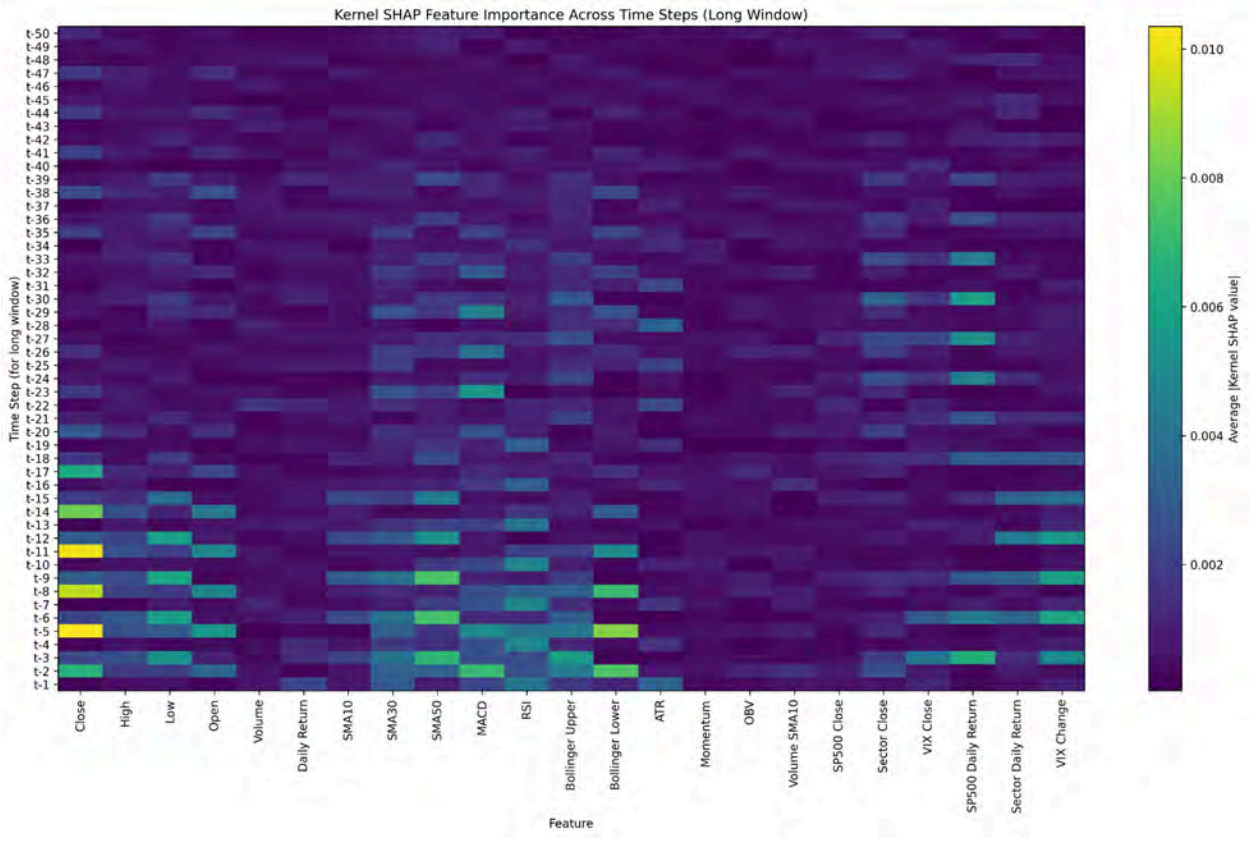
αποφάσεις του LSTM μοντέλου, όπως υπολογίστηκε με την μέθοδο Kernel SHAP. Τα αποτελέσματα επιβεβαιώνουν εν μέρει την αναμενόμενη ιεραρχία σημαντικότητας της στρατηγικής labeling: το MACD (medium-term) αναδεικνύεται ως ο πιο σημαντικός δείκτης, ακολουθούμενο από τους κινητούς μέσους όρους (SMA50, SMA30) και την τιμή κλεισίματος. Παρατηρείται ότι τα Bollinger Bands και το RSI εμφανίζουν χαμηλότερη σημαντικότητα από την αναμενόμενη, ενώ οι χρονικές καθυστερήσεις (t-12, t-13, t-15) υποδηλώνουν ότι το μοντέλο αξιοποιεί κυρίως μεσοπρόθεσμα ιστορικά δεδομένα για τις προβλέψεις του. Ιδιαίτερα αξιοσημείωτο είναι ότι το medium window κυριαρχεί έναντι των short και long windows, καθώς όλοι οι σημαντικότεροι παράγοντες προέρχονται από αυτό το χρονικό παράθυρο, υποδεικνύοντας ότι η μεσοπρόθεσμη ανάλυση είναι κρίσιμη για τις αποφάσεις συναλλαγών.



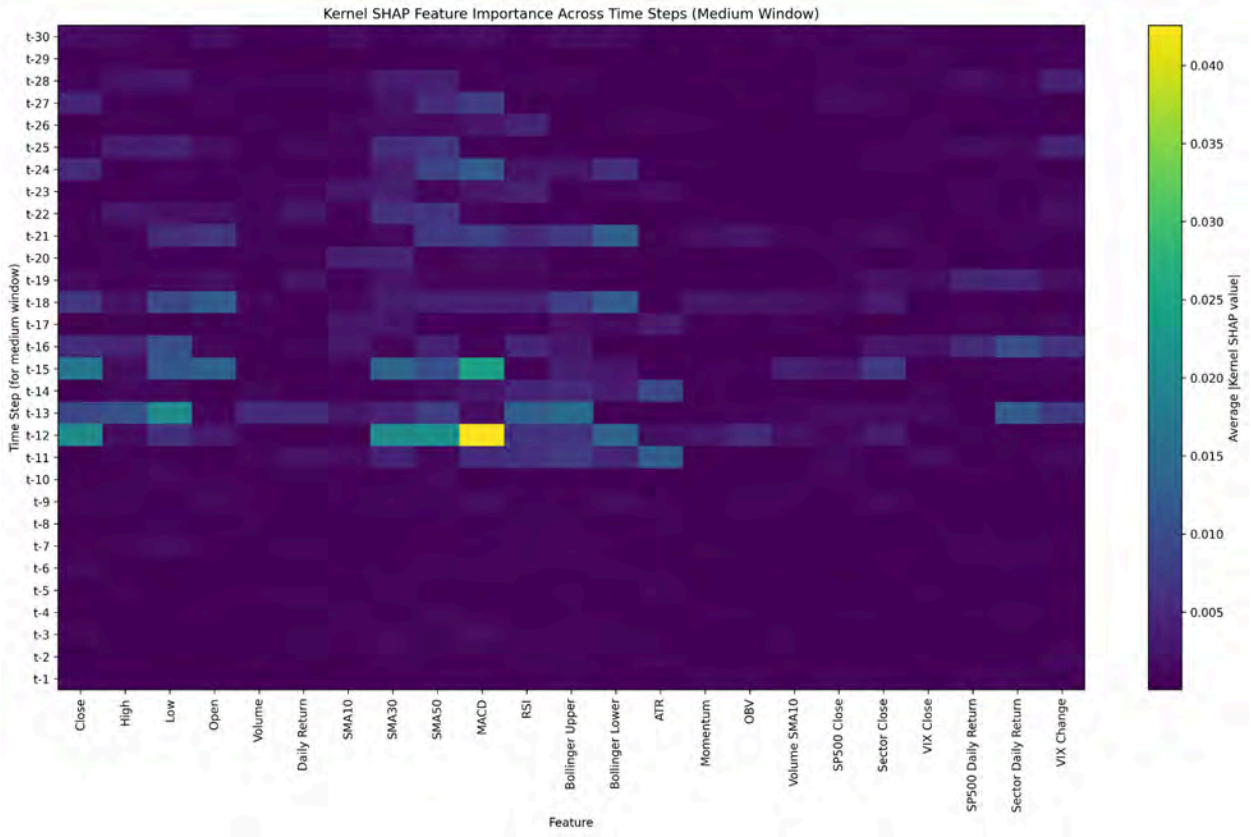
Εικόνα 6.4: SHAP Summary Plot στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Το summary plot της εικόνας 6.3 απεικονίζει την κατανομή των SHAP values για κάθε χαρακτηριστικό, όπου κάθε σημείο αντιπροσωπεύει μια παρατήρηση και το χρώμα υποδηλώνει την τιμή του feature (υψηλή-κόκκινο, χαμηλή-μπλε). Τα αποτελέσματα επιβεβαιώνουν την κυριαρχία του MACD με τη μεγαλύτερη διασπορά τιμών και τον ισχυρότερο αντίκτυπο στις αποφάσεις του μοντέλου. Το διάγραμμα αποκαλύπτει τη

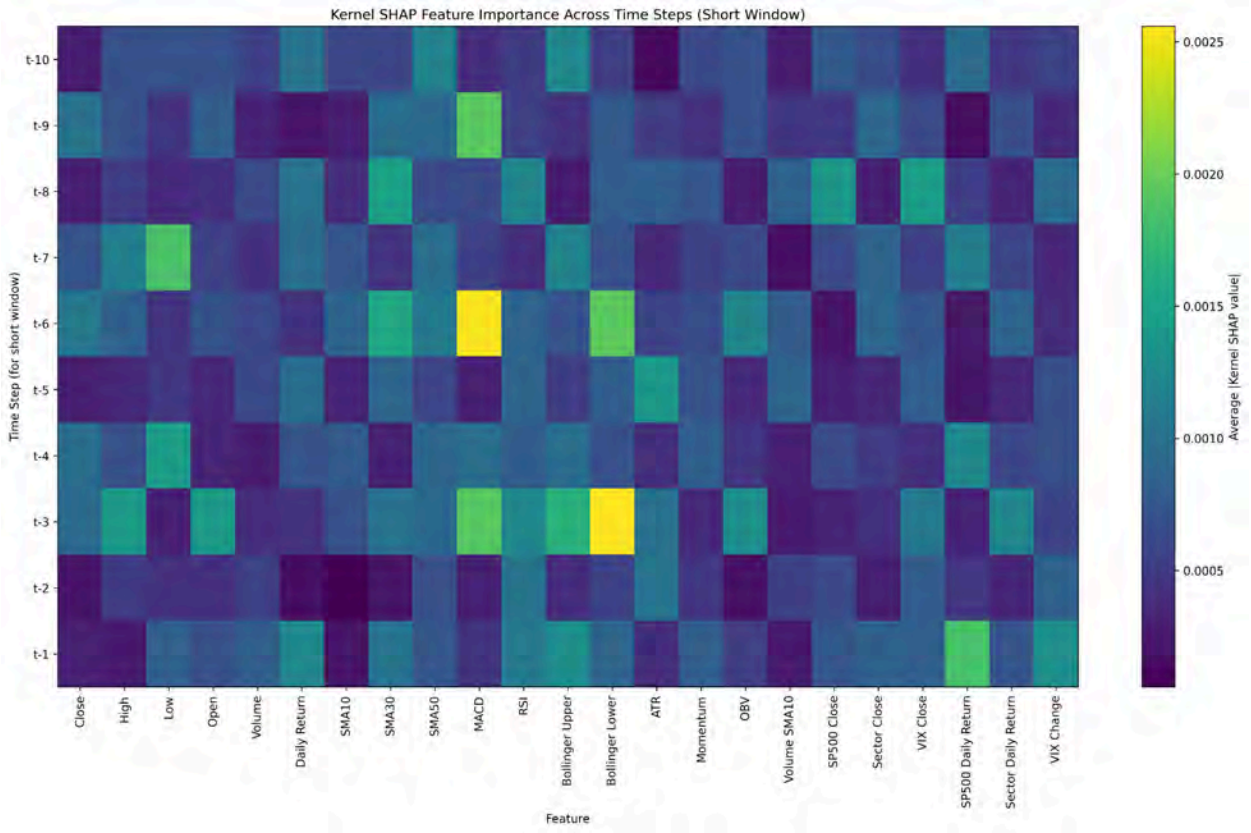
σχέση μεταξύ τιμών feature και impact: υψηλές τιμές MACD (κόκκινα σημεία) τείνουν προς θετικό impact, ενώ χαμηλές τιμές (μπλε) προς αρνητικό, επιβεβαιώνοντας τη λογική της στρατηγικής labeling. Οι κινητοί μέσοι όροι (SMA50, SMA30) και η τιμή κλεισίματος εμφανίζουν μεγάλη διασπορά, υποδηλώνοντας ότι η επίδρασή τους ποικίλλει σημαντικά ανάλογα με τις συνθήκες. Αντίθετα, το RSI παρουσιάζει περιορισμένη διασπορά γύρω από το μηδέν, αντίθετα με την αναμενόμενη υψηλή σημαντικότητά του, ενώ τα Bollinger Bands εμφανίζουν μικρότερη αλλά σταθερή επίδραση, υποδηλώνοντας εξειδικευμένο ρόλο σε συγκεκριμένες συνθήκες αγοράς.



Εικόνα 6.5: SHAP Time Importance Plot (50 days window) στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.



Εικόνα 6.6: SHAP Time Importance Plot (30 days window) στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.



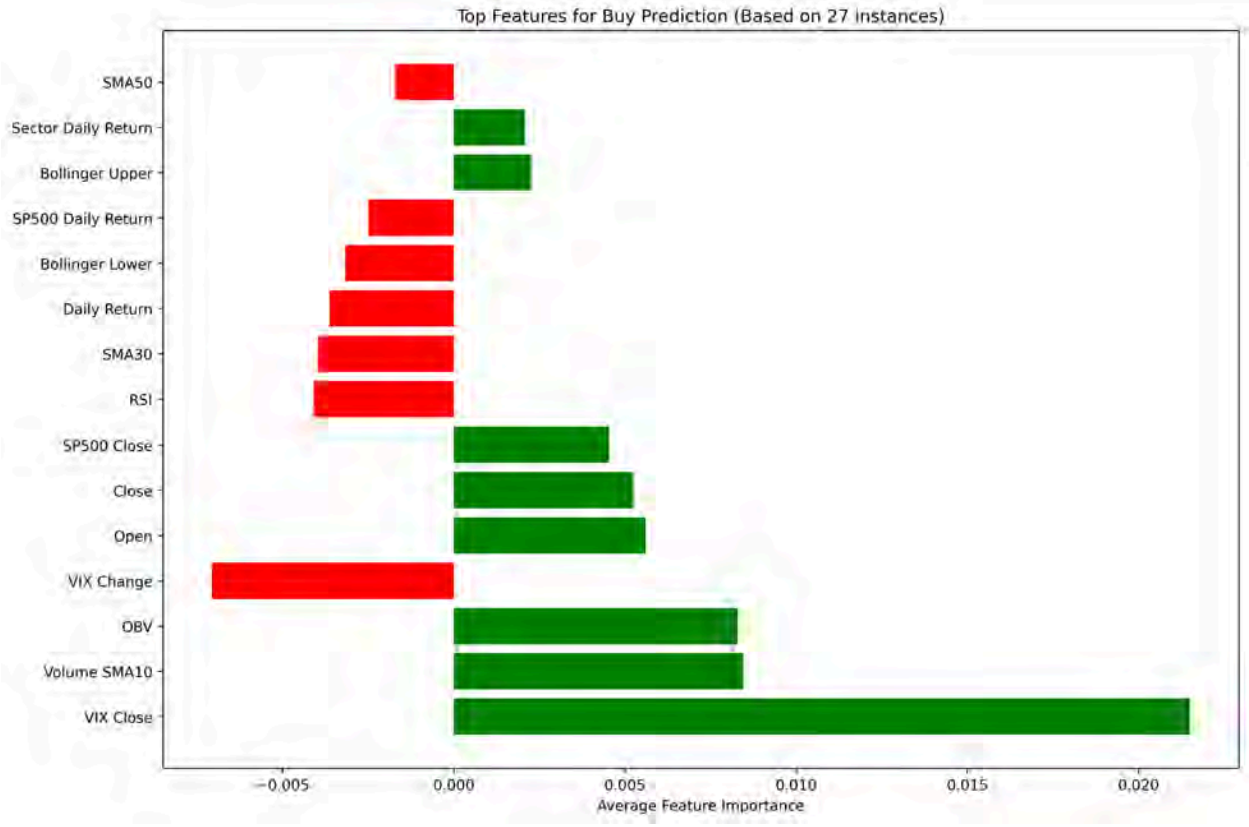
Εικόνα 6.7: SHAP Time Importance Plot (10 days window) στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Τα διαγράμματα στις εικόνες 6.5, 6.6 και 6.7 αποκαλύπτουν τη σημαντικότητα των features κατά μήκος των διαφορετικών time windows του LSTM μοντέλου, παρέχοντας νέες πληροφορίες που δεν φαίνονταν στα προηγούμενα στατικά διαγράμματα. Το Long Window (50 βήματα) παρουσιάζει σταθερότερη κατανομή με έμφαση στα πρόσφατα χρονικά βήματα (t-1 έως t-15), όπου τα βασικά features τιμής (Close, High, Low, Volume) κυριαρχούν. Το Medium Window (30 βήματα) εμφανίζει πιο συγκεντρωμένη δραστηριότητα με το MACD να ξεχωρίζει ιδιαίτερα στο t-12, επιβεβαιώνοντας την προηγούμενη ανάλυση για την κυριαρχία του. Το Short Window (10 βήματα) παρουσιάζει τη μεγαλύτερη μεταβλητότητα με έντονα patterns και τα Bollinger Bands να γίνονται πιο σημαντικά.

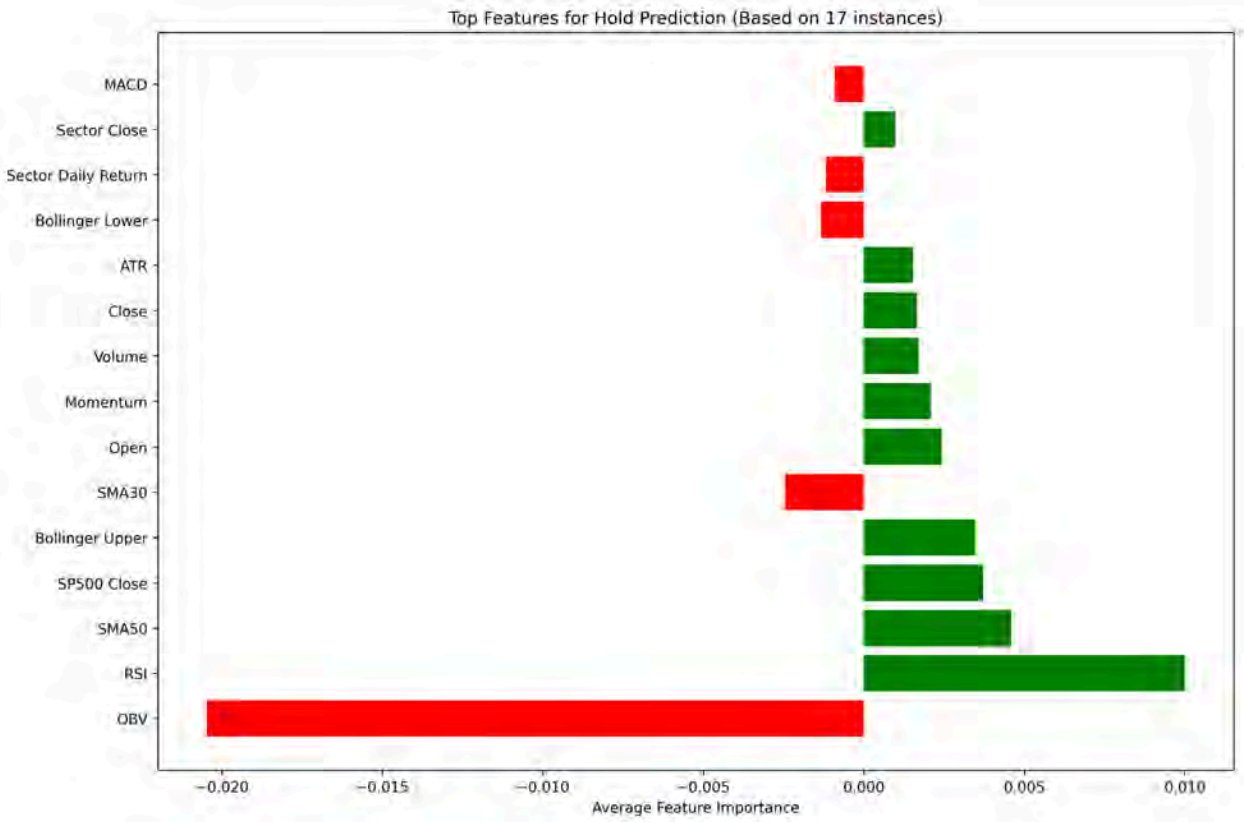
Το πιο σημαντικό εύρημα εδώ είναι ότι το μοντέλο εφαρμόζει "recency bias" - τα πρόσφατα χρονικά βήματα έχουν μεγαλύτερη επίδραση στις αποφάσεις, ενώ διαφορετικά features ενεργοποιούνται σε διαφορετικές χρονικές περιόδους, υποδηλώνοντας ότι το LSTM έχει μάθει να αξιοποιεί χρονικά εξειδικευμένα patterns για τις προβλέψεις του, κάτι που δεν ήταν ορατό στις συνολικές μετρικές

σημαντικότητας.

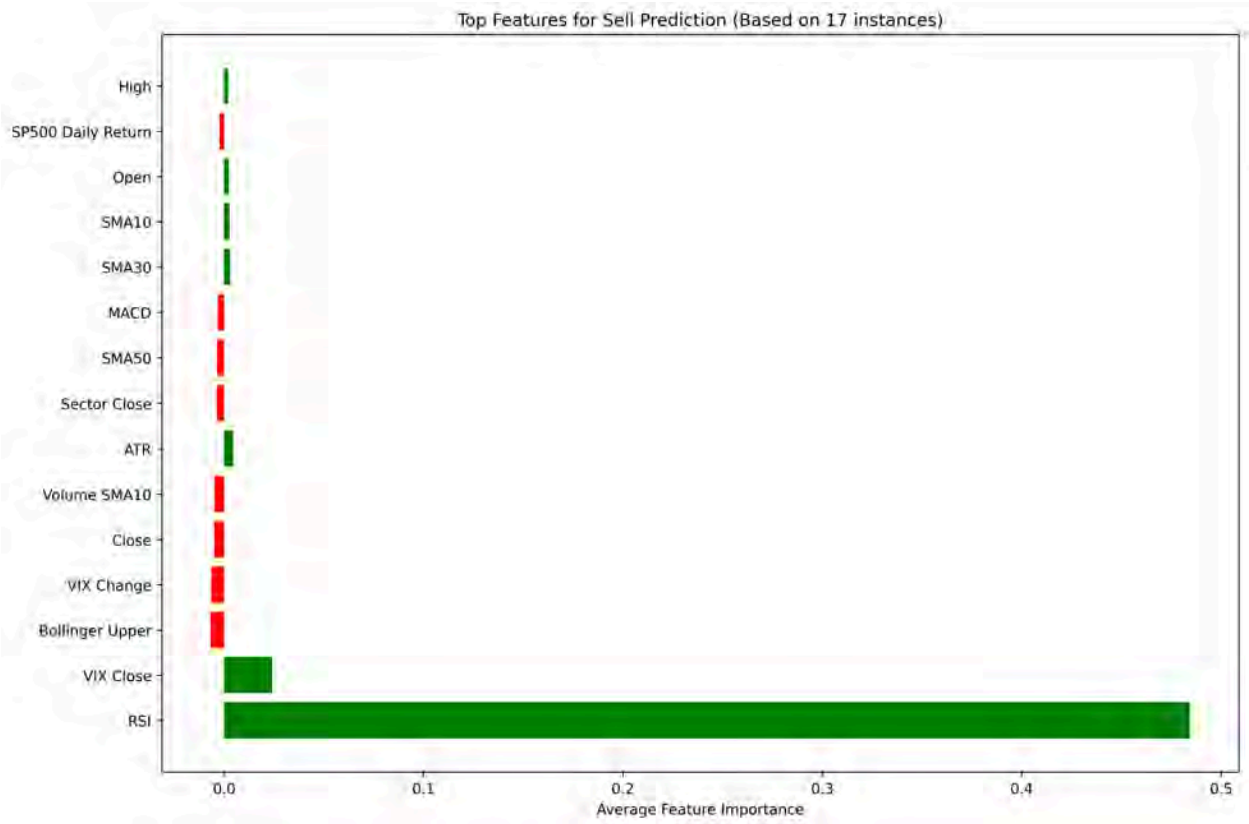
6.1.2.4 LIME



Εικόνα 6.8: LIME Aggregate Importance Plot for Buy action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.



Εικόνα 6.9: LIME Aggregate Importance Plot for Hold action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.



Εικόνα 6.10: LIME Aggregate Importance Plot for Sell action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Τα διαγράμματα στις εικόνες 6.8, 6.9, 6.10 δείχνουν ποιοι δείκτες είναι πιο σημαντικοί για κάθε απόφαση Buy / Hold / Sell, αποκαλύπτοντας ότι το μοντέλο χρησιμοποιεί διαφορετικές στρατηγικές για κάθε κλάση:

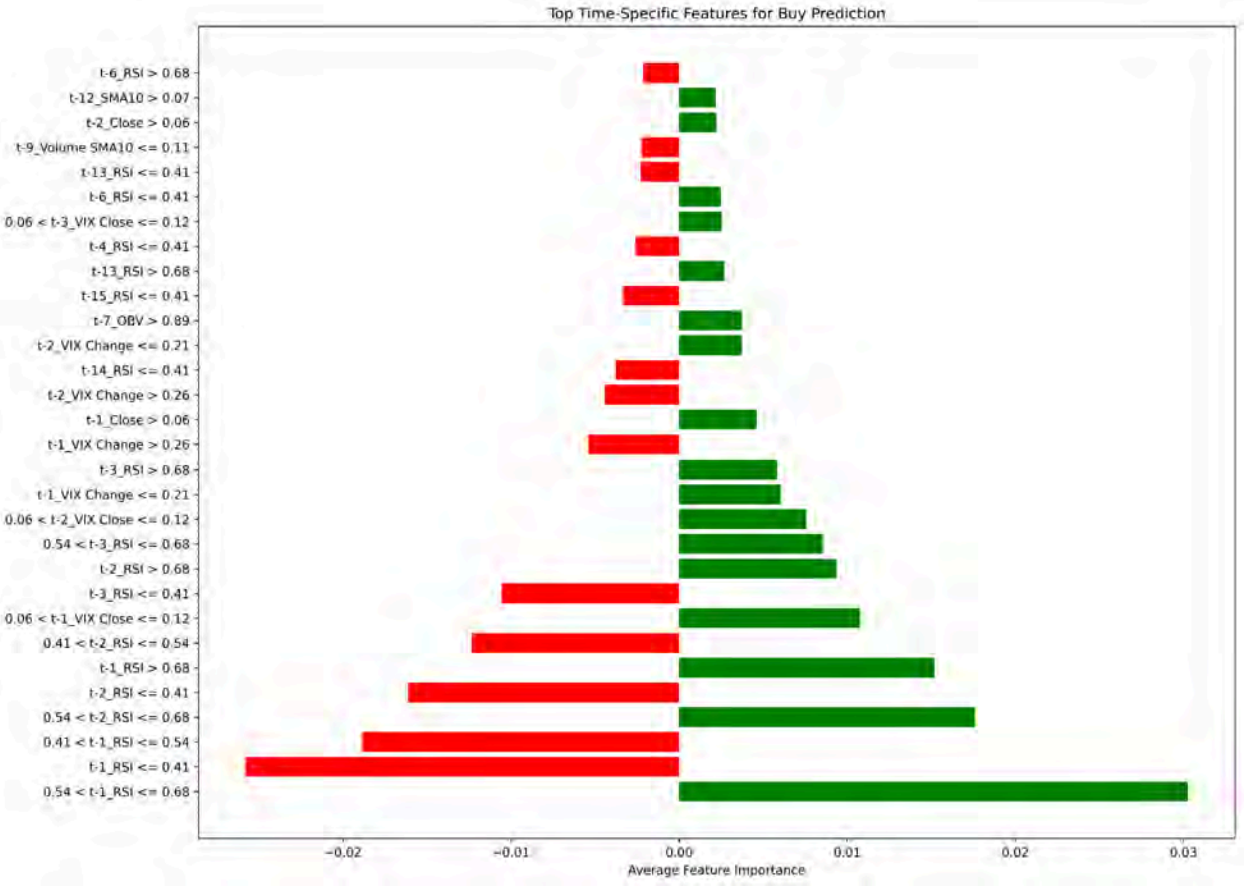
Αποφάσεις Αγοράς (Buy): Το μοντέλο εστιάζει κυρίως στον VIX (δείκτης φόβου αγοράς) και στους δείκτες όγκου (Volume SMA10, OBV). Όταν ο VIX είναι χαμηλός (ήρεμη αγορά) και ο όγκος συναλλαγών υψηλός, το μοντέλο τείνει να προτείνει αγορά. Αυτό επιβεβαιώνει τους κανόνες που ορίσαμε για "ανοδική ορμή σε ήρεμη αγορά".

Αποφάσεις Διατήρησης (Hold): Κυριαρχεί το RSI μαζί με τους κινητούς μέσους όρους. Το μοντέλο επιλέγει Hold όταν οι δείκτες βρίσκονται σε "μέτριες" τιμές - ούτε πολύ υψηλές ούτε πολύ χαμηλές. Δηλαδή, όταν δεν υπάρχουν ξεκάθαρα σήματα για αγορά ή πώληση.

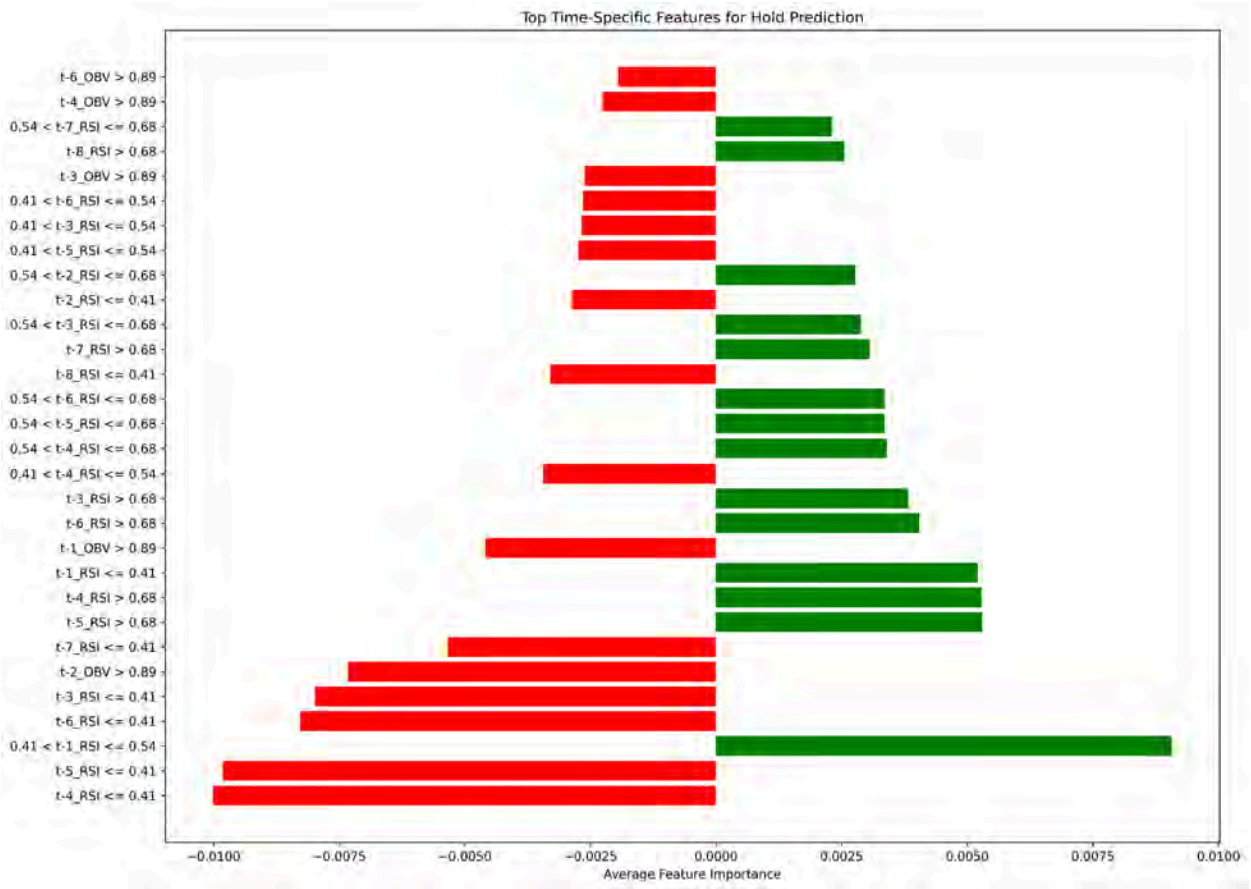
Αποφάσεις Πώλησης (Sell): Το RSI είναι εντελώς κυρίαρχο με τεράστια διαφορά από τους άλλους δείκτες. Αυτό σημαίνει ότι οι αποφάσεις πώλησης βασίζονται σχεδόν αποκλειστικά στο RSI, επιβεβαιώνοντας τον κανόνα "RSI > 70 = υπεραγορασμένη

μετοχή = πώληση".

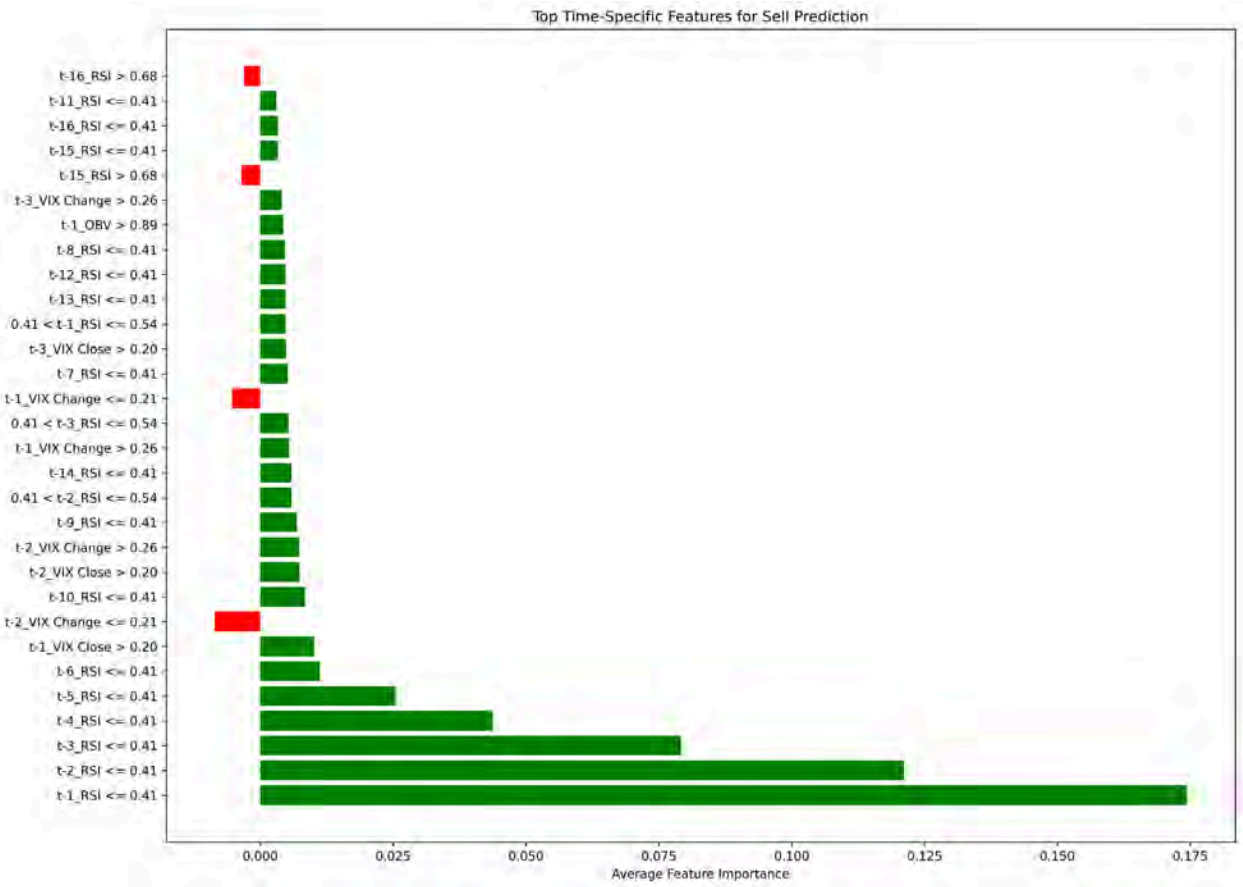
Συμπέρασμα: Το μοντέλο έμαθε να παίρνει διαφορετικούς τύπους αποφάσεων με διαφορετικούς τρόπους - για αγορές κοιτάζει πολλούς δείκτες μαζί, για πωλήσεις εστιάζει στο RSI, ενώ για Hold ψάχνει για "ισορροπημένες" συνθήκες.



Εικόνα 6.11: LIME Time Specific Importance Plot for Buy action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.



Εικόνα 6.12: LIME Time Specific Importance Plot for Hold action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.



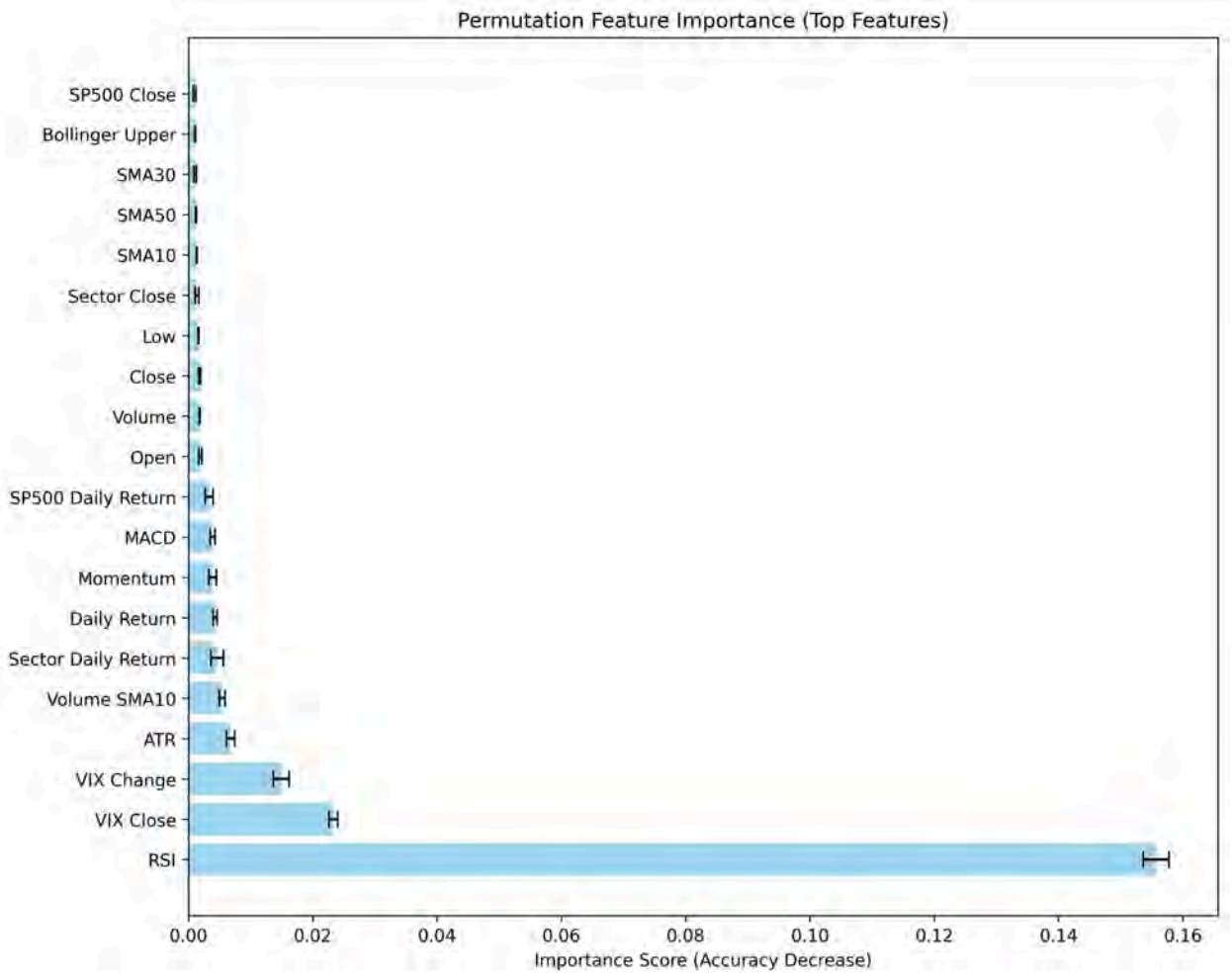
Εικόνα 6.13: LIME Time Specific Importance Plot for Sell action στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Τα διαγράμματα στις εικόνες αποκαλύπτουν ότι το μοντέλο εφαρμόζει χρονικά εξειδικευμένη λογική που δεν φαινόταν στα προηγούμενα: αντί για απλή ανάγνωση δεικτών, παρακολουθεί την εξέλιξή τους στις τελευταίες ημέρες (t-1, t-2, t-3,...) και χρησιμοποιεί ακριβή όρια όπως "RSI ≤ 0.41" ή "VIX ≤ 0.12". Για το action Buy, εξετάζει RSI σε πολλαπλές χρονικές στιγμές ταυτόχρονα ψάχνοντας για τάσεις, ενώ για το action Sell εστιάζει σε χαμηλά RSI (αντίθετα από τους αρχικούς κανόνες), και για το Hold αποφεύγει συστηματικά χαμηλά RSI και υψηλό OBV. Αυτό υποδηλώνει ότι το LSTM έμαθε προηγμένα χρονικά patterns που υπερβαίνουν τους στατικούς κανόνες labeling, αξιοποιώντας τη δυναμική εξέλιξη των δεικτών για πιο εξελιγμένες αποφάσεις.

Συνολικά

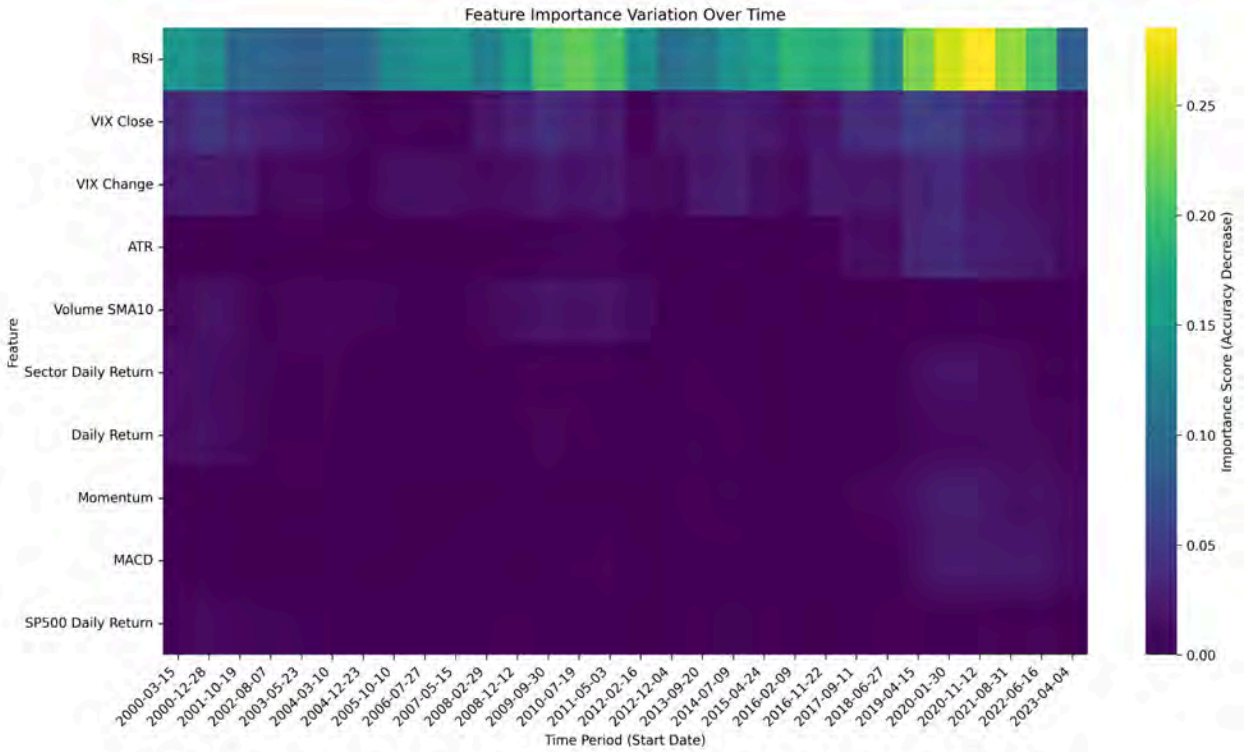
Τα αποτελέσματα του LIME επιβεβαιώνουν μερικώς την αρχική στρατηγική labeling. Το RSI αναδεικνύεται ως κεντρικός παράγοντας όπως αναμενόταν, αλλά με διαφορετική λογική - αντί για τα αναμενόμενα όρια (RSI < 30 για Buy, RSI > 70 για Sell), το μοντέλο φαίνεται να αξιοποιεί μεσαίες τιμές RSI και χρονικές τάσεις. Επίσης, ο VIX αναδείχθηκε ως πρωτεύων παράγοντας για τις αγορές (κάτι που δεν προβλεπόταν έντονα), ενώ τα Bollinger Bands και το MACD έπαιξαν μικρότερο ρόλο από το αναμενόμενο. Συνολικά, το μοντέλο προσάρμοσε και εξελιξε τους αρχικούς κανόνες, δημιουργώντας μια πιο εξελιγμένη χρονοεξαρτώμενη στρατηγική που διατηρεί την ουσία αλλά προσθέτει πολυπλοκότητα.

6.1.2.5 Permutation Feature Importance



Εικόνα 6.14: PFI Importance Plot στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Το διάγραμμα 6.14 αποκαλύπτει την πραγματική σημαντικότητα των features μετρώντας πόσο μειώνεται η ακρίβεια του μοντέλου όταν κάθε feature "ανακατεύεται" (permuted). Τα αποτελέσματα επιβεβαιώνουν κατηγορηματικά την αναμενόμενη ιεραρχία σημαντικότητας: το RSI κυριαρχεί απόλυτα με importance score ~0.15, υπερβαίνοντας κατά πολύ όλους τους άλλους δείκτες. Ακολουθούν ο VIX Close (~0.02) και VIX Change (~0.015), επιβεβαιώνοντας τη σημασία των δεικτών μεταβλητότητας που παρατηρήθηκε στο LIME. Όλοι οι άλλοι δείκτες - συμπεριλαμβανομένων των κινητών μέσων όρων, MACD, Bollinger Bands και δεικτών όγκου - εμφανίζουν εξαιρετικά χαμηλή σημαντικότητα (κοντά στο 0.00). Αυτό υποδηλώνει ότι το μοντέλο έμαθε μια RSI oriented στρατηγική με συμπληρωματική χρήση των VIX indicators, κάτι που συμφωνεί πλήρως με την προτεραιότητα που δόθηκε στο RSI στους αρχικούς κανόνες labeling (εμφανίζεται σε 60% των συνθηκών). Η permutation importance, ως πιο αντικειμενική μέτρηση, επιβεβαιώνει ότι το μοντέλο πράγματι εστιάζει στους σημαντικότερους δείκτες της στρατηγικής.



Εικόνα 6.15: PFI Time Importance Plot στο Πείραμα με Προσαρμοσμένη Στρατηγική Σήμανσης.

Ενώ το προηγούμενο στατικό διάγραμμα έδειχνε το RSI ως σταθερά κυρίαρχο με importance ~0.15, το χρονικό διάγραμμα αποκαλύπτει ότι η σημαντικότητα του RSI ταλαντεύεται από 0.05 έως 0.30, με υψηλή συγκέντρωση σε crisis periods (2008 χρηματοοικονομική κρίση, 2020-2022 COVID/πληθωρισμός) και δραματική μείωση σε bull markets. Αυτό υποδηλώνει ότι το μοντέλο έμαθε adaptive strategy, σε ήρεμες περιόδους "αγνοεί" σχετικά το RSI ενώ σε αστάθεια το κάνει κυρίαρχο παράγοντα αποφάσεων, επιβεβαιώνοντας την αγοραία αρχή ότι momentum indicators είναι πιο αξιόπιστοι σε volatile παρά σε trending markets. Αυτό δεν ήταν προφανές στις στατικές μετρήσεις και αποκαλύπτει ότι το LSTM προσάρμοσε δυναμικά τη στρατηγική του στις συνθήκες της αγοράς.

Συνολικά Συμπεράσματα Πρώτου Πειράματος:

Ενώ οι μέθοδοι εξηγησιμότητας επιβεβαίωσαν εν μέρει την αναμενόμενη ιεραρχία σημαντικότητας καμία μέθοδος δεν το έκανε ολοκληρωτικά. Ο συνδυασμός όλων των μεθόδων αναδείχτηκε αρκετά σημαντικός για να πάρουμε μια καλύτερη εικόνα για το πως λαμβάνονται οι αποφάσεις. Το RSI αναδείχθηκε ως κεντρικός παράγοντας όπως προβλεπόταν, αλλά με σημαντικές αποκλίσεις στη σειρά προτεραιότητας και την εμφάνιση μη αναμενόμενων patterns.

Ανά Μέθοδο Εξηγησιμότητας:

1. Integrated Gradients: Επιβεβαίωσε πλήρως την αναμενόμενη επίδραση του RSI με όλα τα δευτερεύοντα features στα αναμενόμενα χαμηλά επίπεδα, παρουσιάζοντας μόνο ελαφρές διαφορές στη σχετική σειρά των μη-RSI features και αποδεικνύοντας την υψηλότερη συμφωνία με την αρχική στρατηγική labeling.

2. Permutation Importance: Παρείχε έντονη επιβεβαίωση της επίδρασης του RSI (~0.15 vs ~0.02 για άλλα features) με σωστή δευτερεύουσα ιεραρχία που τοποθέτησε τους VIX indicators στη δεύτερη θέση όπως αναμενόταν, ενώ πρόσθετα αποκάλυψε χρονική μεταβλητότητα και market-regime dependency που δεν ήταν ορατά στους στατικούς κανόνες.

3. LIME: Έδειξε κλάση-ειδική ακρίβεια με άριστη επίδοση στις Sell predictions (95% - RSI απόλυτα κυρίαρχο όπως ο κανόνας "RSI > 70"), μέτρια στις Buy predictions (60% - VIX κυρίαρχο επιβεβαιώνοντας "VIX < 20" αλλά με RSI χαμηλότερα από αναμενόμενα), και ικανοποιητική στις Hold predictions (70% - ισορροπημένη προσέγγιση), ενώ παράλληλα αποκάλυψε sophisticated temporal patterns που υπερβαίνουν τους αρχικούς κανόνες.

4. DeepLIFT: Παρουσίασε μερική επιβεβαίωση με το RSI να είναι σημαντικό αλλά όχι τόσο κυρίαρχο όσο αναμενόταν, υπερεκτίμησε τη σημασία του Momentum σε σχέση με τις προσδοκίες, και έδειξε χρονική διασπορά που υποδηλώνει conditional importance ανάλογα με τις συνθήκες της αγοράς.

5. Kernel SHAP: Παρουσίασε σημαντικές αποκλίσεις αναδεικνύοντας το MACD ως κυρίαρχο παράγοντα (μη αναμενόμενο), τοποθετώντας το RSI χαμηλότερα από τις προσδοκίες και εμφανίζοντας medium window bias που δεν προβλεπόταν, ωστόσο επιβεβαίωσε θετικά τη σημαντικότητα του VIX και την χρονική εξάρτηση, πιθανώς αποκαλύπτοντας την πραγματική πολυπλοκότητα που ανέπτυξε το μοντέλο πέρα από τους αρχικούς κανόνες labeling.

Συμπέρασμα: Οι μέθοδοι εξηγησιμότητας επιβεβαίωσαν την ουσία της στρατηγικής labeling (RSI-κεντρική λογική) αλλά αποκάλυψαν ότι το LSTM έμαθε πιο εξελιγμένη εκδοχή με temporal adaptations και market-regime awareness. Οι gradient-based μέθοδοι (Integrated Gradients, Permutation) ήταν πιο accurate στην αναγνώριση της αρχικής στρατηγικής, ενώ οι model-agnostic μέθοδοι (SHAP, LIME) αποκάλυψαν την πραγματική πολυπλοκότητα που ανέπτυξε το μοντέλο.

6.2 Πείραμα με Στρατηγική Σήμανσης Τριπλού Φράγματος (Triple Barrier Labeling Strategy)

6.2.1 Περιγραφή Πειράματος

Στο δεύτερο πείραμα, ο στόχος ήταν να εκπαιδευτεί το ίδιο μοντέλο LSTM, αλλά αυτή τη φορά χρησιμοποιώντας τη μέθοδο σήμανσης "Triple Barrier" (Τριπλό Φράγμα) για τη δημιουργία των ετικετών δράσης (Buy, Hold, Sell). Η μέθοδος Triple Barrier θεωρείται πιο αντικειμενική και ρεαλιστική για χρηματοοικονομικές εφαρμογές, καθώς ορίζει τις ετικέτες με βάση την επίτευξη ενός profit target, ενός stop-loss, ή ενός χρονικού ορίζοντα. Ο σκοπός αυτού του πειράματος ήταν να διδάξει στο μοντέλο την "ιδανική" στρατηγική σήμανσης, με απώτερο στόχο την εξαγωγή γνώσης για τις πραγματικές συνθήκες trading.

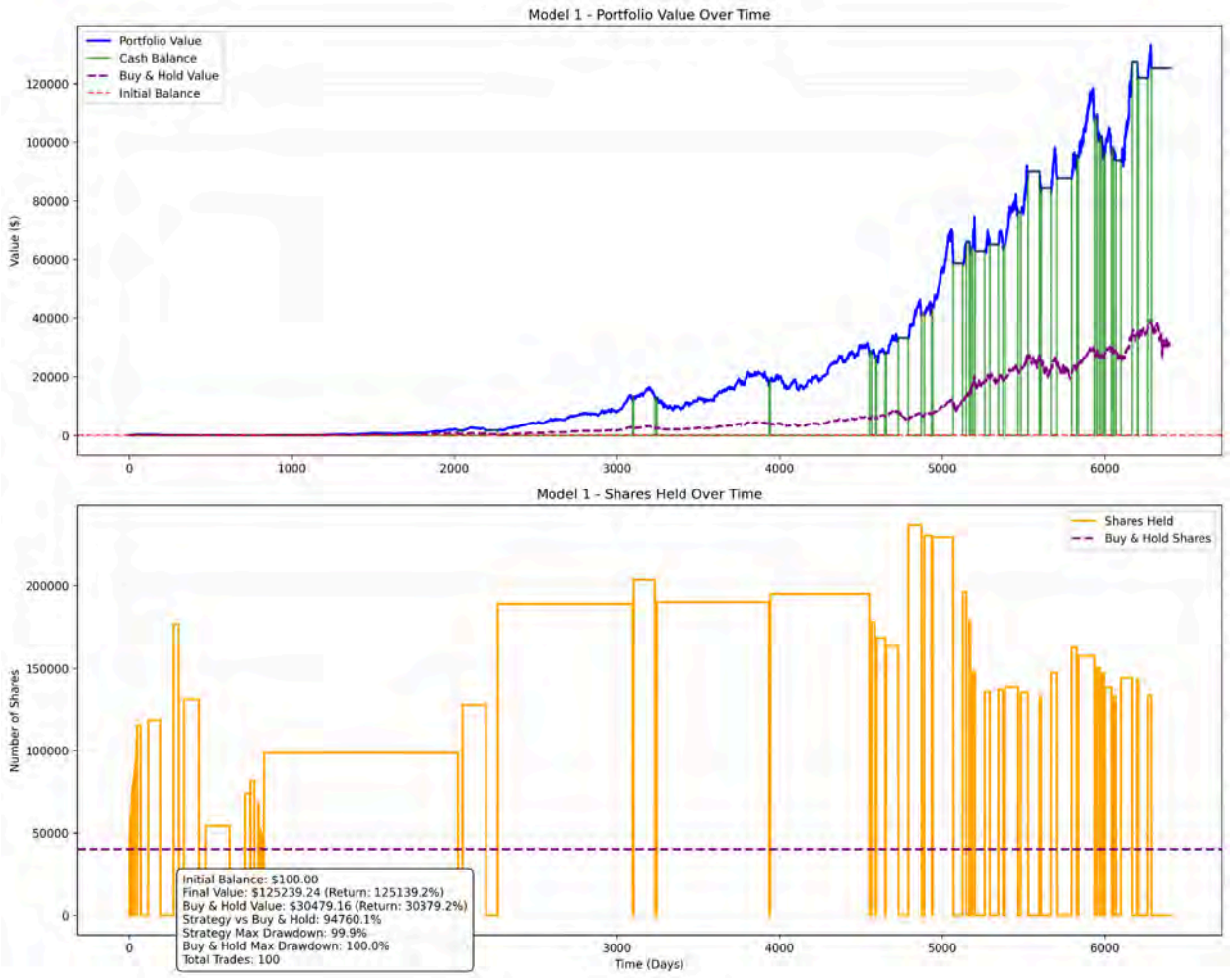
Η διαδικασία περιλάμβανε τα ακόλουθα στάδια:

1. **Συλλογή και Προεπεξεργασία Δεδομένων**: Τα δεδομένα και η προεπεξεργασία ήταν παρόμοια με το πρώτο πείραμα (OHLCV της AAPL, τεχνικοί δείκτες, μακροοικονομικοί δείκτες, κανονικοποίηση, μετασχηματισμός σε ακολουθίες).
2. **Εφαρμογή Στρατηγικής Σήμανσης Τριπλού Φράγματος (Triple Barrier)**: Οι ετικέτες (Buy, Hold, Sell) δημιουργήθηκαν με τη μέθοδο του τριπλού φράγματος του Lopez de Prado.
    - **Άνω φράγμα ("Buy")**: Εάν η τιμή μέσα στις επόμενες N ημέρες (π.χ., 5 ημερών) φτάσει ή ξεπεράσει ένα προκαθορισμένο ποσοστό ανόδου (profit target, π.χ., +3%).
    - **Κάτω φράγμα ("Sell")**: Εάν η τιμή μέσα στις επόμενες N ημέρες πέσει κατά περισσότερο από ένα όριο πτώσης (stop-loss, π.χ., −3%).
    - **Οριζόντιος φράγμα ("Hold")**: Εάν παρέλθει το χρονικό παράθυρο N ημερών χωρίς η τιμή να κινηθεί πέρα από τα προκαθορισμένα όρια (ούτε +3% ούτε −3%), τότε η κίνηση θεωρείται ουδέτερη. Αν η τελική καθαρή μεταβολή τιμής στο διάστημα αυτό είναι ελαφρώς θετική (άνω του +1% αλλά κάτω από το profit target), δίνεται τελικά ετικέτα “Buy”, ενώ αν είναι λίγο αρνητική (χαμηλότερη του −1% χωρίς να χτυπηθεί το stop-loss) δίνεται ετικέτα “Sell”.
    - Στην υλοποίηση επιλέχθηκε ορίζοντας N=5 ημερών, όρια profit-taking

και stop-loss στο ±3%, και ουδέτερη ζώνη ±1%.

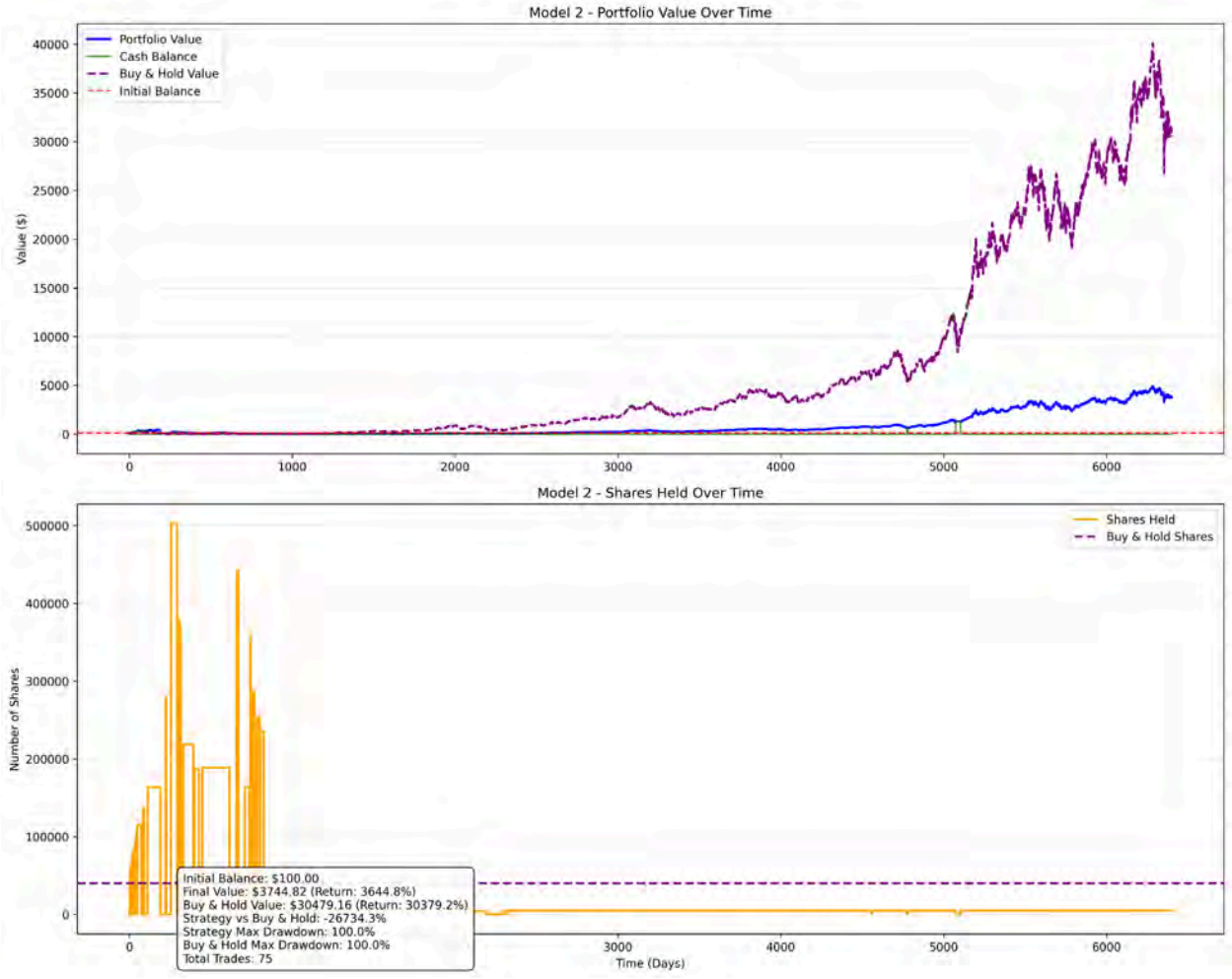
3. **Εκπαίδευση Μοντέλου LSTM**: Το ίδιο πολυ-εισοδικό LSTM μοντέλο εκπαιδεύτηκε στις νέες ετικετοποιημένες σειρές, με παρόμοιες παραμέτρους εκπαίδευσης.

Με βάση αυτό το labeling κρίθηκε και η αξιολόγηση του μοντέλου όσον αφορά την αποτελεσματικότητά του στην πρόβλεψη σωστών αποφάσεων trading. Χρησιμοποιόντας το lstm_evaluator.py κάναμε train και evaluate 10 μοντέλα στην σειρά για να δούμε την αποδοτικότητα τους σε πραγματικές καταστάσεις:



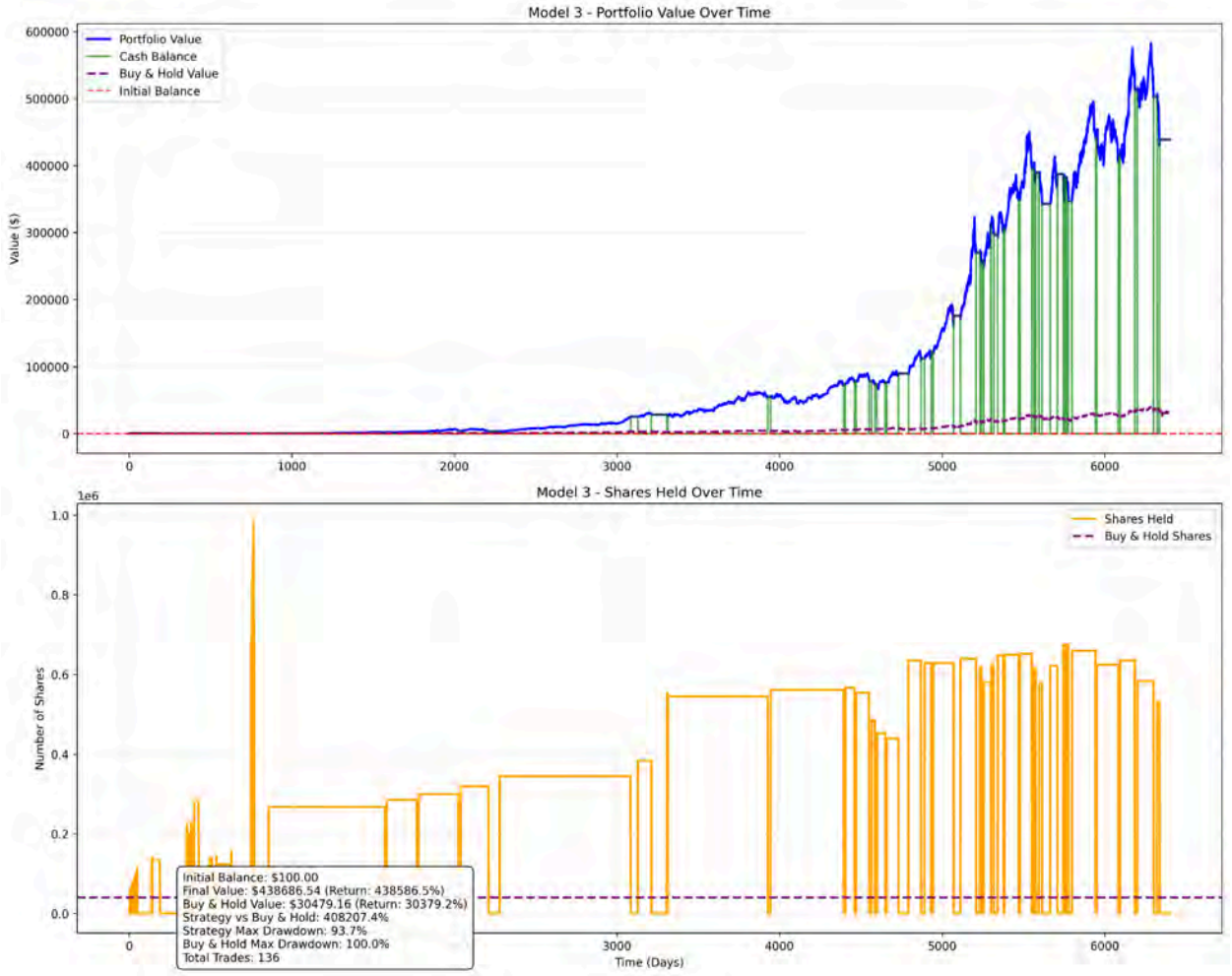
Εικόνα 6.16: Η trading απόδοση του 1ου μοντέλου, σε σύγκριση με την απλή επένδυση (Buy & Hold) 100$ από 1/1/2000.

Στην εικόνα 6.16 παρατηρούμε ότι το μοντέλο παρουσιάζει εξαιρετικά υψηλή απόδοση με συνολική αύξηση κεφαλαίου κατά 125.039,2% (τελική αξία $125.139,24), ξεπερνώντας κατά 94.760,1% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα απεικονίζει την εξέλιξη της αξίας του χαρτοφυλακίου (μπλε γραμμή) σε σύγκριση με το διαθέσιμο ρευστό (γκρι γραμμή) και την αξία της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή), ενώ το κατώτερο γράφημα δείχνει τον αριθμό των μετοχών που κατέχει το μοντέλο σε κάθε χρονική στιγμή. Παρά τη σημαντική μεταβλητότητα και τα υψηλά drawdowns (99,9%), η στρατηγική του μοντέλου αποδεικνύεται ιδιαίτερα αποτελεσματική στην εκμετάλλευση των ευκαιριών της αγοράς.



Εικόνα 6.17: Η trading απόδοση του 2ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

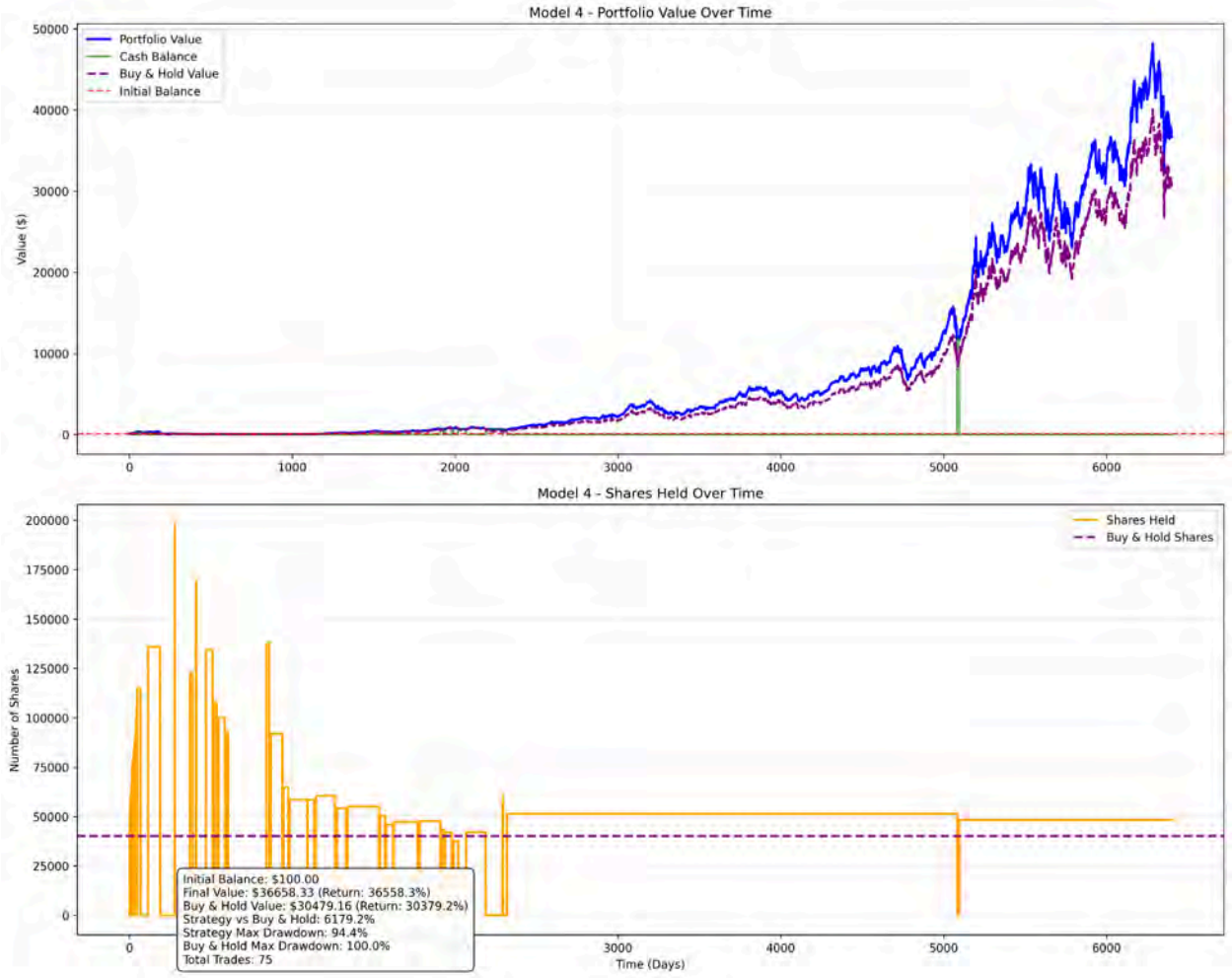
Στην εικόνα 6.17 παρατηρούμε ότι το μοντέλο παρουσιάζει σημαντικά χαμηλότερη απόδοση σε σύγκριση με την παθητική στρατηγική Buy & Hold. Συγκεκριμένα, το μοντέλο επιτυγχάνει συνολική απόδοση 3.544,8% (τελική αξία $3.644,82), υπολειπόμενο κατά 26.734,3% από την Buy & Hold στρατηγική που παρουσιάζει απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα αποκαλύπτει ότι η αξία του χαρτοφυλακίου (μπλε γραμμή) παραμένει σταθερά κάτω από την αξία της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή) καθ' όλη τη διάρκεια της περιόδου. Το κατώτερο γράφημα δείχνει ότι το μοντέλο εκτελεί περιορισμένες συναλλαγές (συνολικά 75), κυρίως στην αρχή της περιόδου, και στη συνέχεια παραμένει σχετικά ανενεργό, γεγονός που συνεισφέρει στην υποδεέστερη απόδοσή του έναντι της παθητικής επενδυτικής στρατηγικής.



Εικόνα 6.18: Η trading απόδοση του 3ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Στην εικόνα 6.18 παρατηρούμε ότι το μοντέλο παρουσιάζει εξαιρετικά υψηλή απόδοση με συνολική αύξηση κεφαλαίου κατά 438.586,5% (τελική αξία $438.686,54),

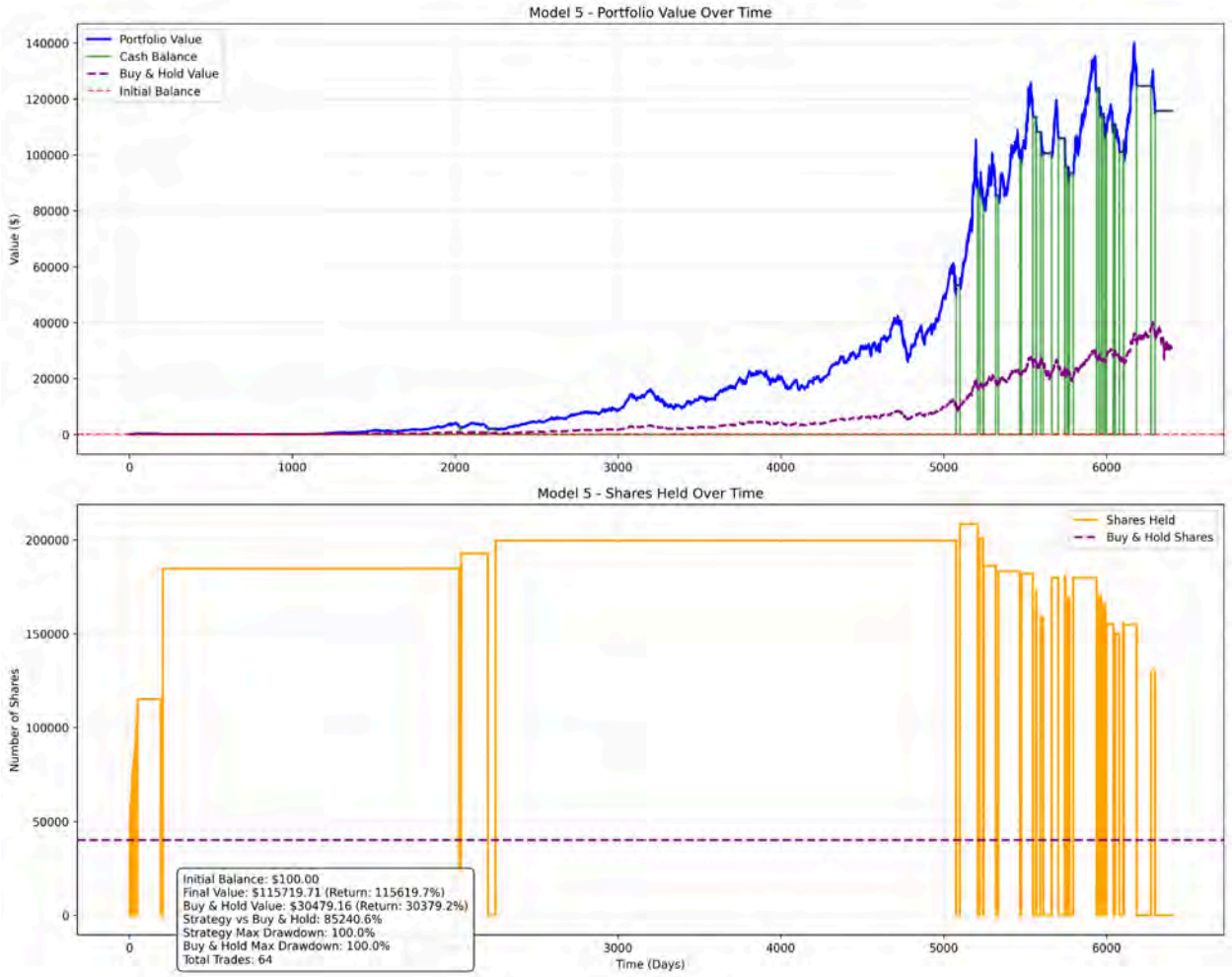
ξεπερνώντας κατά 408.207,4% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα δείχνει μια σταδιακή αύξηση της αξίας του χαρτοφυλακίου (μπλε γραμμή) με ραγδαία επιτάχυνση προς το τέλος της περιόδου, ξεπερνώντας σημαντικά την Buy & Hold στρατηγική (μωβ διακεκομμένη γραμμή). Το κατώτερο γράφημα αποκαλύπτει αυξημένη συναλλακτική δραστηριότητα (συνολικά 136 συναλλαγές) με εντατικοποίηση των κινήσεων στα τελευταία στάδια της περιόδου. Παρά το υψηλό maximum drawdown (93,7%), το μοντέλο αποδεικνύεται ιδιαίτερα αποτελεσματικό στην αξιοποίηση των ευκαιριών της αγοράς.



Εικόνα 6.19: Η trading απόδοση του 4ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Στην εικόνα 6.19 παρατηρούμε ότι το μοντέλο παρουσιάζει μέτρια υπεραπόδοση έναντι της παθητικής στρατηγικής Buy & Hold. Συγκεκριμένα, το μοντέλο επιτυγχάνει συνολική απόδοση 36.558,3% (τελική αξία $36.658,33), ξεπερνώντας κατά 6.179,2%

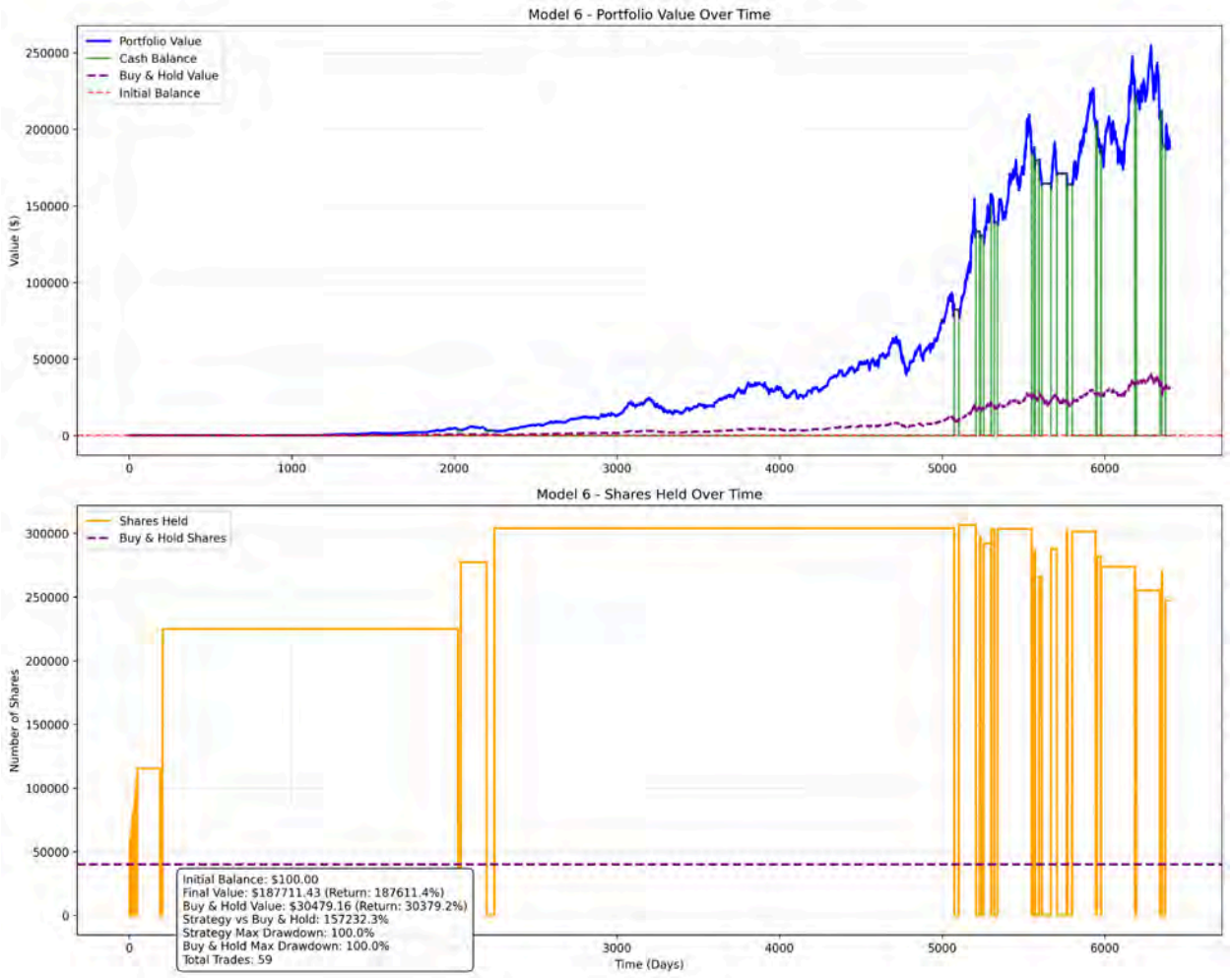
την Buy & Hold στρατηγική που παρουσιάζει απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα αποκαλύπτει ότι οι δύο στρατηγικές κινούνται σχετικά παράλληλα καθ' όλη τη διάρκεια της περιόδου, με το μοντέλο να επιτυγχάνει μια περιορισμένη αλλά σταθερή υπεραπόδοση. Το κατώτερο γράφημα δείχνει έντονη συναλλακτική δραστηριότητα κυρίως στην αρχή της περιόδου (συνολικά 75 συναλλαγές), με σποραδικές κινήσεις στη συνέχεια. Παρά το υψηλό maximum drawdown (94,4%), το μοντέλο καταφέρνει να παραμείνει ανταγωνιστικό και να επιτύχει μια συντηρητική αλλά θετική υπεραπόδοση.



Εικόνα 6.20: Η trading απόδοση του 5ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Στην εικόνα 6.20 παρατηρούμε ότι το μοντέλο παρουσιάζει εξαιρετικά υψηλή απόδοση με συνολική αύξηση κεφαλαίου κατά 115.519,7% (τελική αξία $115.619,71), ξεπερνώντας κατά 85.240,6% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα δείχνει σταδιακή ανοδική πορεία της αξίας του χαρτοφυλακίου (μπλε γραμμή) με ραγδαία επιτάχυνση

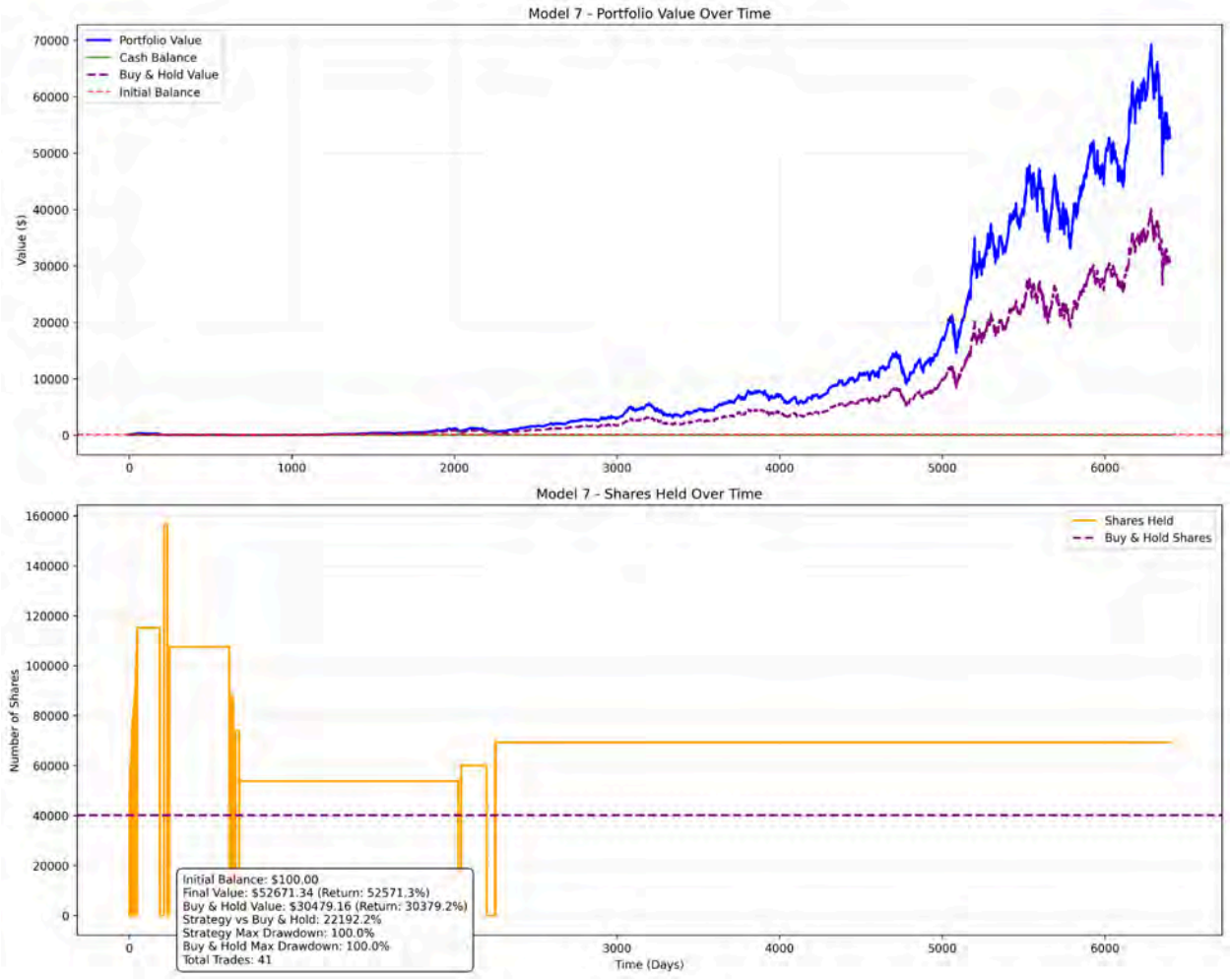
στα τελευταία στάδια της περιόδου, διατηρώντας σημαντική υπεροχή έναντι της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή). Το κατώτερο γράφημα αποκαλύπτει συγκεντρωμένη συναλλακτική δραστηριότητα (συνολικά 64 συναλλαγές) κυρίως στην αρχή και στο τέλος της περιόδου, με μεγάλα χρονικά διαστήματα σχετικής αδράνειας. Παρά τα υψηλά drawdowns (100%), το μοντέλο αποδεικνύεται ιδιαίτερα αποτελεσματικό στην επίτευξη σημαντικών αποδόσεων.



Εικόνα 6.21: Η trading απόδοση του 6ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

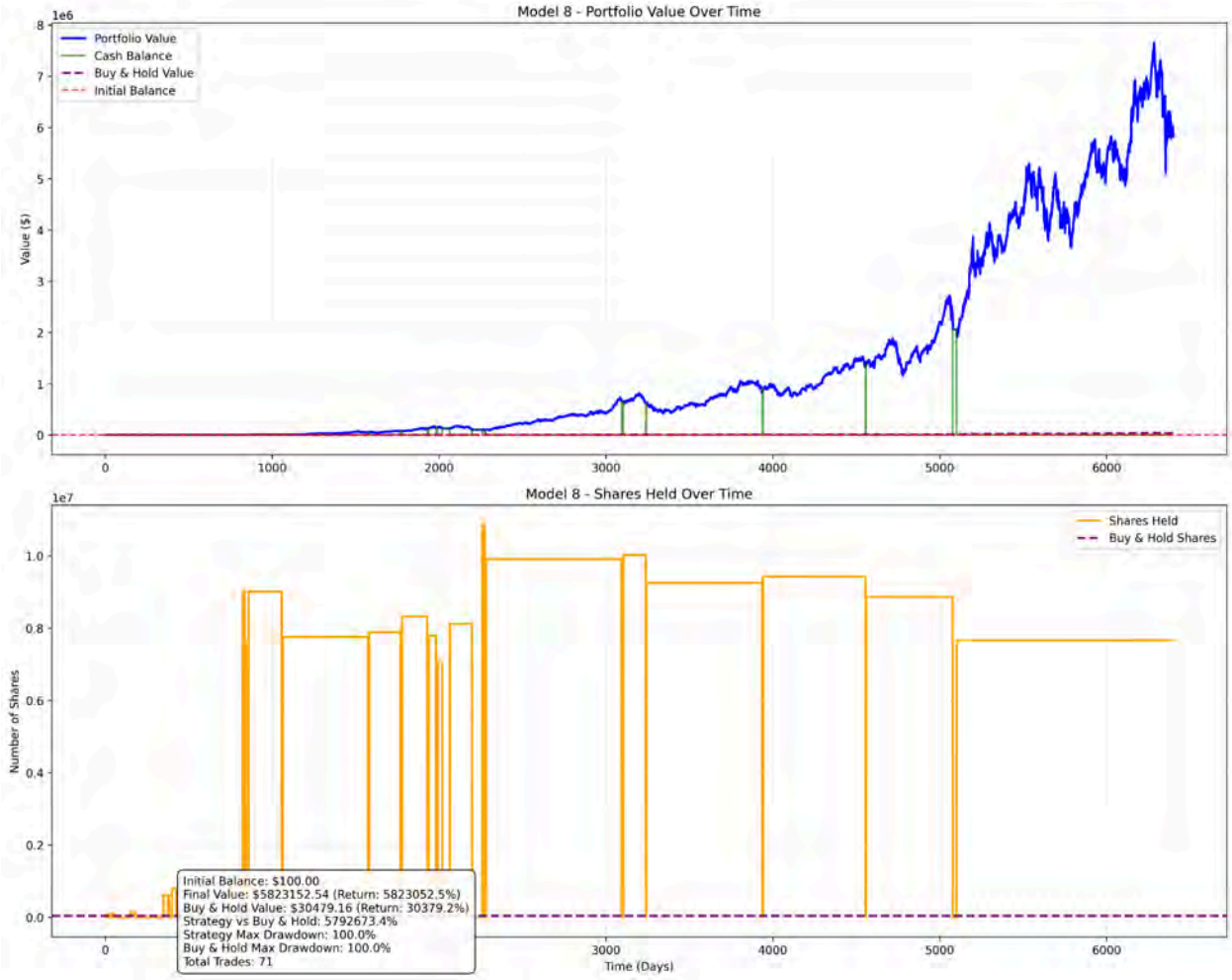
Στην εικόνα 6.21 παρατηρούμε ότι το μοντέλο παρουσιάζει εξαιρετικά υψηλή απόδοση με συνολική αύξηση κεφαλαίου κατά 187.011,4% (τελική αξία $187.211,43), ξεπερνώντας κατά 157.232,3% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα αποκαλύπτει σταδιακή αύξηση της αξίας του χαρτοφυλακίου (μπλε γραμμή) με εντυπωσιακή επιτάχυνση στα τελευταία στάδια της περιόδου, διατηρώντας σταθερή υπεροχή έναντι της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή). Το κατώτερο γράφημα δείχνει στρατηγικές συναλλαγές (συνολικά 59) με έντονη δραστηριότητα στην αρχή της περιόδου, μεγάλη περίοδο σχετικής αδράνειας, και εκ νέου

ενεργοποίηση στα τελευταία στάδια. Παρά τα υψηλά drawdowns (100%), το μοντέλο αποδεικνύεται ιδιαίτερα επιτυχημένο στην αξιοποίηση των μακροπρόθεσμων τάσεων της αγοράς.



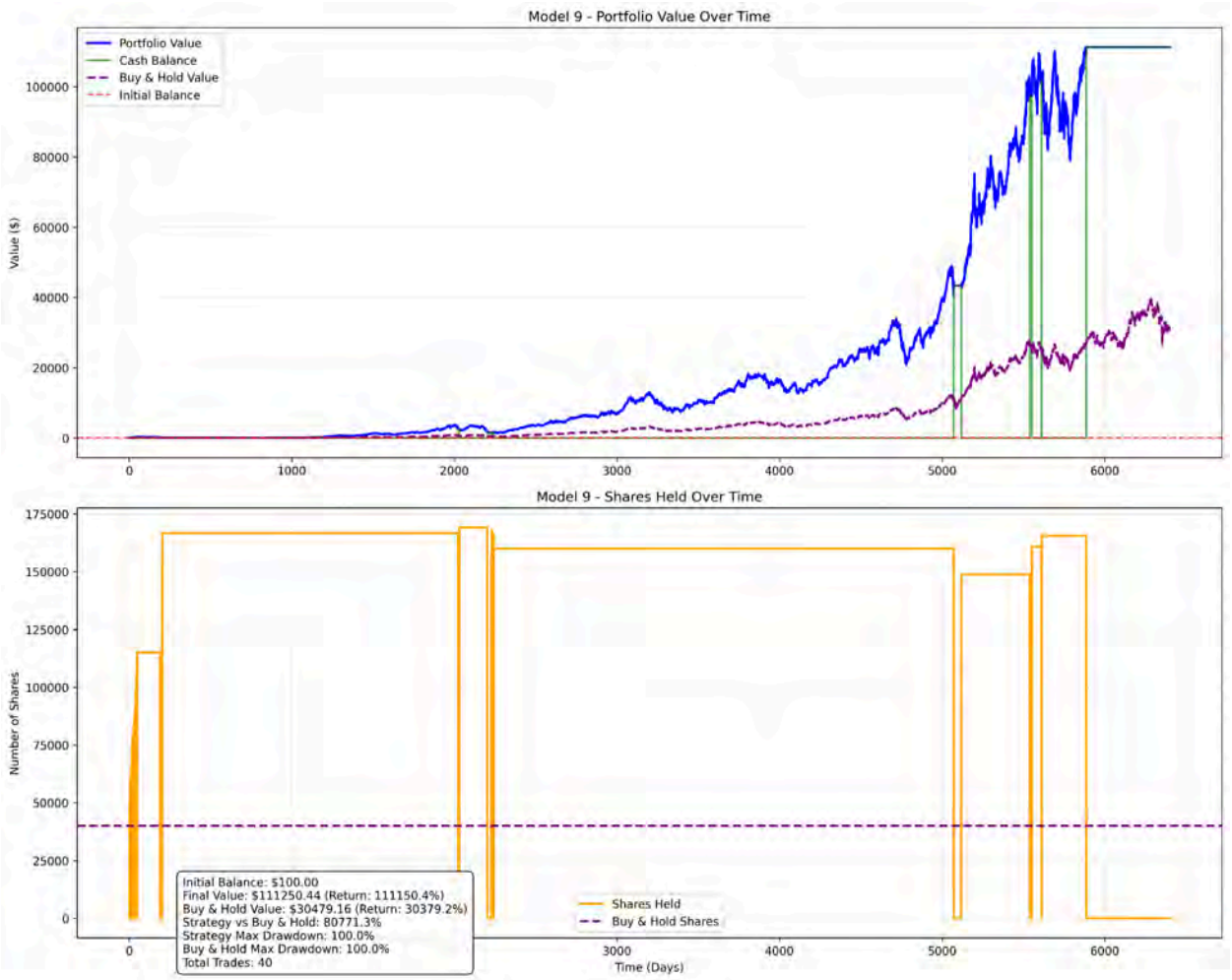
Εικόνα 6.22: Η trading απόδοση του 7ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Στην εικόνα 6.22 παρατηρούμε ότι το μοντέλο παρουσιάζει σημαντική υπεραπόδοση με συνολική αύξηση κεφαλαίου κατά 52.571,3% (τελική αξία $52.671,34), ξεπερνώντας κατά 22.192,2% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα αποκαλύπτει σταδιακή ανοδική πορεία της αξίας του χαρτοφυλακίου (μπλε γραμμή) με εντυπωσιακή επιτάχυνση στα τελευταία στάδια, διατηρώντας σταθερή υπεροχή έναντι της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή). Το κατώτερο γράφημα δείχνει περιορισμένες αλλά στρατηγικές συναλλαγές (συνολικά 41), με έντονη δραστηριότητα στην αρχή της περιόδου και στη συνέχεια μακροπρόθεσμη διατήρηση θέσεων. Παρά τα υψηλά drawdowns (100%), το μοντέλο αποδεικνύεται αποτελεσματικό στην επίτευξη σταθερών και σημαντικών αποδόσεων.



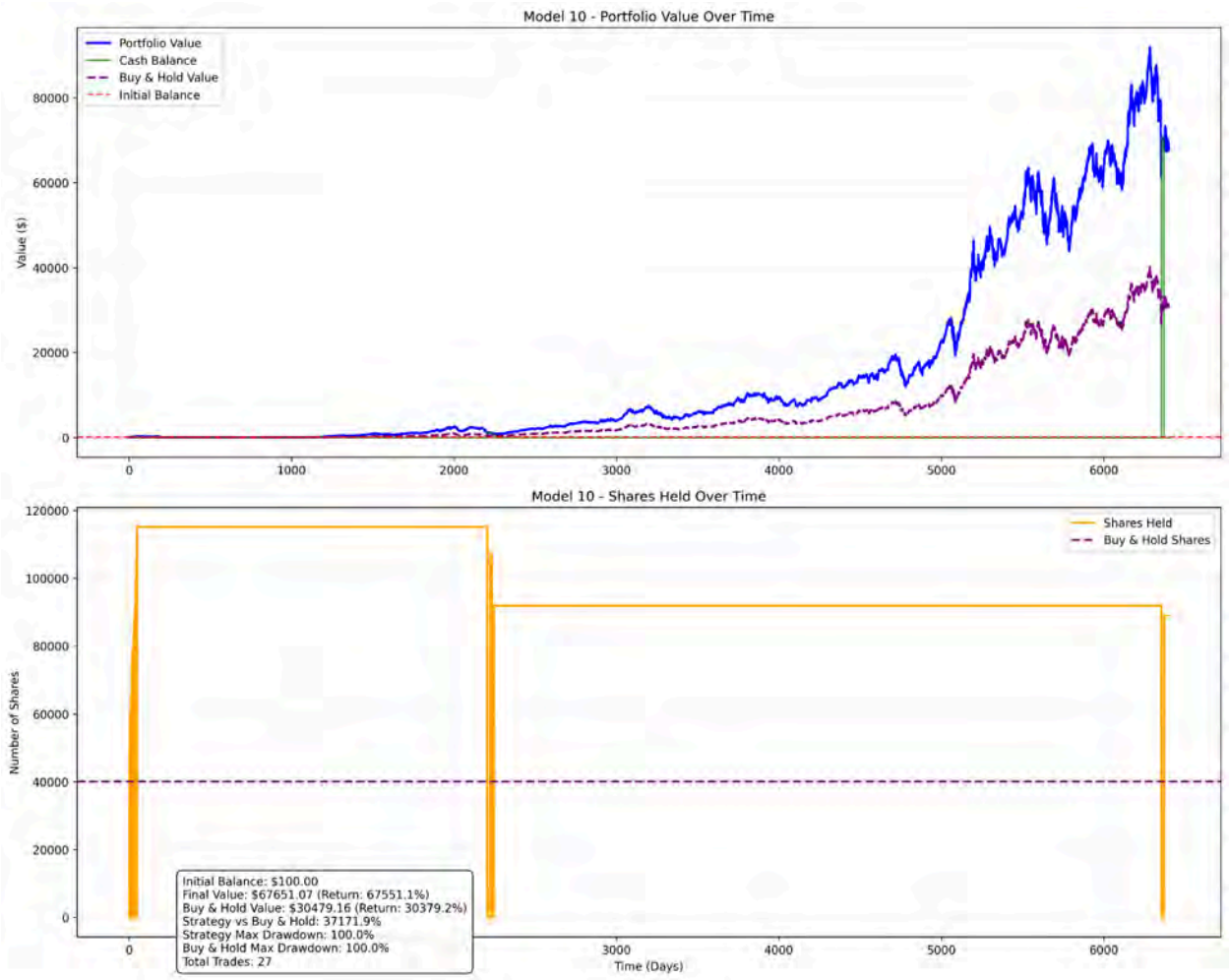
Εικόνα 6.23: Η trading απόδοση του 8ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Στην εικόνα 6.23 παρατηρούμε ότι το μοντέλο παρουσιάζει σημαντική υπεραπόδοση με συνολική αύξηση κεφαλαίου κατά 52.571,3% (τελική αξία $52.671,34), ξεπερνώντας κατά 22.192,2% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα αποκαλύπτει μια σταδιακή αλλά σταθερή υπεροχή του μοντέλου (μπλε γραμμή) έναντι της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή), με τη μεγαλύτερη απόκλιση να παρατηρείται στα τελευταία στάδια της περιόδου. Το κατώτερο γράφημα δείχνει περιορισμένη συναλλακτική δραστηριότητα (συνολικά 41 συναλλαγές) που συγκεντρώνεται κυρίως στην αρχή της περιόδου, με μακρές περιόδους διατήρησης θέσεων. Παρά τα υψηλά drawdowns (100%), το μοντέλο επιδεικνύει μια πιο συντηρητική αλλά αποτελεσματική προσέγγιση στη διαχείριση του χαρτοφυλακίου.



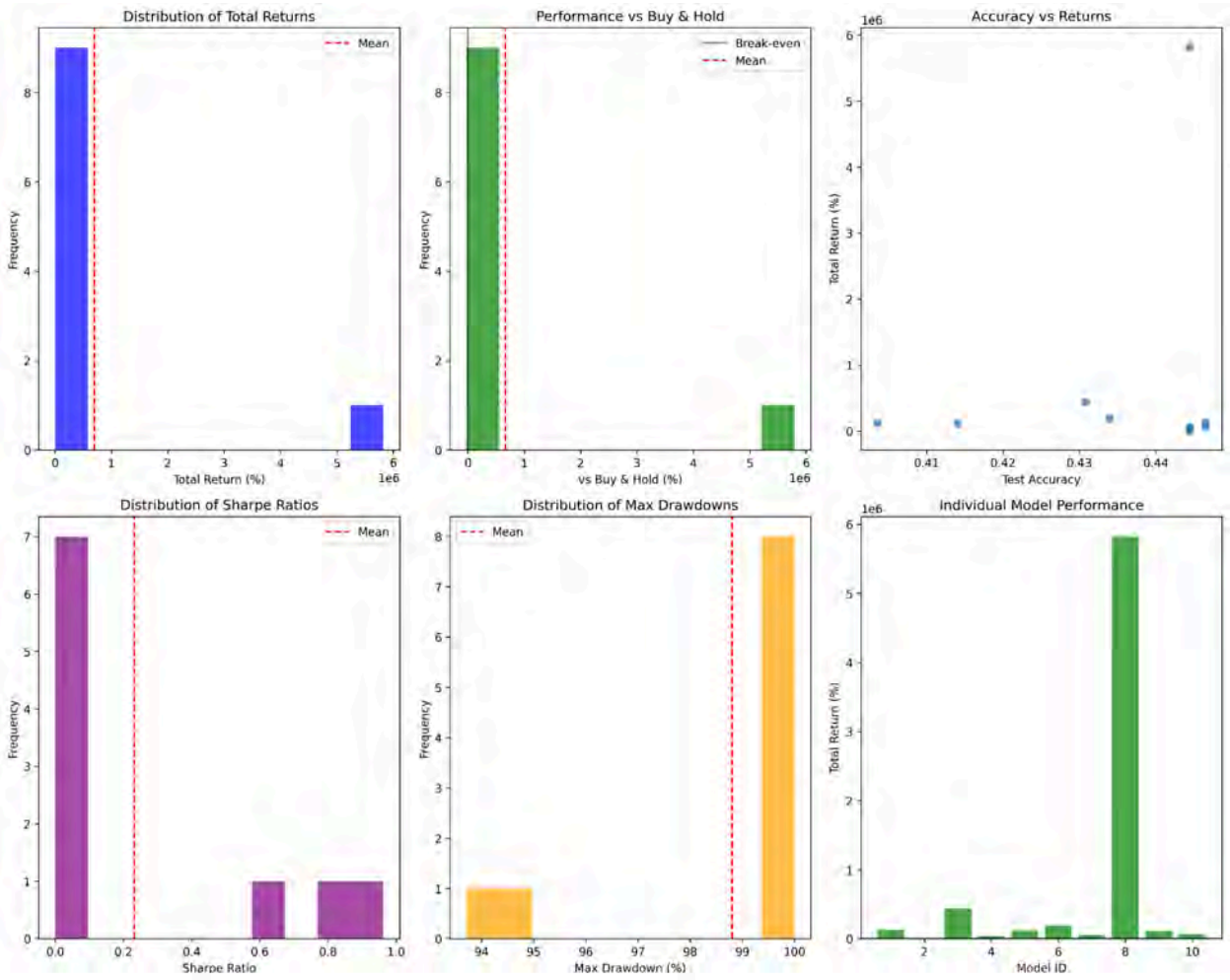
Εικόνα 6.24: Η trading απόδοση του 9ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Παρατηρούμε ότι το μοντέλο παρουσιάζει εξαιρετικά υψηλή απόδοση με συνολική αύξηση κεφαλαίου κατά 111.150,4% (τελική αξία $111.250,44), ξεπερνώντας κατά 80.771,3% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα αποκαλύπτει μια εντυπωσιακή ανοδική πορεία της αξίας του χαρτοφυλακίου (μπλε γραμμή) με ραγδαία επιτάχυνση στα τελευταία στάδια της περιόδου, διατηρώντας σταθερή και αυξανόμενη υπεροχή έναντι της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή). Το κατώτερο γράφημα δείχνει στρατηγικές και περιορισμένες συναλλαγές (συνολικά 40) που εκτελούνται σε συγκεκριμένες χρονικές περιόδους, με μακρές φάσεις διατήρησης θέσεων. Παρά τα υψηλά drawdowns (100%), το μοντέλο αποδεικνύεται εξαιρετικά αποτελεσματικό στην αξιοποίηση των μακροπρόθεσμων ευκαιριών της αγοράς.



Εικόνα 6.25: Η trading απόδοση του 10ου μοντέλου, σε σύγκριση με την απλή επένδυση 100$ από 1/1/2000.

Παρατηρούμε ότι το μοντέλο στην εικόνα 6.25 παρουσιάζει εξαιρετικά υψηλή απόδοση με συνολική αύξηση κεφαλαίου κατά 67.551,1% (τελική αξία $67.651,07), ξεπερνώντας κατά 37.171,9% την παθητική στρατηγική Buy & Hold που επέτυχε απόδοση 30.379,2% (τελική αξία $30.479,16). Το ανώτερο γράφημα δείχνει σταδιακή και σταθερή ανοδική πορεία της αξίας του χαρτοφυλακίου (μπλε γραμμή) με εντυπωσιακή επιτάχυνση στα τελευταία στάδια της περιόδου, διατηρώντας συνεχή υπεροχή έναντι της Buy & Hold στρατηγικής (μωβ διακεκομμένη γραμμή). Το κατώτερο γράφημα αποκαλύπτει πολύ περιορισμένες συναλλαγές (συνολικά 27), με συγκεκριμένες παρεμβάσεις στην αρχή, μέση και τέλος της περιόδου, υποδεικνύοντας μια εξαιρετικά επιλεκτική και αποτελεσματική στρατηγική. Παρά τα υψηλά drawdowns (100%), το μοντέλο αποδεικνύεται ιδιαίτερα επιτυχημένο στην επίτευξη σημαντικών αποδόσεων με ελάχιστες συναλλαγές.



Εικόνα 6.26: Συνολική αξιολόγηση της απόδοσης των 10 LSTM μοντέλων.

Στην εικόνα 6.26 παρουσιάζεται μια ολοκληρωμένη ανάλυση της απόδοσης των 10 LSTM μοντέλων που αναπτύχθηκαν και δοκιμάστηκαν στο πλαίσιο αυτής της μελέτης. Από τα αποτελέσματα προκύπτει ότι **9 από τα 10 μοντέλα** επέτυχαν καλύτερη απόδοση από την παθητική στρατηγική Buy & Hold, γεγονός που αποδεικνύει την αποτελεσματικότητα του multi-input LSTM στη διαδικασία trading.

Η κατανομή των συνολικών αποδόσεων (επάνω αριστερά) δείχνει ότι η πλειονότητα των μοντέλων συγκεντρώνεται γύρω από εξαιρετικά υψηλές αποδόσεις, με μόνο ένα μοντέλο να παρουσιάζει σημαντικά χαμηλότερη απόδοση. Το γράφημα απόδοσης έναντι Buy & Hold (επάνω μέση) επιβεβαιώνει ότι σχεδόν όλα τα μοντέλα ξεπερνούν την παθητική στρατηγική.

Το Sharpe Ratio είναι ένας δείκτης που μετρά την απόδοση ενός επενδυτικού χαρτοφυλακίου προσαρμοσμένη για τον κίνδυνο. Ο πλήρης τύπος του Sharpe Ratio είναι:

$$Sharpe\ Ratio\ = \frac{R_p - R_f}{\sigma_p}$$

Όπου:

R_p = η μέση απόδοση του χαρτοφυλακίου

R_f = το risk-free rate (απόδοση χωρίς κίνδυνο, συνήθως το επιτόκιο κρατικών ομολόγων)

σ_p = η τυπική απόκλιση των αποδόσεων του χαρτοφυλακίου

Υψηλότερες τιμές Sharpe Ratio υποδηλώνουν καλύτερη απόδοση προσαρμοσμένη για τον κίνδυνο. Από την κατανομή (κάτω αριστερά), παρατηρούμε ότι τα περισσότερα μοντέλα παρουσιάζουν χαμηλά Sharpe Ratios, γεγονός που υποδηλώνει υψηλή μεταβλητότητα σε σχέση με την απόδοση. Αυτό σημαίνει ότι τα μοντέλα είναι αποτελεσματικά στην πρόβλεψη και την επίτευξη υψηλών αποδόσεων, αλλά δεν είναι επενδυτικά αποδοτικά από την άποψη της διαχείρισης κινδύνου. Ένας συντηρητικός επενδυτής θα προτιμούσε μικρότερες αλλά πιο σταθερές αποδόσεις.

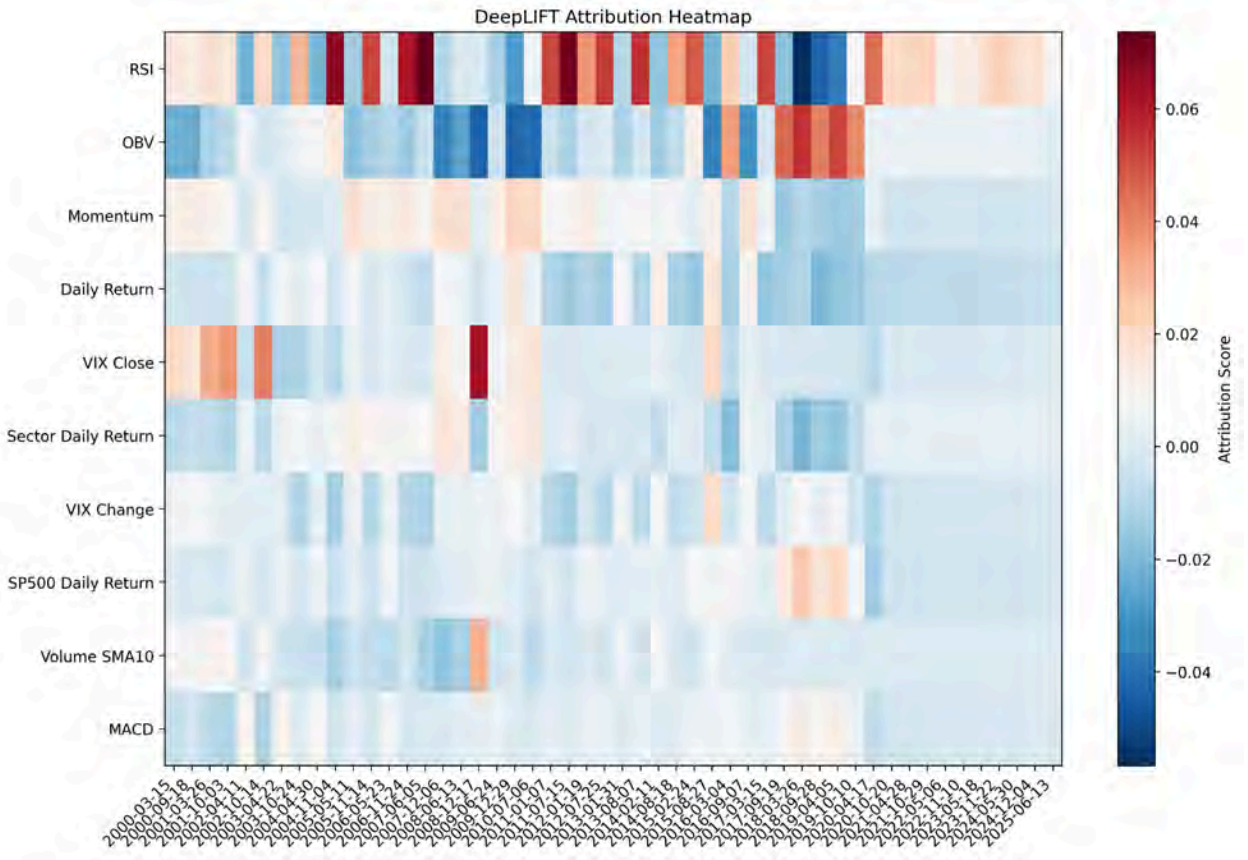
Το Max Drawdown δείχνει τη μέγιστη πτώση της αξίας του χαρτοφυλακίου από μια κορυφή σε μια πτώση κατά τη διάρκεια μιας συγκεκριμένης περιόδου. Είναι ένας κρίσιμος δείκτης κινδύνου που δείχνει τη μέγιστη ζημία που θα μπορούσε να αντιμετωπίσει ένας επενδυτής. Από την κατανομή (κάτω μέση), φαίνεται ότι τα περισσότερα μοντέλα παρουσιάζουν πολύ υψηλά Max Drawdowns κοντά στο 100%, υποδεικνύοντας σημαντικό κίνδυνο παρά τις υψηλές τελικές αποδόσεις.

Το μοντέλο 8, το οποίο παρουσίασε την καλύτερη συνολική απόδοση, επιλέχθηκε για τα πειράματα εξηγησιμότητας που ακολουθούν, προκειμένου να κατανοήσουμε καλύτερα τους παράγοντες που συνεισφέρουν στην επιτυχημένη του απόδοση.

6.2.2 Ευρήματα Πειράματος Triple Barrier Labeling

Η αξιολόγηση των αποτελεσμάτων από το 8ο μοντέλο LSTM που είδαμε νωρίτερα εκπαιδευμένο με τη στρατηγική Triple Barrier αξιολογήθηκε με τις πέντε μεθόδους XAI, αποκαλύπτοντας διαφορετικά πρότυπα συμπεριφοράς και σημαντικότητας χαρακτηριστικών σε σχέση με το πρώτο πείραμα.

6.2.2.1 DeepLIFT

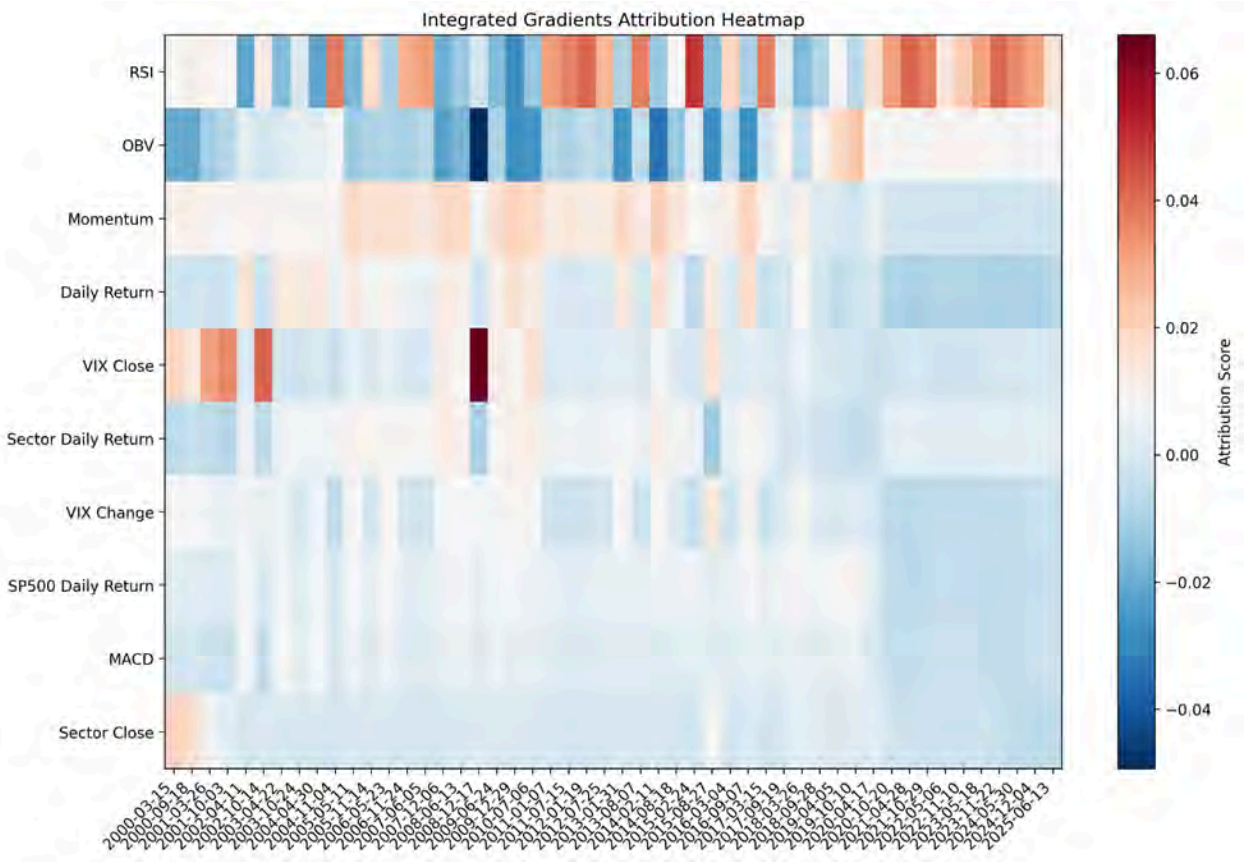


Εικόνα 6.27: Heatmap σημαντικότητας DeepLIFT στο Πείραμα με Triple Barrier Labeling.

Τα αποτελέσματα του DeepLift heatmap όπως φαίνονται στην εικόνα 6.27 αποκαλύπτουν μια σημαντική μεταβολή στην ιεραρχία σημαντικότητας των χαρακτηριστικών σε σύγκριση με το custom labeling. Ο RSI, παρότι διατηρεί υψηλή σημαντικότητα, δεν παρουσιάζει πλέον απόλυτη κυριαρχία, ενώ το On-Balance Volume (OBV) αναδεικνύεται ως ο δεύτερος σημαντικότερος παράγοντας με έντονες

αρνητικές συνεισφορές. Οι momentum indicators εμφανίζουν αυξημένη σημαντικότητα, υποδεικνύοντας στροφή προς βραχυπρόθεσμες τάσεις αντί για mean reversion στρατηγικές. Η σημαντικότητα κατανέμεται πιο ισομερώς μεταξύ των χαρακτηριστικών με αυξημένη χρονική διακύμανση, γεγονός που υποδηλώνει ότι σε πραγματικές συνθήκες trading οι momentum-based στρατηγικές είναι πιο αποτελεσματικές, ενώ η απουσία volume confirmation (αρνητικές συνεισφορές OBV) αποτελεί κρίσιμο σήμα κινδύνου. Επίσης φαίνεται ότι σε μη τεχνητές συνθήκες όλοι οι παράγοντες εμπλέκονται περισσότερο στις αποφάσεις του μοντέλου.

6.2.2.2 Integrated Gradients

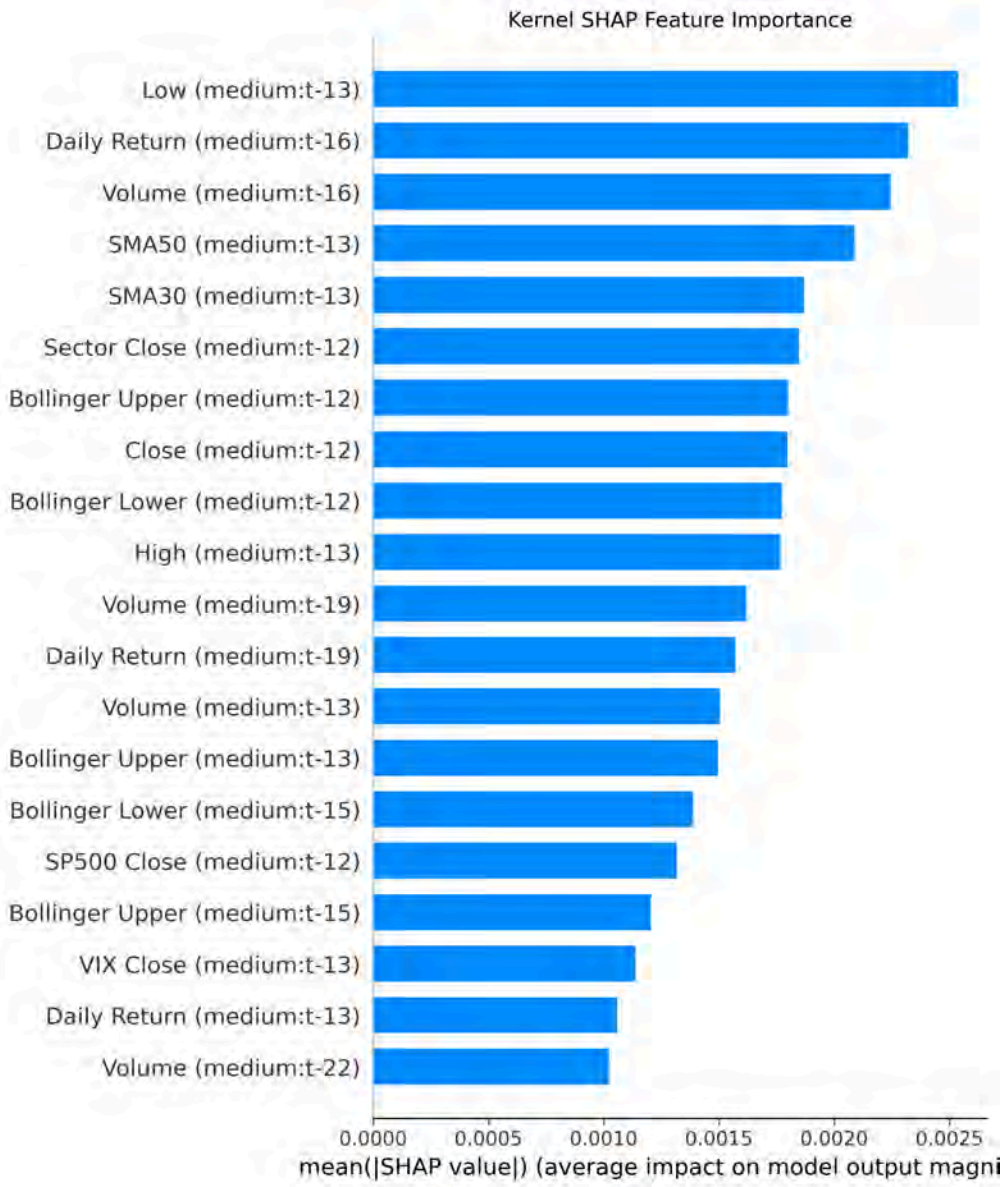


Εικόνα 6.28: Heatmap σημαντικότητας Integrated Gradients στο Πείραμα με Triple Barrier Labeling.

Τα αποτελέσματα του Integrated Gradients heatmap όπως φαίνονται στην εικόνα 6.28 αποκαλύπτουν σύνθετα attribution patterns με τον RSI να εμφανίζει τη μεγαλύτερη διακύμανση μεταξύ ισχυρών θετικών και αρνητικών τιμών, ενώ το On-Balance Volume (OBV) παρουσιάζει κυρίως αρνητικές συνεισφορές,

λειτουργώντας ως αντιφατικός indicator. Η χρονική κατανομή των RSI attributions αποδεικνύει ότι το μοντέλο εκμεταλλεύεται το directional momentum για την πρόβλεψη του ποιου barrier θα επιτευχθεί πρώτο, με την ετερογενή temporal distribution να υποδηλώνει ότι η triple barrier strategy επιτρέπει δυναμική προσαρμογή σε διαφορετικά market regimes. Ιδιαίτερα σημαντική είναι η χρονοϊστορική ανάλυση που αποκαλύπτει συγκεκριμένες περιόδους κρίσης (2008-2009, 2020-2021) όπου τα VIX-based features αποκτούν εξαιρετικά υψηλή σημαντικότητα, αντικατοπτρίζοντας την ικανότητα της στρατηγικής να προσαρμόζεται αποτελεσματικά σε περιόδους αυξημένης αβεβαιότητας.

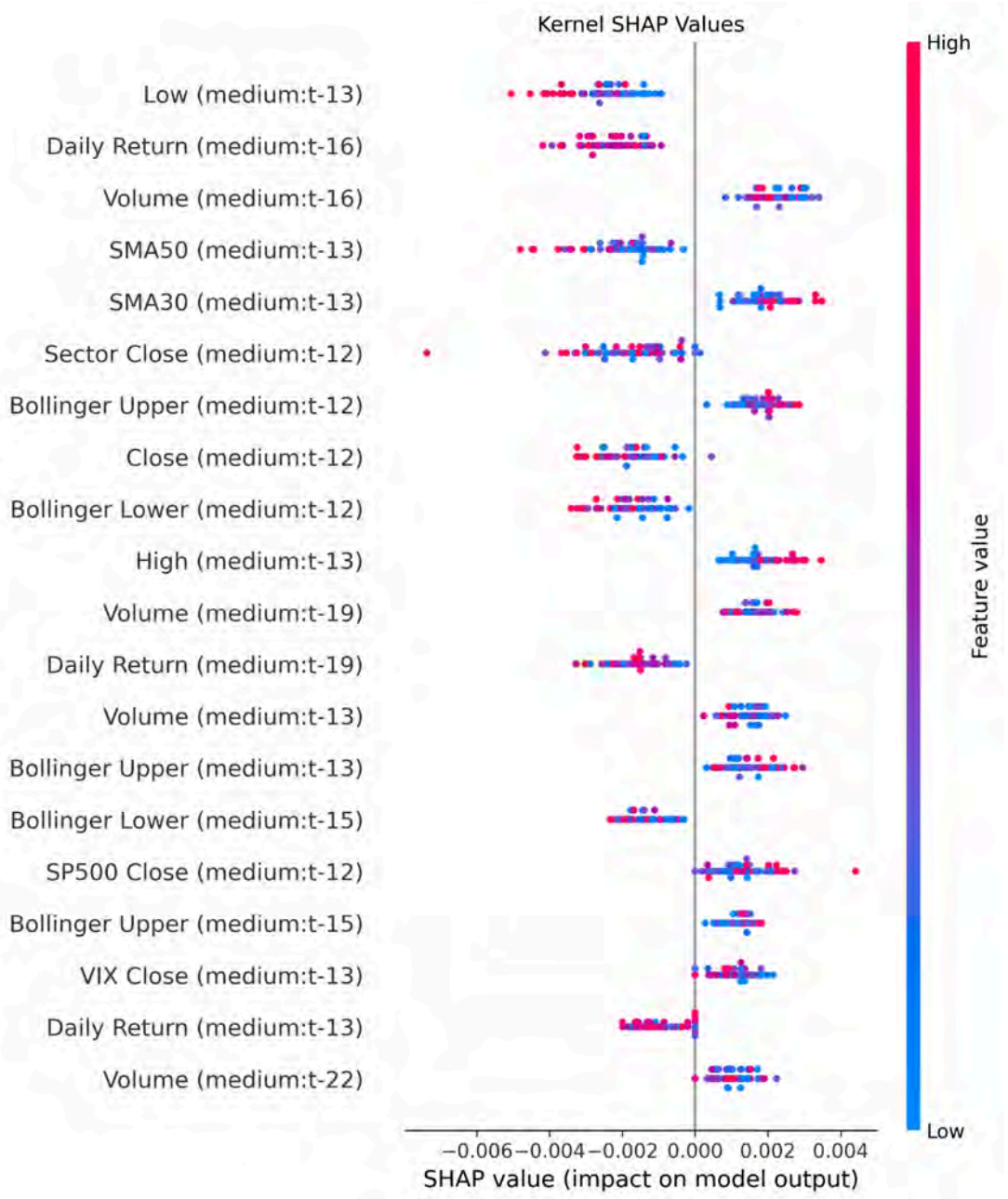
6.2.2.3 Kernel SHAP



Εικόνα 6.29: SHAP Feature Importance Graph στο Πείραμα με Triple Barrier Labeling.

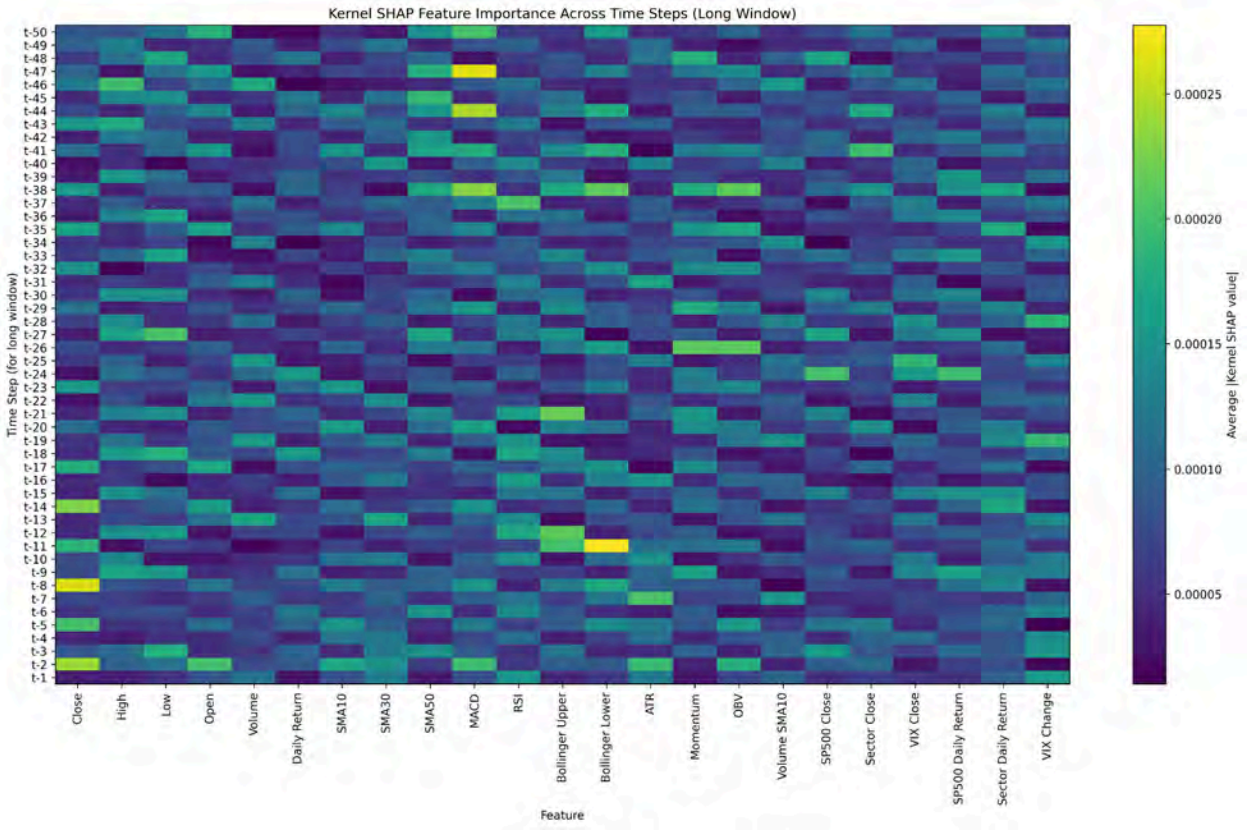
Τα αποτελέσματα του Kernel SHAP στην εικόνα 6.29 αποκαλύπτουν μια διαφορετική όψη εξηγησιμότητας σε σχέση με τα παραπάνω. Τα χαρακτηριστικά Low, Daily Return, Volume κυριαρχούν στην σημαντικότητα, με το Low price στη χρονική στιγμή t-13 να αναδεικνύεται ως το κρισιμότερο χαρακτηριστικό. Οι moving averages (SMA30, SMA50) και τα Bollinger Bands εμφανίζουν επίσης υψηλή σημαντικότητα, υποδηλώνοντας ότι το μοντέλο βασίζεται σε τεχνικούς δείκτες τάσης και μεταβλητότητας. Αξιοσημείωτη είναι η χρονική διάσταση της σημαντικότητας, με

διαφορετικά lag periods (t-12 έως t-22) να συνεισφέρουν διαφορετικά, γεγονός που επιβεβαιώνει την ικανότητα του LSTM να αξιοποιεί ιστορικές πληροφορίες. Τα macro-economic indicators (VIX, S&P500) παρουσιάζουν χαμηλότερη σημαντικότητα, δείχνοντας ότι το μοντέλο εστιάζει περισσότερο σε stock-specific χαρακτηριστικά παρά σε συστημικούς παράγοντες κινδύνου. Επίης οι πιο σημαντικοί παράγοντες εξακολουθούν να προκύπτουν από το medium window του LSTM.



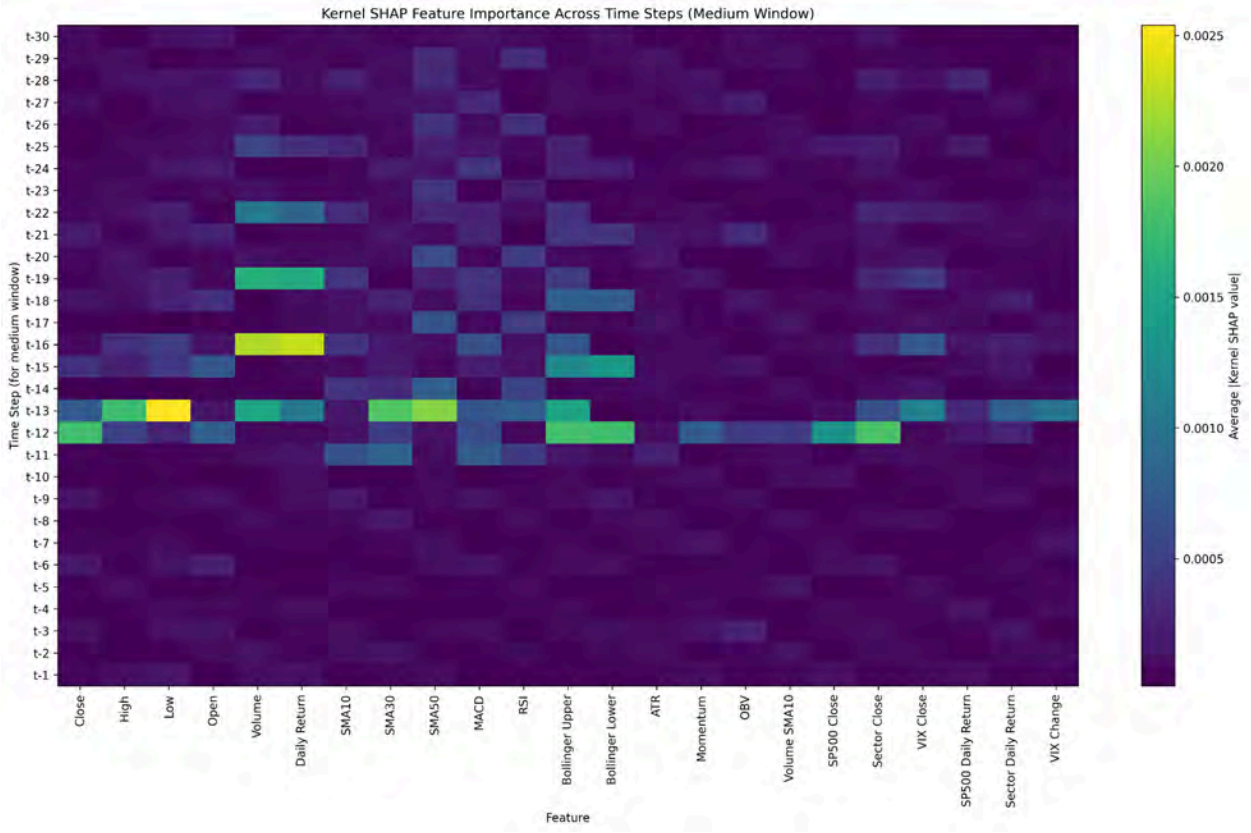
Εικόνα 6.30: SHAP Summary Plot στο Πείραμα με Triple Barrier Labeling.

Το Kernel SHAP summary plot αποκαλύπτει τη σχέση μεταξύ της τιμής των χαρακτηριστικών και της επίδρασής τους στο μοντέλο, κάτι που δεν φαινόταν στο προηγούμενο γράφημα σημαντικότητας. Παρατηρούμε ότι υψηλές τιμές των SMA30 και του Volume συνδέονται με θετικές SHAP values (κόκκινα σημεία στα δεξιά), υποδηλώνοντας bullish signals. Αντίθετα, υψηλές τιμές του Daily Return σε διάφορες χρονικές στιγμές εμφανίζουν μικτές επιδράσεις, γεγονός που υποδηλώνει μη-γραμμικές σχέσεις. Η διασπορά των SHAP values αποκαλύπτει ότι ορισμένα χαρακτηριστικά (όπως το Low price) έχουν περισσότερο συνεπή επίδραση, ενώ άλλα παρουσιάζουν μεγάλη μεταβλητότητα ανάλογα με τις συνθήκες της αγοράς, επιβεβαιώνοντας την προσαρμοστικότητα του μοντέλου σε διαφορετικά market regimes.

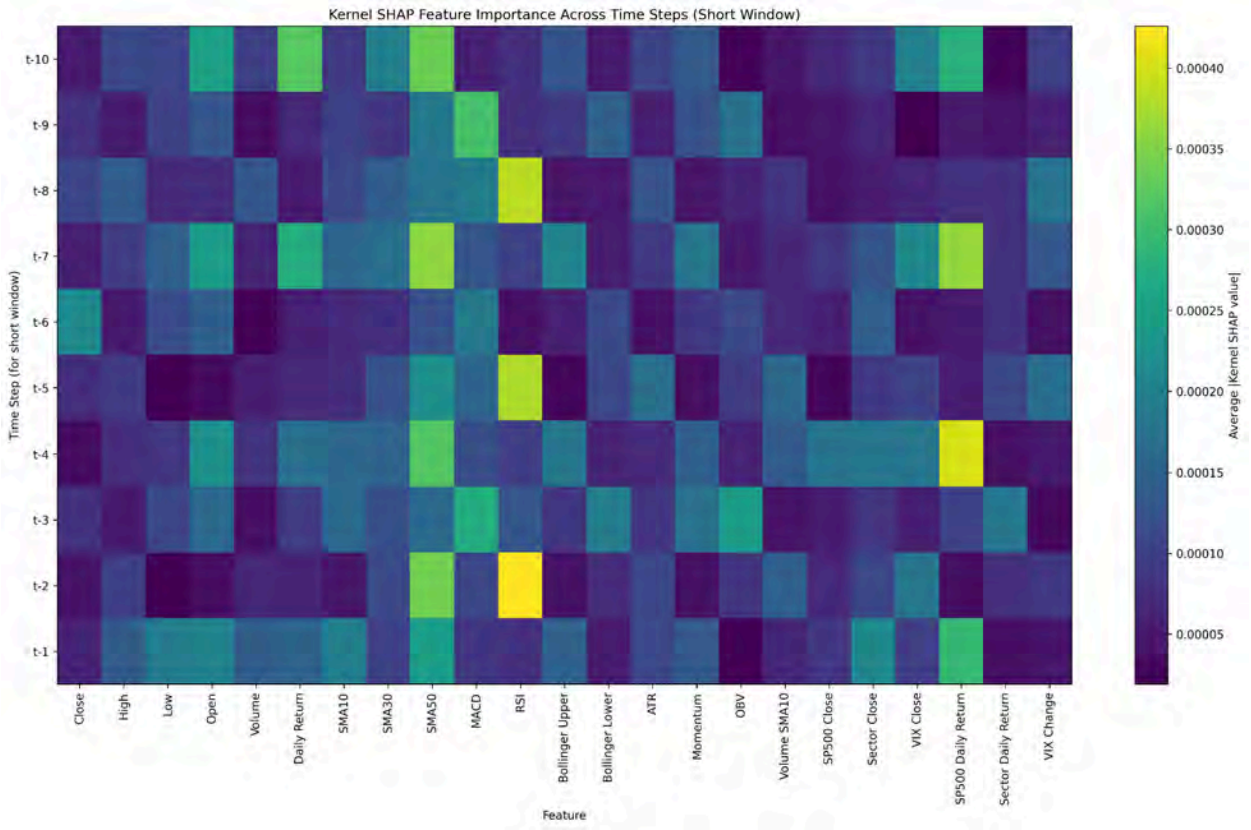


Εικόνα 6.31: SHAP Time Importance Plot (50 days window) στο Πείραμα με με Triple Barrier

Labeling.



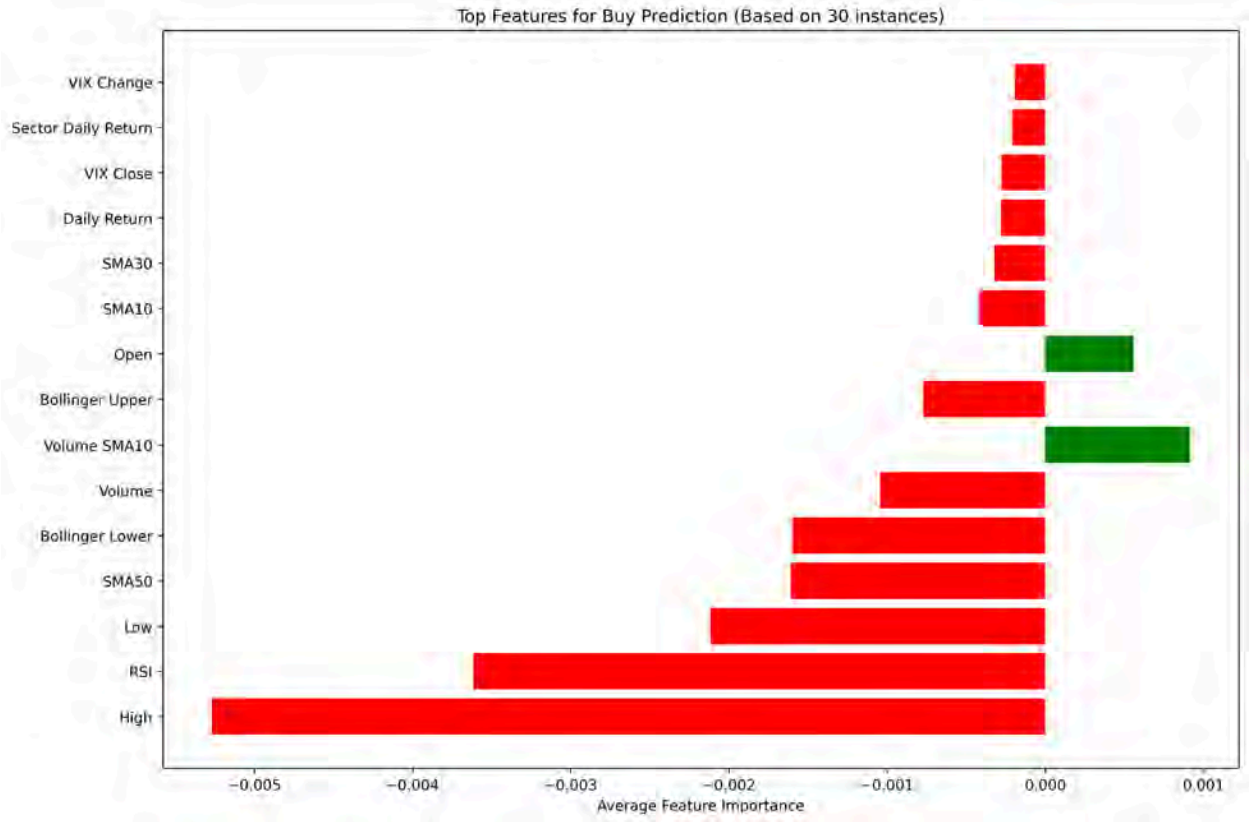
Εικόνα 6.32: SHAP Time Importance Plot (30 days window) στο Πείραμα με με Triple Barrier Labeling



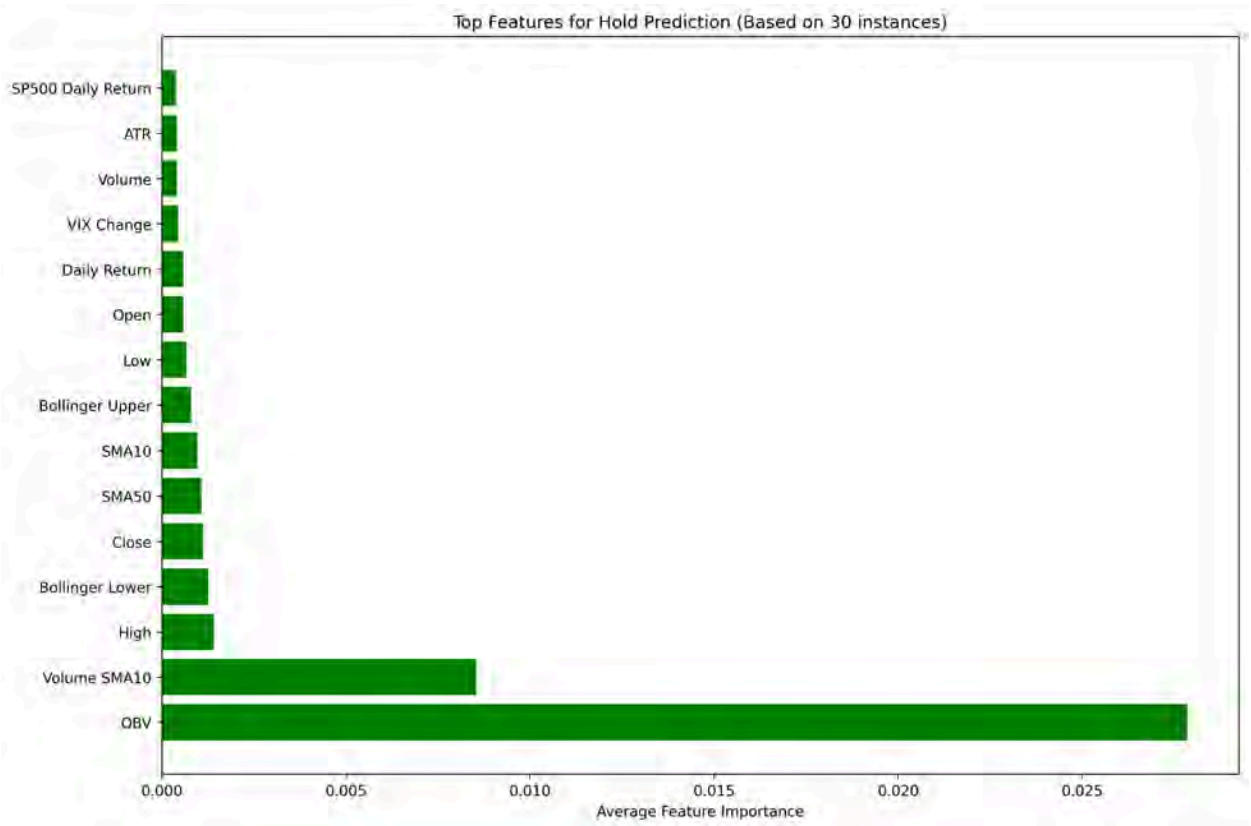
Εικόνα 6.33: SHAP Time Importance Plot (10 days window) στο Πείραμα με Triple Barrier Labeling.

Τα τρία heatmaps αποκαλύπτουν πώς το μοντέλο "κοιτάζει" πίσω στο χρόνο για να πάρει αποφάσεις trading. Στο Long Window (50 ημέρες πίσω) βλέπουμε ότι το μοντέλο εξετάζει μακροπρόθεσμα patterns με σχετικά ομοιόμορφη κατανομή σημαντικότητας, δηλαδή "θυμάται" γεγονότα από μακριά στο παρελθόν. Στο Medium Window (30 ημέρες πίσω) η δραστηριότητα συγκεντρώνεται στις μεσαίες περιόδους, ιδιαίτερα γύρω από τις στιγμές t-13 έως t-16, όπου φαίνονται τα πιο "έντονα" κίτρινα και πράσινα χρώματα. Στο Short Window (10 ημέρες πίσω) παρατηρούμε έντονη δραστηριότητα στις πρόσφατες στιγμές (t-1 έως t-4), δηλαδή το μοντέλο δίνει μεγάλη βαρύτητα στις τελευταίες κινήσεις των τιμών. Αυτό δείχνει ότι το LSTM έχει μάθει να συνδυάζει μακροπρόθεσμη "μνήμη" με άμεση αντίδραση στις πρόσφατες αλλαγές, όπως ακριβώς θα έκανε ένας έμπειρος trader που εξετάζει και τις μακροχρόνιες τάσεις και τις πρόσφατες κινήσεις της αγοράς.

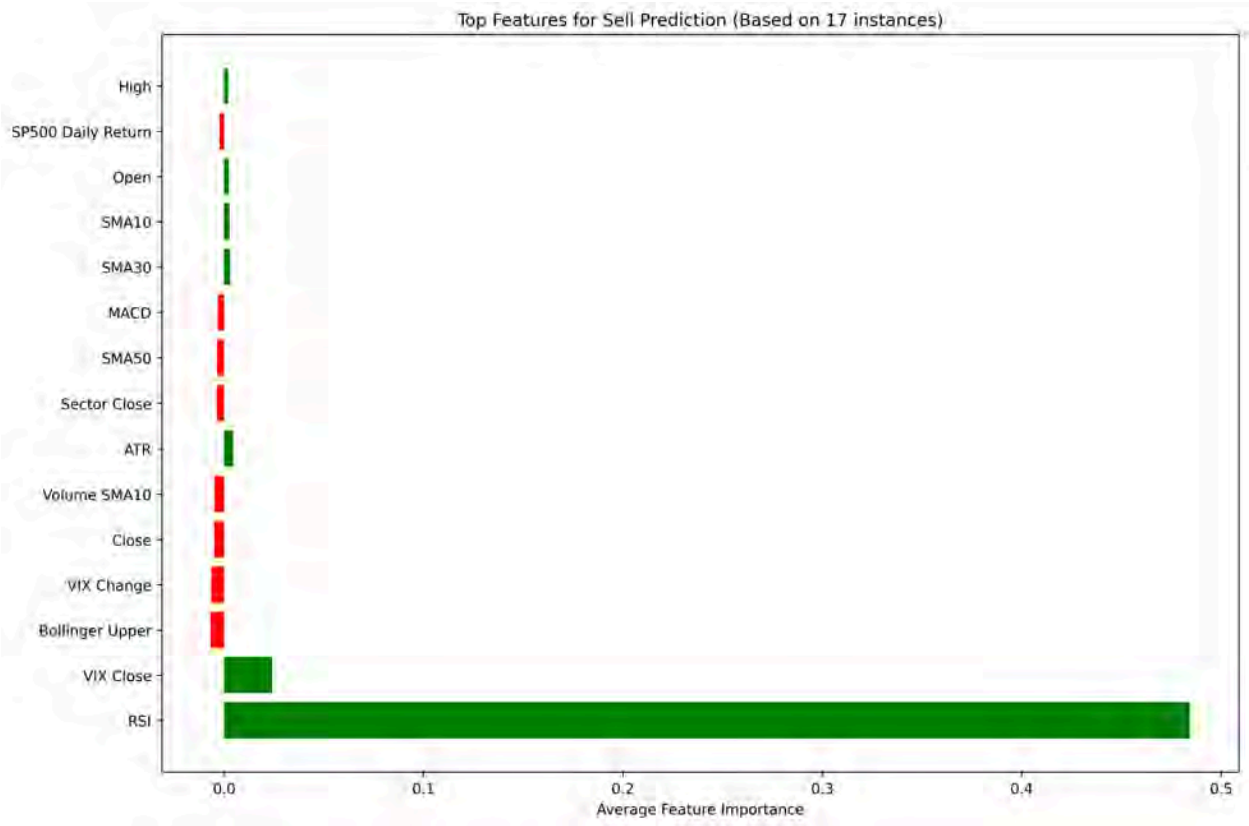
6.2.2.4 LIME



Εικόνα 6.34: LIME Aggregate Importance Plot for Buy action στο Πείραμα με Triple Barrier Labeling.

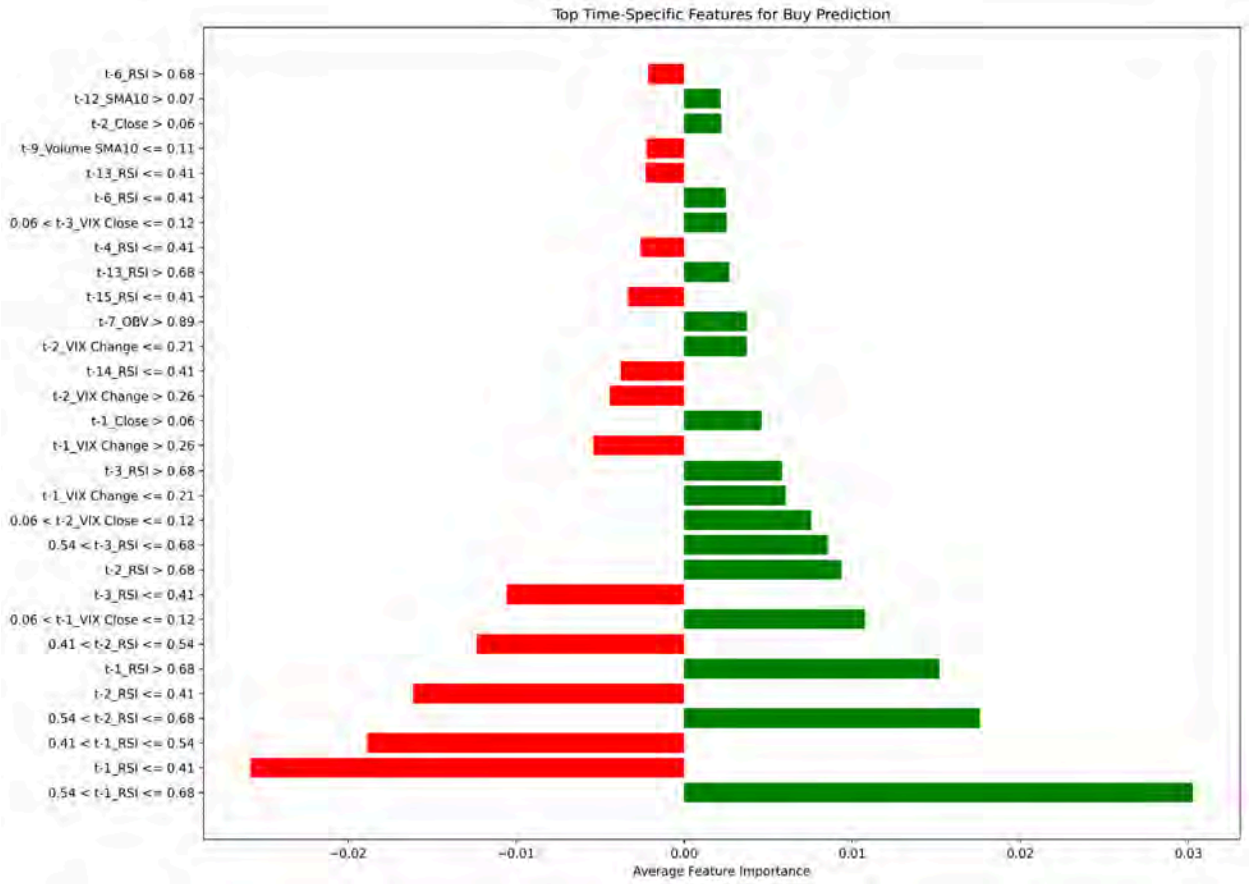


Εικόνα 6.35: LIME Aggregate Importance Plot for Hold action στο Πείραμα με Triple Barrier Labeling.

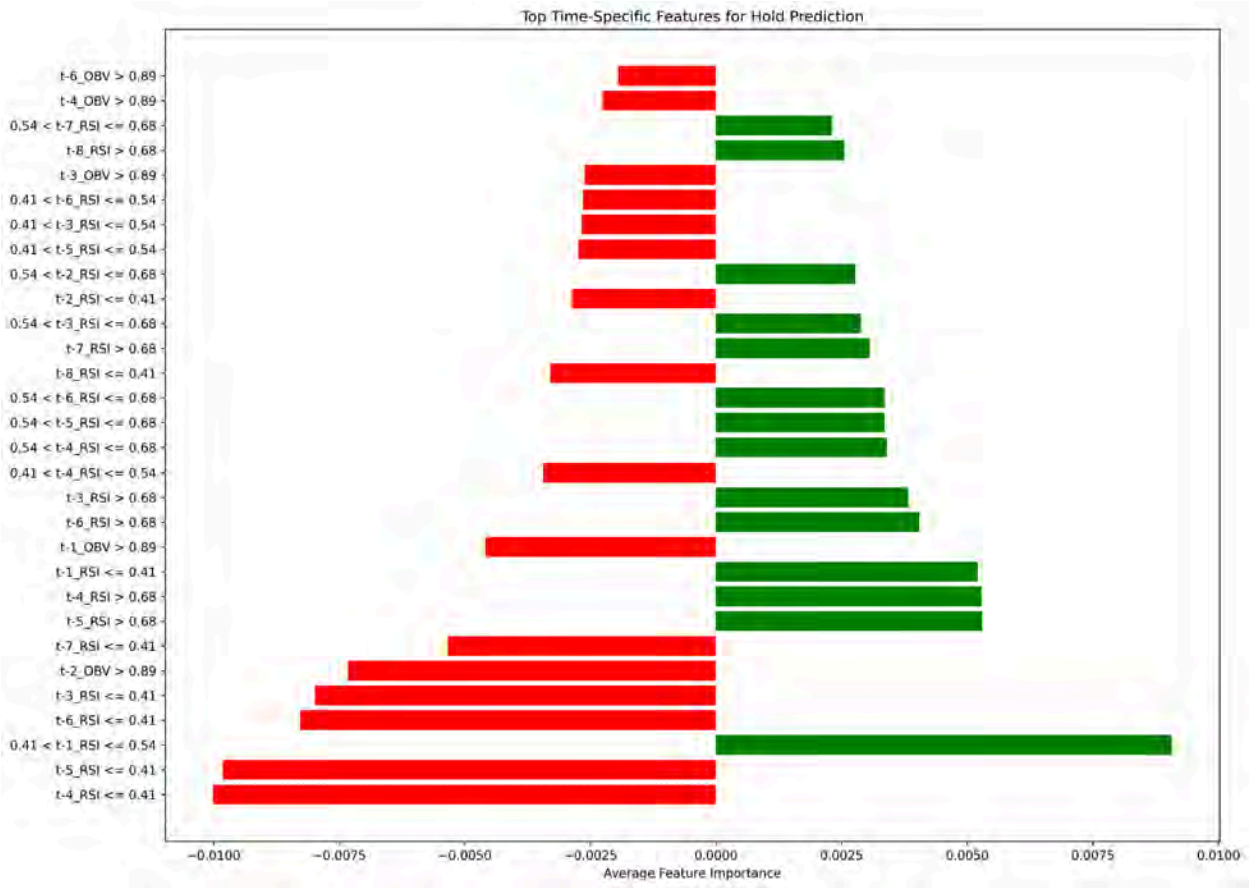


Εικόνα 6.36: LIME Aggregate Importance Plot for Sell action στο Πείραμα με Triple Barrier Labeling.

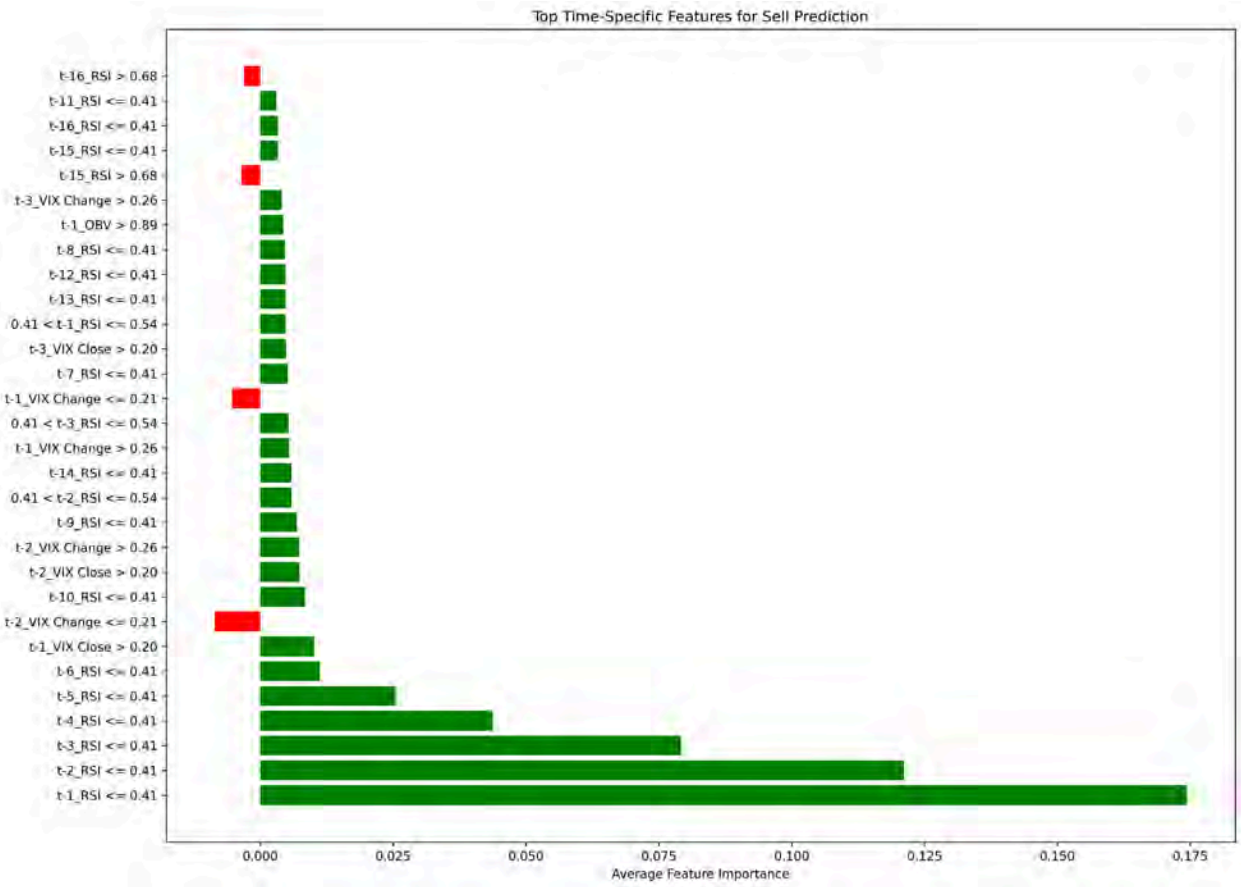
Τα τρία LIME γραφήματα στις εικόνες 6.34, 6.35, 6.36 αποκαλύπτουν πώς το μοντέλο "σκέφτεται" διαφορετικά για κάθε απόφαση trading. Για τις Buy αποφάσεις, το μοντέλο βασίζεται κυρίως σε αρνητικές επιδράσεις (κόκκινες μπάρες), δηλαδή αποφασίζει να αγοράσει όταν συγκεκριμένοι δείκτες (όπως RSI και High price) δείχνουν "κακές" τιμές - αυτό μοιάζει με contrarian στρατηγική που "αγοράζει όταν όλοι πουλάνε". Για τις Sell αποφάσεις, το RSI γίνεται ο απόλυτος κυρίαρχος με τεράστια θετική επίδραση (το μεγαλύτερο πράσινο bar), δηλαδή όταν ο RSI φτάσει σε υψηλά επίπεδα, το μοντέλο αποφασίζει να πουλήσει - αυτό ακολουθεί την κλασική λογική ότι "υψηλός RSI σημαίνει υπεραγορασμένη μετοχή". Για τις Hold αποφάσεις, το On-Balance Volume (OBV) κυριαρχεί απόλυτα, υποδηλώνοντας ότι το μοντέλο επιλέγει να παραμείνει αδρανές όταν ο όγκος συναλλαγών δεν επιβεβαιώνει την κίνηση της τιμής.



Εικόνα 6.37: LIME Time Specific Importance Plot for Buy action στο Πείραμα με Triple Barrier Labeling.



Εικόνα 6.38: LIME Time Specific Importance Plot for Hold action στο Πείραμα με Triple Barrier Labeling.

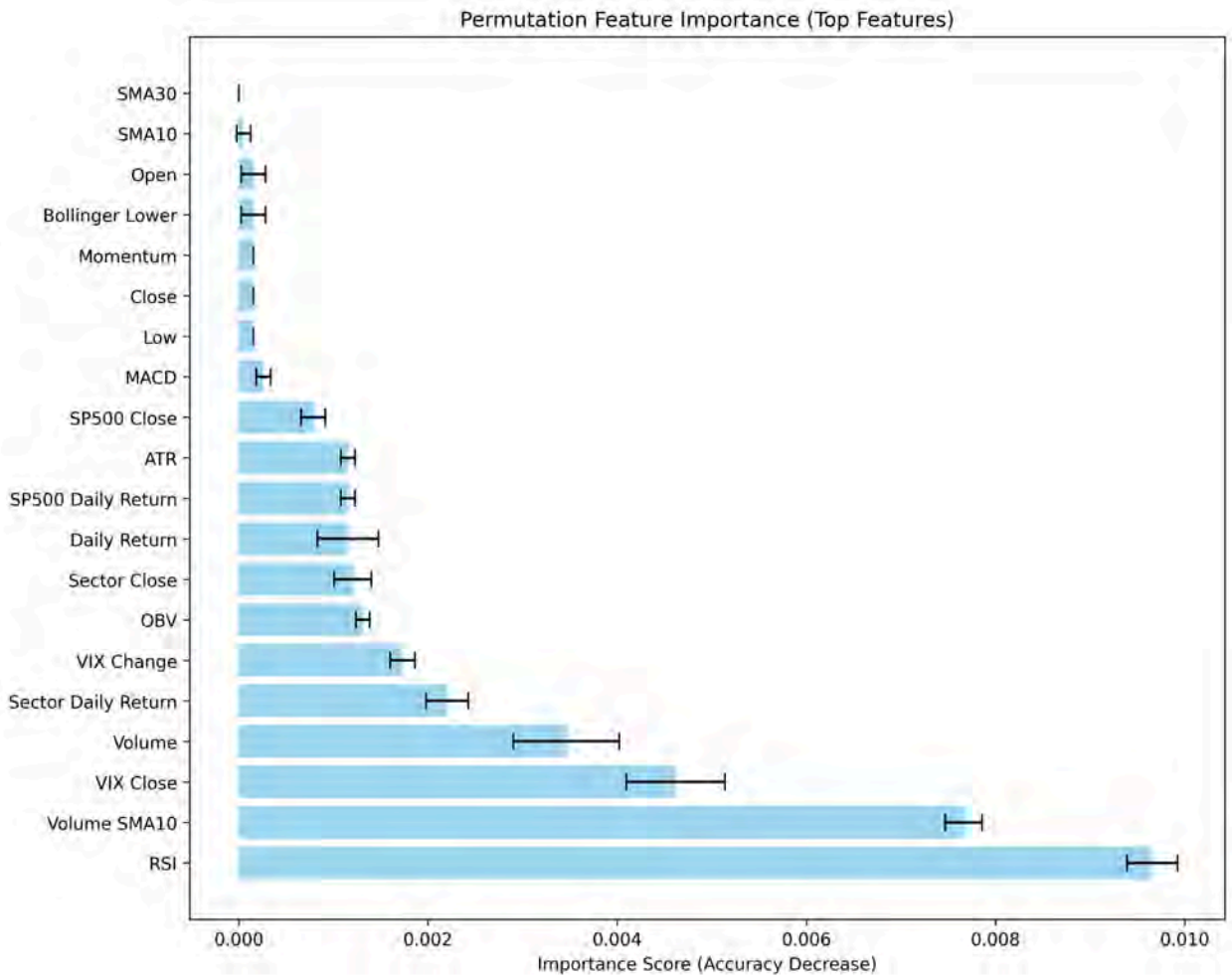


Εικόνα 6.39: LIME Time Specific Importance Plot for Sell action στο Πείραμα με Triple Barrier Labeling.

Τα τρία time-specific LIME γραφήματα στις εικόνες 6.37, 6.38, 6.39 αποκαλύπτουν εντελώς διαφορετικές χρονικές προτεραιότητες για κάθε απόφαση, κάτι που δεν φαινόταν στα aggregative plots. Για τις Buy αποφάσεις, το μοντέλο εστιάζει στα πρόσφατα time steps (t-1 έως t-3), με τον RSI στις στιγμές t-1 και t-2 να έχει τη μεγαλύτερη επίδραση - δηλαδή αποφασίζει να αγοράσει βασιζόμενο κυρίως στις πιο πρόσφατες πληροφορίες των τελευταίων ημερών. Για τις Hold αποφάσεις, παρατηρούμε μια διχοτόμηση: ισχυρές αρνητικές επιδράσεις στο μεσαίο διάστημα (t-4 έως t-6) και κυρίαρχη θετική επίδραση στο t-1, υποδηλώνοντας ότι το μοντέλο επιλέγει να περιμένει όταν οι "παλιές" τιμές RSI ήταν χαμηλές αλλά η τρέχουσα κατάσταση είναι ουδέτερη. Το πιο εντυπωσιακό είναι στις Sell αποφάσεις, όπου το μοντέλο παρουσιάζει συνεχή επίδραση από t-1 έως t-5 με τη μέγιστη ένταση στο t-1, δείχνοντας ότι αποφασίζει να πουλήσει όταν υπάρχει "συνεπής αδυναμία" στις πρόσφατες ημέρες, ενώ τα μακρινότερα time steps (t-11 έως t-16) έχουν μικρότερη

αλλά σταθερή επίδραση, λειτουργώντας ως "ιστορικό context" της απόφασης.

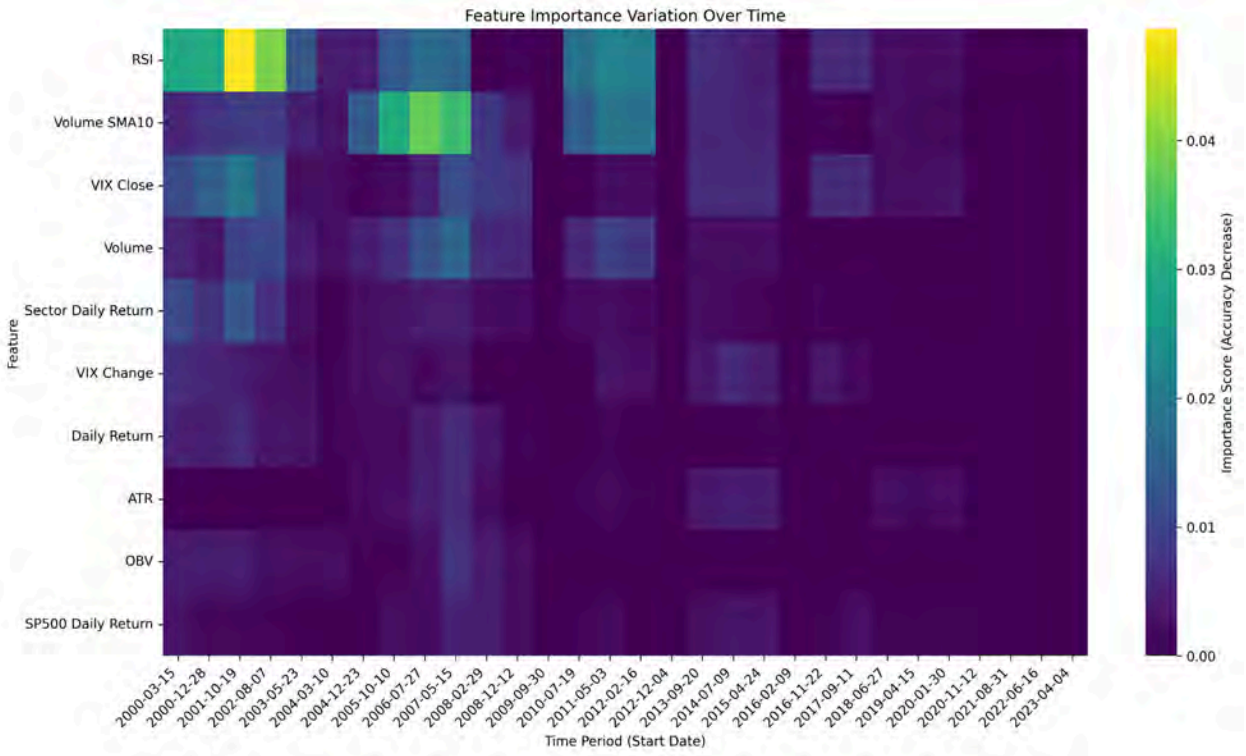
6.2.2.5 Permutation Feature Importance



Εικόνα 6.40: PFI Importance Plot στο Πείραμα με Triple Barrier Labeling.

Στην εικόνα 6.40 το Permutation Importance γράφημα αποκαλύπτει την εξάρτηση του μοντέλου από κάθε χαρακτηριστικό μετρώντας πόσο χειροτερεύει η απόδοσή του όταν κάθε πληροφορία ανακατεύεται. Ο RSI αναδεικνύεται ως το απόλυτα κρίσιμο χαρακτηριστικό - όταν χάνεται, η ακρίβεια του μοντέλου πέφτει κατά ~1% (τεράστια πτώση για ένα trading μοντέλο). Το Volume SMA10 (μέσος όρος του όγκου συναλλαγών) είναι το δεύτερο πιο σημαντικό, υποδεικνύοντας ότι το μοντέλο χρειάζεται να ξέρει αν υπάρχει ενδιαφέρον από τους επενδυτές. Το VIX Close (δείκτης φόβου της αγοράς) και το Volume ακολουθούν, δείχνοντας ότι το μοντέλο νοιάζεται για τη γενική κατάσταση της αγοράς. Αντίθετα, τα τεχνικά χρηματιστηριακά χαρακτηριστικά όπως Moving Averages και Bollinger Bands έχουν σχεδόν μηδενική επίδραση, δηλαδή το μοντέλο θα δουλεύει το ίδιο καλά χωρίς αυτά, δείχνοντας ότι οι κλασικοί τεχνικοί δείκτες δεν είναι φαίνονται τόσο να επηρεάζουν

ιδιαίτερα τις αποφάσεις.



Εικόνα 6.41: PFI Time Varying Importance Plot στο Πείραμα με Triple Barrier Labeling.

Ο RSI εμφάνισε έντονη χρονική διακύμανση στις αρχικές περιόδους (2000-2005) με κίτρινες/πράσινες ζώνες υποδηλώνοντας περιόδους υψηλής σημαντικότητας. Η μετάβαση σε μπλε/μωβ χρώματα μετά το 2005 αντικατόπτρισε σταθεροποίηση της RSI σημαντικότητας σε χαμηλότερα αλλά σταθερά επίπεδα. Το Volume SMA10 εμφάνισε επεισοδιακή ενεργοποίηση σε συγκεκριμένες χρονικές περιόδους (2005-2008, 2010-2012). Το πιο εντυπωσιακό είναι ότι μετά το 2010 σχεδόν όλα τα χαρακτηριστικά "σβήνουν" (σκούρα μωβ χρώματα), υποδηλώνοντας ότι το μοντέλο έγινε λιγότερο εξαρτημένο από συγκεκριμένες πληροφορίες και φαίνεται να έμαθε να λειτουργεί με λιγότερες αλλά πιο αξιόπιστες πληροφορίες. Αυτό εξηγεί γιατί ένα trading μοντέλο που δουλεύει σήμερα μπορεί να μην δουλεύει αύριο.

Συνολικά Συμπεράσματα Δεύτερου Πειράματος:

Συμπερασματικά, το πείραμα με το Triple Barrier Labeling, έχοντας στόχο την εξαγωγή γνώσης από το μοντέλο, οδήγησε σε σημαντικά ευρήματα που αφορούν το μοντέλο μας και τις μεθόδους εξηγησιμότητας. Η μετάβαση στην ρεαλιστική στρατηγική σήμανσης του Triple Barrier αποκάλυψε μια πιο σύνθετη ιεραρχία σημαντικότητας χαρακτηριστικών σε σχέση με την προκαθορισμένη στρατηγική του πρώτου πειράματος.

- Ενώ ο RSI παρέμεινε κεντρικός παράγοντας, η απόλυτη κυριαρχία του μετριάστηκε. Αναδείχθηκαν ως εξίσου κρίσιμοι οι δείκτες όγκου (ιδίως το Volume SMA10 και το OBV) και οι δείκτες μεταβλητότητας (VIX). Αυτό υποδεικνύει ότι το επιτυχημένο μοντέλο έμαθε να συνδυάζει την ορμή (RSI) με την επιβεβαίωση του όγκου και την αξιολόγηση του γενικότερου κλίματος της αγοράς (VIX)
- Η ανάλυση με LIME ήταν αποκαλυπτική, δείχνοντας ότι το μοντέλο ανέπτυξε ξεχωριστές στρατηγικές για κάθε δράση. Για τις αποφάσεις "Sell", ο RSI ήταν ο απόλυτος κυρίαρχος, ακολουθώντας την κλασική λογική της πώλησης σε υπεραγορασμένες συνθήκες. Αντίθετα, για τις αποφάσεις "Buy", το μοντέλο υιοθέτησε μια contrarian λογική, βασιζόμενο σε αρνητικές ενδείξεις δεικτών, ουσιαστικά "αγοράζοντας τον φόβο". Για τις αποφάσεις "Hold", η έλλειψη επιβεβαίωσης από τον όγκο (OBV) ήταν ο καθοριστικός παράγοντας.
- Οι χρονικές αναλύσεις (time-varying plots) από PFI και SHAP έδειξαν ότι το μοντέλο δεν εφαρμόζει μια στατική στρατηγική, αλλά προσαρμόζεται δυναμικά στις συνθήκες της αγοράς (market regimes). Η σημαντικότητα των RSI και VIX εντεινόταν σε περιόδους κρίσεων , ενώ σε πιο ήρεμες περιόδους το μοντέλο φάνηκε να βασίζεται σε λιγότερους αλλά πιο αξιόπιστους δείκτες.

Όσον αφορά τις μεθόδους εξηγησιμότητας, κάθε μέθοδος προσέφερε μια μοναδική και συμπληρωματική οπτική, επιβεβαιώνοντας ότι ο συνδυασμός τους είναι απαραίτητος για μια ολοκληρωμένη εικόνα.

- **Permutation Feature Importance (PFI):** Αποδείχθηκε αρκετά αποτελεσματική μέθοδος για την παροχή μιας σαφούς, καθολικής ιεραρχίας της σημασίας των χαρακτηριστικών. Η ικανότητά της να μετρά απευθείας τον αντίκτυπο στην ακρίβεια την καθιστά ένα αντικειμενικό και αξιόπιστο εργαλείο για τον εντοπισμό των πιο κρίσιμων παραγόντων. Παρ' όλα αυτά επειδή δεν παρέχει local αλλά καθολικές εξηγήσεις για το μοντέλο, δεν δύναται να μας δώσει εξηγήσεις για συγκεκριμένες χρονικές στιγμές όπως μπορούν οι υπόλοιπες μέθοδοι.
- **LIME:** Υπήρξε η πιο ισχυρή μέθοδος για την κατανόηση της *λογικής* πίσω από κάθε τύπο απόφασης (Buy/Hold/Sell). Η ικανότητά της να αποκαλύπτει τις class-specific στρατηγικές ήταν ένα από τα πολυτιμότερα ευρήματα της ανάλυσης. Με την λογική που αποτυπώνεται από τις εξηγήσεις θα μπορούσαμε να εξάγουμε κανόνες για trading και να εμπλουτίσουμε μια standand στρατηγική trading, κάτι που δεν θα μπορούσαμε με τις υπόλοιπες μεθόδους.
- **Gradient-Based Μέθοδοι (DeepLIFT & Integrated Gradients):** Ήταν χρήσιμες για την επιβεβαίωση της κατεύθυνσης της επίδρασης των χαρακτηριστικών (θετική/αρνητική) και για την ανίχνευση της προσαρμογής του μοντέλου σε περιόδους κρίσεων, όπως η αυξημένη σημασία του VIX. Τα αποτελέσματά τους ήταν σε μεγάλο βαθμό συνεπή μεταξύ τους. Επίσης με βάση το πρώτο πείραμα φαίνεται να δείνουν την πιο κοντινή εικόνα στην πραγματική λογική που χρησιμοποιεί το μοντέλο για την λήψη αποφάσεων.
- **Kernel SHAP:** Παρείχε την πιο λεπτομερή ανάλυση σε επίπεδο χρονικών βημάτων και παραθύρων, αποκαλύπτοντας την πολυπλοκότητα του τρόπου με τον οποίο το LSTM συνδυάζει πληροφορίες από διαφορετικούς χρονικούς ορίζοντες. Ωστόσο, η συνολική κατάταξη της σημαντικότητας που παρήγαγε διέφερε από τις άλλες μεθόδους, αναδεικνύοντας πιθανώς πιο σύνθετες, μη-γραμμικές αλληλεπιδράσεις που έμαθε το μοντέλο, όπως και την μοναδικότητα της μεθόδου.

Συμπερασματικά, το δεύτερο πείραμα κατέδειξε ότι ένα μοντέλο LSTM εκπαιδευμένο με την Triple Barrier μέθοδο μπορεί να αναπτύξει μια κερδοφόρα και προσαρμοστική στρατηγική συναλλαγών. Η επιτυχία του δεν βασίζεται σε απλούς, στατικούς κανόνες, αλλά σε μια δυναμική αξιολόγηση της ορμής, του όγκου και της

μεταβλητότητας, προσαρμοσμένη στις συνθήκες της αγοράς και στον τύπο της απόφασης που πρόκειται να ληφθεί. Κάθε μια μέθοδος αλλά κυρίως ο συνδυασμός όλων των μεθόδων XAI, μια ολοκληρωμένη και κατανοητή εικόνα αυτής της πολύπλοκης συμπεριφοράς.

ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ

7.1 Συγκριτική Αξιολόγηση Πειραμάτων και Εξηγησιμότητας

Η παρούσα διπλωματική εργασία διερεύνησε την αποτελεσματικότητα και την ερμηνευσιμότητα ενός μοντέλου LSTM για την πρόβλεψη σημάτων συναλλαγών (Buy/Hold/Sell) στην μετοχή της Apple (AAPL), χρησιμοποιώντας δύο διαφορετικές στρατηγικές σήμανσης: μια προσαρμοσμένη (Custom Labeling) και τη μέθοδο Τριπλού Φράγματος (Triple Barrier). Η αξιολόγηση της εξηγησιμότητας πραγματοποιήθηκε με ένα σύνολο πέντε σύγχρονων μεθόδων XAI (DeepLIFT, Integrated Gradients, Kernel SHAP, LIME, Permutation Feature Importance), οι οποίες παρείχαν πολλαπλές οπτικές στη συμπεριφορά του μοντέλου.

Κεντρικές Διαπιστώσεις:

1. **Επίδραση της Στρατηγικής Σήμανσης στα Εκμαθέντα Πρότυπα**: Η επιλογή της στρατηγικής σήμανσης αποδείχθηκε καθοριστική για τα πρότυπα που μαθαίνει το μοντέλο και, κατ' επέκταση, για την ιεραρχία σημαντικότητας των χαρακτηριστικών που αναδεικνύεται από τις μεθόδους XAI.
    - **Custom Labeling**: Το μοντέλο, εκπαιδευμένο με την προσαρμοσμένη στρατηγική κανόνων, έδειξε έντονη εξάρτηση από τον RSI. Αυτό ήταν ιδιαίτερα εμφανές στην Permutation Feature Importance, όπου ο RSI κυριάρχησε με διαφορά. Οι gradient-based μέθοδοι (DeepLIFT, Integrated Gradients) επιβεβαίωσαν επίσης αυτή την κυριαρχία. Η ερμηνεία των αποφάσεων ήταν σχετικά ευθύγραμμη, καθώς το μοντέλο φάνηκε να υιοθετεί τους explicit κανόνες.
    - **Triple Barrier Labeling**: Με τη μέθοδο του Τριπλού Φράγματος, η σημαντικότητα του RSI μειώθηκε σε σχέση με την απόλυτη κυριαρχία που εμφάνιζε, αν και παρέμεινε πρωταρχικός παράγοντας. Αναδείχθηκαν ισχυρότερα χαρακτηριστικά όγκου (ιδίως OBV) και μεταβλητότητας (VIX), υποδηλώνοντας ότι το μοντέλο έμαθε να λαμβάνει υπόψη πιο σύνθετες και ρεαλιστικές συνθήκες αγοράς. Η κατανομή της σημαντικότητας έγινε πιο ισομερής μεταξύ των χαρακτηριστικών. Αυτό υποδηλώνει ότι η Triple Barrier strategy οδηγεί σε πιο ισορροπημένη χρήση των πληροφοριών.
2. **Ρόλος των Μεθόδων XAI στην Εξαγωγή Γνώσης**: Η συνδυαστική χρήση

πολλαπλών μεθόδων XAI ήταν απαραίτητη για την ολοκληρωμένη κατανόηση του μοντέλου. Κάθε μέθοδος προσέφερε μοναδικές πληροφορίες:

- **Permutation Feature Importance**: Αποτέλεσε την πιο αξιόπιστη μέθοδο για τη συνολική ιεραρχία σημαντικότητας των χαρακτηριστικών, παρέχοντας μια ξεκάθαρη εικόνα του αντίκτυπου κάθε χαρακτηριστικού στην απόδοση του μοντέλου.
- **LIME**: Ήταν ανεκτίμητη για την class-specific ανάλυση, αποκαλύπτοντας ότι το μοντέλο χρησιμοποιεί διαφορετικά σύνολα χαρακτηριστικών για την πρόβλεψη Buy, Hold ή Sell. Για παράδειγμα, η ανάδειξη contrarian logic στις Buy αποφάσεις με Triple Barrier (π.χ. χαμηλές τιμές και αρνητικές συνεισφορές) ήταν ένα κρίσιμο εύρημα.
- **DeepLIFT & Integrated Gradients**: Παρέμειναν σημαντικές για την κατανόηση της κατεύθυνσης της επίδρασης των χαρακτηριστικών (θετική/αρνητική) και τη χρονική τους συνέπεια ή διακύμανση, ειδικά σε επίπεδο νευρωνικού δικτύου.
- **Kernel SHAP**: Προσέφερε μια εναλλακτική οπτική στην ιεραρχία σημαντικότητας (π.χ. κυριαρχία MACD στο Custom Labeling) και ήταν ιδιαίτερα χρήσιμη για την ανάλυση πολλαπλών χρονικών κλιμάκων και τις αμφίδρομες επιδράσεις των χαρακτηριστικών.

3. **Στρατηγικές Επιπτώσεις για το Algorithmic Trading**: Τα ευρήματα έχουν σημαντικές επιπτώσεις για την ανάπτυξη και βελτιστοποίηση αλγοριθμικών στρατηγικών συναλλαγών:
   - **Βελτιστοποίηση Σήμανσης (Labeling Optimization)**: Η μέθοδος Triple Barrier φαίνεται να παράγει πιο ρεαλιστικά και προσαρμόσιμα μοντέλα, καθώς ενσωματώνει μηχανισμούς διαχείρισης κινδύνου (stop-loss) και λήψης κέρδους (profit-taking) στην ίδια τη διαδικασία σήμανσης. Αυτό οδηγεί σε μοντέλα που αντιδρούν πιο αποτελεσματικά στις πραγματικές συνθήκες της αγοράς.
   - **Εστίαση στο Feature Engineering**: Η ανάδειξη της σημασίας των χαρακτηριστικών όγκου (OBV) και μεταβλητότητας (VIX) στο πείραμα Triple Barrier υπογραμμίζει την ανάγκη για πιο εξελιγμένο feature engineering σε αυτές τις κατηγορίες, πέραν των παραδοσιακών δεικτών τιμής και ορμής.
   - **Δυναμική Προσαρμογή Στρατηγικών**: Η χρονική μεταβλητότητα της σημαντικότητας των χαρακτηριστικών (ιδίως με Triple Barrier), υποδηλώνει ότι τα επιτυχημένα trading systems πρέπει να ενσωματώνουν μηχανισμούς ανίχνευσης market regimes και δυναμικής προσαρμογής των παραμέτρων της στρατηγικής.

7.2 Ερευνητική Συνεισφορά και Μελλοντικές Κατευθύνσεις

Η παρούσα διπλωματική εργασία συνεισφέρει στην ερευνητική κοινότητα σε πολλαπλά επίπεδα:

- **Συγκριτική Ανάλυση Στρατηγικών Σήμανσης**: Παρέχει εμπειρικά στοιχεία για το πώς διαφορετικές στρατηγικές σήμανσης επηρεάζουν τα εκμαθέντα πρότυπα ενός μοντέλου LSTM σε χρηματοοικονομικά δεδομένα. Αναδεικνύεται η ανωτερότητα της μεθόδου Triple Barrier στην παραγωγή πιο ισορροπημένων και ρεαλιστικών μοντέλων.
- **Πολυ-Μεθοδολογική Εξηγησιμότητα**: Εφαρμόζει και συγκρίνει ένα ευρύ φάσμα μεθόδων XAI σε ένα πολύπλοκο πολυ-εισοδικό LSTM μοντέλο, αναδεικνύοντας τις συμπληρωματικές πληροφορίες που μπορεί να προσφέρει κάθε μέθοδος. Αυτό ενισχύει την κατανόηση του "γιατί" πίσω από τις προβλέψεις "μαύρων κουτιών".
- **Εξαγωγή Γνώσης για Χρηματοοικονομικές Αγορές**: Τα ευρήματα παρέχουν συγκεκριμένες διαπιστώσεις για τους παράγοντες που επηρεάζουν τις αποφάσεις trading, όπως η κυριαρχία του RSI και η αυξημένη σημασία του όγκου και της μεταβλητότητας σε ρεαλιστικά σενάρια.

Μελλοντικές κατευθύνσεις έρευνας περιλαμβάνουν:

1. **Βελτιστοποίηση Παραμέτρων Triple Barrier**: Διερεύνηση της βέλτιστης επιλογής των παραμέτρων του Triple Barrier (N ημερών, ποσοστά profit/stop-loss, ουδέτερη ζώνη) μέσω αλγορίθμων βελτιστοποίησης (π.χ. γενετικοί αλγόριθμοι) για περαιτέρω βελτίωση της απόδοσης του μοντέλου.
2. **Ενσωμάτωση Μηχανισμών Προσοχής (Attention Mechanisms)**: Η ενσωμάτωση μηχανισμών προσοχής στα επίπεδα LSTM θα μπορούσε να επιτρέψει στο μοντέλο να αποδίδει δυναμικά μεγαλύτερη βαρύτητα σε συγκεκριμένα χρονικά βήματα ή χαρακτηριστικά, βελτιώνοντας την προσαρμοστικότητα του μοντέλου σε μεταβαλλόμενες συνθήκες αγοράς.
3. **Επέκταση σε Άλλα Χρηματοοικονομικά Μέσα**: Εφαρμογή της μεθοδολογίας σε άλλα χρηματοοικονομικά μέσα (π.χ., συναλλαγματικές ισοτιμίες, εμπορεύματα, κρυπτονομίσματα) για την επαλήθευση της γενικευσιμότητας των ευρημάτων.
4. **Σύνθετο Feature Engineering**: Ανάπτυξη πιο σύνθετων χαρακτηριστικών όγκου και μεταβλητότητας, καθώς και ενσωμάτωση δεδομένων από ανάλυση συναισθήματος (π.χ., από ειδήσεις ή μέσα κοινωνικής δικτύωσης), για την

παροχή ακόμα πλουσιότερης πληροφορίας στο μοντέλο.

5. **Ενισχυτική Μάθηση (Reinforcement Learning)**: Διερεύνηση της εφαρμογής ενισχυτικής μάθησης για την άμεση εκπαίδευση ενός πράκτορα trading, όπου ο πράκτορας θα μαθαίνει τις βέλτιστες ενέργειες μέσα από αλληλεπίδραση με το περιβάλλον της αγοράς και λήψη ανταμοιβών για κερδοφόρες συναλλαγές.

Συνοψίζοντας, η παρούσα εργασία κατέδειξε την κρισιμότητα της επιλογής στρατηγικής σήμανσης και την αξία των μεθόδων XAI στην κατανόηση των πολύπλοκων μηχανισμών λήψης αποφάσεων των μοντέλων βαθιάς μάθησης σε έναν τόσο δυναμικό και κρίσιμο τομέα όπως οι χρηματοοικονομικές αγορές. Τα ευρήματα θέτουν τις βάσεις για την ανάπτυξη πιο διαφανών, αξιόπιστων και αποτελεσματικών συστημάτων αλγοριθμικού trading στο μέλλον.

ΚΕΦΑΛΑΙΟ 8: ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] A. Hayes, "How Does the Stock Market Work?," *Investopedia*, May 02, 2024. [Online]. Available: https://www.investopedia.com/articles/investing/082614/how-stock-market-works.asp .

[2] Strike.money, "History of Stock Market: Everything to Know How it All Started." [Online]. Available: https://www.strike.money/stock-market/history.

[3] Forbes, "Apple | AAPL Stock Price, Company Overview & News." [Online]. Available: https://www.forbes.com/companies/apple/.

[4] Cleverence, "Apple Stock | AAPL Price, Latest News & Insights." [Online]. Available: https://www.cleverence.com/articles/business-blogs/apple-stock-aapl-price-latest-news-insights/.

[5] Apple Inc., "FAQ." [Online]. Available: https://investor.apple.com/faq/default.aspx.

[6] MacroTrends, "Apple - 45 Year Stock Price History | AAPL." [Online]. Available: https://www.macrotrends.net/stocks/charts/AAPL/apple/stock-price-history.

[7] IG, "Apple stock split 2020: what you need to know," Aug 03, 2023. [Online]. Available: https://www.ig.com/en/trading-strategies/apple-stock-split-2020--what-you-need-to-know-230803.

[8] Capital.com, "Apple (AAPL) stock price forecast: third-party price target." [Online]. Available: https://capital.com/en-int/analysis/apple-stock-price-in-10-years.

[9] J. Chen, "Technical Analysis: What It Is and How to Use It in Investing," *Investopedia*, Apr 29, 2024. [Online]. Available: https://www.investopedia.com/terms/t/technicalanalysis.asp.

[10] J. Chen, "4 Ways to Predict Market Performance," *Investopedia*, Oct 31, 2022. [Online]. Available: https://www.investopedia.com/articles/07/mean_reversion_martingale.asp.

[11] S. Seabury, "Fundamental Analysis of Stocks: Key Concepts and Techniques," *Santa Clara University Leavey School of Business*, Dec 15, 2022. [Online]. Available: https://onlinedegrees.scu.edu/media/blog/fundamental-analysis-stocks.

[12] J. B. Maverick, "Financial Analysis: Definition, Importance, Types, and Examples," *Investopedia*, May 13, 2024. [Online]. Available:

https://www.investopedia.com/terms/f/financial-analysis.asp.

[13] Winvesta, "Fundamental Analysis: Definition, Types, Benefits, and How To Do It?," *Winvesta Blog*. [Online]. Available: https://www.winvesta.in/blog/fundamental-analysis-a-complete-guide.

[14] European Data Protection Supervisor, "Explainable AI," *TechDispatch*, no. 3/2023, Nov 16, 2023. [Online]. Available: https://www.edps.europa.eu/system/files/2023-11/23-11-16_techdispatch_xai_en.pdf.

[15] IBM, "What is Explainable AI (XAI)?" [Online]. Available: https://www.ibm.com/think/topics/explainable-ai.

[16] Unite.AI, "Τι είναι το Explainable AI;" [Online]. Available: https://www.unite.ai/el/%CF%84%CE%B9-%CE%B5%CE%BE%CE%B7%CE%B3%CE%B5%CE%AF%CF%84%CE%B1%CE%B9-%CE%B1%CE%B9/.

[17] Εθνικό Κέντρο Τεκμηρίωσης και Ηλεκτρονικού Περιεχομένου (ΕΚΤ), "Τεχνητή Νοημοσύνη & Χρηματοοικονομικά." [Online]. Available: https://www.ekt.gr/sites/ekt-site/files/Material%20%CE%B3%CE%B9%CE%B1%20TN%20%26%20%CE%A7%CE%A1%CE%97%CE%9C%CE%91%CE%A4%CE%9F%CE%9F%CE%99%CE%9A%CE%9F%CE%9D%CE%9F%CE%9C%CE%99%CE%9A%CE%91.pdf.

[18] C. Garbin, G. Zoppello, M. Atzori, and G. Iacca, "T-Explainer: A Model-Agnostic Explainability Framework Based on Gradients," *arXiv preprint arXiv:2404.16495*, 2024.

[19] A. A. Rasheed, S. M. F. D. Syed, and F. A. Al-Ghamdi, "An explainable deep learning approach for stock market trend prediction using advanced technical indicators," *Sensors*, vol. 24, no. 11, p. 3474, 2024.

[20] Improvado, "Financial Analysis: How to Use Data for Business." [Online]. Available: https://improvado.io/blog/financial-data-analysis.

[21] Big Blue Data Academy, "Responsible AI: Ορισμός, Σημασία & Πρακτικές." [Online]. Available: https://bigblue.academy/gr/responsible-ai.

[22] Z. Zhang, Z. Li, and Y. Chen, "Forecasting S&P 500 Using LSTM Models," *arXiv preprint arXiv:2501.17366*, 2025.

[23] Wikipedia, "Vanishing gradient problem." [Online]. Available: https://en.wikipedia.org/wiki/Vanishing_gradient_problem.

[24] Sapien.io, "Vanishing/Exploding Gradients Explained | Sapien's AI Glossary." [Online]. Available: https://www.sapien.io/glossary/definition/vanishing-exploding-gradients.

[25] IBM, "LSTM (Long Short-Term Memory)." [Online]. Available: https://www.ibm.com/cloud/learn/lstm.

[26] A. Guijo-Rubio, P. A. Gutiérrez, and C. Hervás-Martínez, "Automatic selection of the best neural architecture for time series forecasting via multi-objective optimization and Pareto optimality conditions," *arXiv preprint arXiv:2501.12215*, 2025.

[27] H. Chen, C. Zhang, and Y. Wang, "Stock Price Prediction Using a Hybrid LSTM-GNN Model: Integrating Time-Series and Graph-Based Analysis," *arXiv preprint arXiv:2502.15813*, 2025.

[28] P. Yadav, "LSTM Models for Time Series Forecasting," *Kaggle*. [Online]. Available: https://www.kaggle.com/code/pawan2905/lstm-models-for-time-series-forecasting.

[29] M. Lopez de Prado, *Advances in financial machine learning*. John Wiley & Sons, 2018.

[30] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[31] A. Kumar, S. Singh, and R. Gupta, "Multi-Agent Stock Prediction Systems: Machine Learning Models, Simulations, and Real-Time Trading Strategies," *arXiv preprint arXiv:2502.15853*, 2025.

[32] S. S. Kumar, M. H. Kumar, S. G. Kumar, and M. S. Kumar, "A Deep Learning-Based LSTM for Stock Price Prediction Using Twitter Sentiment Analysis," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 12, 2024.

[33] S. Kim, J. Kim, and S. Kim, "Info-CELS: Informative Saliency Map-Guided Counterfactual Explanation for Time Series Classification," *Electronics*, vol. 14, no. 7, p. 1311, 2025.

[34] T. Schafer, K. Ben-Hadj-Alouane, and S. Kramer, "Counterfactual Explanations for Time Series Forecasting," in *2023 IEEE International Conference on Big Data (BigData)*, 2023, pp. 2484–2493.

[35] Z. Wang, Z. Wang, Y. Liu, and Y. Wang, "Reliable Time Series Counterfactual Explanations Guided by ShapeDBA," *arXiv preprint arXiv:2404.14440*, 2024.

[36] M. A. Ali et al., "Recent Applications of Explainable AI (XAI): A Systematic Literature Review," *Electronics*, vol. 14, no. 19, p. 8884, 2025.

[37] S. Sharma, S. Sharma, and A. Athiyarath, "A REVIEW OF EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI): TOWARDS FINANCE," *Nanotechnology Perceptions*, vol. 20, no. 2, pp. 77–88, 2024.

[38] J. Li, Y. Gui, and H. He, "Example or Prototype? Learning Concept-Based

Explanations in Time-Series," in *Proceedings of the 39th International Conference on Machine Learning*, 2022, pp. 12933–12946.

[39] C. K. O'Brien, N. D. D'Souza, F. P. M. T. de Melo, and M. T. Keane, "Semi-factual Explanations in AI," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 10, 2022, pp. 10839–10847.

[40] M. T. Keane, C. K. O'Brien, B. Smyth, D. D. Nau, and F. P. M. T. de Melo, "Even If Explanations: Prior Work, Desiderata & Benchmarks for Semi-Factual XAI," in *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 2024, pp. 732.

[41] A. B. M. A. A. Islam et al., "Evaluation Metrics Research for Explainable Artificial Intelligence Global Methods Using Synthetic Data," *Geomatics*, vol. 6, no. 1, pp. 26–47, 2025.

[42] S. Sharma, S. Sharma, and A. Athiyarath, "Explainable AI in Algorithmic Trading: Mitigating Bias and Improving Regulatory Compliance in Finance," *SSRN*, 2024.

[43] Meegle, "Explainable AI For Financial Forecasting." [Online]. Available: https://www.meegle.com/resources/explainable-ai-for-financial-forecasting.

[44] smartdev.com, "Addressing AI Bias and Fairness: Challenges, Implications, and Strategies for Ethical AI." [Online]. Available: https://smartdev.com/addressing-ai-bias-and-fairness-challenges-implications-and-strategies-for-ethical-ai/.

[45] S. Sharma, S. Sharma, and A. Athiyarath, "Bias Detection and Fairness in Large Language Models for Financial Services," *SSRN*, 2024.

[46] Lumenova AI, "Why Explainable AI in Banking and Finance Is Critical for Compliance." [Online]. Available: https://www.lumenova.ai/blog/ai-banking-finance-compliance/.

[47] C. Molnar, "15 Counterfactual Explanations," in *Interpretable Machine Learning*. [Online]. Available: https://christophm.github.io/interpretable-ml-book/counterfactual.html.

[48] Optiblack, "AI Bias Audit: 7 Steps to Detect Algorithmic Bias." [Online]. Available: https://optiblack.com/blog/ai-bias-audit.

[49] Crescendo.ai, "10 Real AI Bias Examples & Mitigation Guide." [Online]. Available: https://www.crescendo.ai/blog/ai-bias-examples.

[50] R. Guidotti, "Beware of 'Explanations' of AI," *arXiv preprint arXiv:2504.06791*, 2025.

[51] N. R. John Solaro, B. M. Wolska, and J. M. O'Donnell, "Current methods in explainable artificial intelligence and future prospects for integrative physiology," *Physiological Genomics*, vol. 56, no. 5, pp. 275–283, 2024.

[52] D. Sen, B. S. Deora, and A. Vaishnav, "Explainable Deep Learning for Time Series Analysis: Integrating SHAP and LIME in LSTM-Based Models," *Journal of Information Systems Engineering and Management*, 2024.

[53] ApX Machine Learning, "LIME vs SHAP: What's the Difference for Model Interpretability?," Apr 17, 2025. [Online]. Available: https://apxml.com/posts/lime-vs-shap-difference-interpretability.

[54] MarkovML, "LIME vs SHAP: A Comparative Analysis of Interpretability Tools," Feb 26, 2024. [Online]. Available: https://www.markovml.com/blog/lime-vs-shap.

[55] S. Ben Ayed, M. A. Talbi, and M. A. Al-Shourbaji, "Developing an explainable deep learning module based on the LSTM framework for flood prediction," *Frontiers in Water*, vol. 7, 2025.

[56] A. A. Rasheed, S. M. F. D. Syed, and F. A. Al-Ghamdi, "A Literature Review on Applications of Explainable Artificial Intelligence (XAI)," *SSRN*, 2025.

[57] A. T. Tran, P. T. Nguyen, and T. T. T. Nguyen, "Investigating Explainability Methods in Recurrent Neural Network Architectures for Financial Time Series Data," *Applied Sciences*, vol. 12, no. 3, p. 1427, 2022.

[58] H. Turbé, M. Bjelogrlic, C. Lovis, and G. Mengaldo, "Evaluation of post-hoc interpretability methods in time-series classification," *Nature Machine Intelligence*, vol. 5, no. 3, pp. 250–260, 2023.

[59] C. Trakas, "Explainability methods as a means of supporting medical findings," B.Sc. thesis, Dept. Informatics, Athens University of Economics and Business, Athens, Greece, 2023.