

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ, ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

«BEHAVIORAL FINANCE MEETS DATA ANALYTICS: PREDICTING MARKET TRENDS WITH SENTIMENT ANALYSIS AND INFORMATION THEORY»

ΟΝΟΜΑΤΕΠΩΝΥΜΟ: ΚΩΝΣΤΑΝΤΙΝΟΣ ΣΑΜΑΡΙΔΗΣ

ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΕΠΙΒΛΕΠΟΝΤΟΣ ΚΑΘΗΓΗΤΗ: ΜΑΣΤΟΡΑΚΗΣ ΓΕΩΡΓΙΟΣ

ΙΟΥΝΙΟΣ 2025



ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ, ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ

«BEHAVIORAL FINANCE MEETS DATA ANALYTICS: PREDICTING MARKET TRENDS WITH SENTIMENT ANALYSIS AND INFORMATION THEORY»

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

ΟΝΟΜΑΤΕΠΩΝΥΜΟ: ΚΩΝΣΤΑΝΤΙΝΟΣ ΣΑΜΑΡΙΔΗΣ

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

ΜΑΣΤΟΡΑΚΗΣ ΓΕΩΡΓΙΟΣ

ΛΕΜΟΝΑΚΗΣ ΧΡΗΣΤΟΣ

ΚΩΝΣΤΑΝΤΙΝΟΣ ΠΑΝΑΓΙΩΤΑΚΗΣ

ΙΟΥΝΙΟΣ 2025

Πνευματικά δικαιώματα

Copyright © Κωνσταντίνος Σαμαρίδης, 2025
Με επιφύλαξη παντός δικαιώματος. All Rights reserved.
Η έγκριση της Διπλωματικής Εργασίας από το Πανεπιστημίου Νεάπολις δεν υποδηλώνει απαραιτήτως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Πανεπιστημίου.

Σελίδα εγκυρότητας

Ονοματεπώνυμο Φοιτητή: Κωνσταντίνος Σαμαρίδης

Τίτλος Διπλωματικής Εργασίας: Behavioral Finance Meets Data Analytics: Predicting Market Trends with Sentiment Analysis and Information Theory

Η παρούσα Διπλωματική Εργασία εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση εξ αποστάσεως μεταπτυχιακού τίτλου στο Πανεπιστήμιο Νεάπολις και εγκρίθηκε στις 20/6/2025 από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή:

Επιβλέπων (Πανεπιστήμιο Νεάπολις Πάφος): Γεώργιος Μαστοράκης, Αναπλ. Καθηγητής, Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας, Ελληνικό Μεσογειακό Πανεπιστήμιο

Μέλος Εξεταστικής Επιτροπής: Χρήστος Λεμονάκης, Αναπλ. Καθηγητής, Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας, Ελληνικό Μεσογειακό Πανεπιστήμιο

Μέλος Εξεταστικής Επιτροπής: Κωνσταντίνος Παναγιωτάκης, Καθηγητής, Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας, Ελληνικό Μεσογειακό Πανεπιστήμιο

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Ο Σαμαρίδης Κωνσταντίνος, γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα ότι η παρούσα εργασία με τίτλο «Behavioral Finance Meets Data Analytics: Predicting Market Trends with Sentiment Analysis and Information Theory», αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές που έχω χρησιμοποιήσει, έχουν δηλωθεί κατάλληλα στις βιβλιογραφικές παραπομπές και αναφορές. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

Ο Δηλών

Κωνσταντίνος Σαμαρίδης

Πρόλογος

Από πάντα με ενδιέφερε το πώς οι άνθρωποι παίρνουν αποφάσεις. Πώς επηρεάζονται, πώς δρουν υπο πίεση και πώς το συναίσθημα μπορεί να επηρεάσει τη λογική τους. Αυτός ο διαρκής προβληματισμός με οδήγησε φυσικά στο πεδίο της Συμπεριφορικής Χρηματοοικονομικής, ενός κλάδου που συνδυάζει την ψυχολογία με την οικονομική θεωρία. Έτσι, η επιλογή του θέματος της παρούσας διπλωματικής εργασίας δεν ήταν τυχαία, αλλά καρπός προσωπικού ενδιαφέροντος και ακαδημαϊκής αναζήτησης.

Η ενασχόληση με το πρακτικό μέρος της μελέτης αποτέλεσε ιδιαίτερη πρόκληση. Η πρόσβαση σε χρηματοοικονομικά δεδομένα, ειδικά σε δεδομένα ποιότητας και έκτασης κατάλληλης για ανάλυση, αποδείχθηκε χρονοβόρα και σε αρκετές περιπτώσεις κοστοβόρα διαδικασία. Παρ' όλες τις δυσκολίες, η εμπειρία ήταν εξαιρετικά διδακτική και με βοήθησε να αποκτήσω πολύτιμες δεξιότητες στην πράξη.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά την οικογένειά μου και τους φίλους μου για τη διαρκή στήριξή τους καθ' όλη τη διάρκεια αυτής της διαδρομής. Ιδιαίτερες ευχαριστίες οφείλω στον Δημήτρη Ζησιάδη για την ηθική υποστήριξη και την ενθάρρυνση σε κρίσιμες στιγμές. Τέλος, θα ήθελα να εκφράσω την ευγνωμοσύνη μου προς τον επιβλέποντα καθηγητή μου, κ. Γεώργιο Μαστοράκη, για την καθοδήγησή του, τις παρατηρήσεις του και την πολύτιμη συμβολή του στην ολοκλήρωση της παρούσας εργασίας.

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει τη συμβολή της Συμπεριφορικής Χρηματοοικονομικής και της Ανάλυσης Δεδομένων στην πρόβλεψη των τάσεων των χρηματοπιστωτικών αγορών, με έμφαση στην ανάλυση συναισθήματος (Sentiment Analysis) και στην εφαρμογή της Θεωρίας Πληροφορίας (Information Theory). Βασικός στόχος είναι η αξιολόγηση της αποτελεσματικότητας συνδυαστικών μοντέλων που ενσωματώνουν συναισθηματικά και πληροφοριακά χαρακτηριστικά.

Η μεθοδολογία περιλαμβάνει τη συλλογή και επεξεργασία δεδομένων τιμών μετοχών και σχετικών ειδήσεων για την περίοδο 2022–2024 (εκπαίδευση) και το πρώτο εξάμηνο του 2025 (δοκιμή). Η ανάλυση επικεντρώνεται σε έξι τεχνολογικές μετοχές: Apple (AAPL), Microsoft (MSFT), Nvidia (NVDA), Tesla (TSLA), Meta (META) και Amazon (AMZN). Χρησιμοποιήθηκαν τέσσερα μοντέλα μηχανικής μάθησης: Logistic Regression, Random Forest, XGBoost και Support Vector Machines (SVM). Η αξιολόγηση πραγματοποιήθηκε με χρήση των δεικτών Accuracy, Precision, Recall και F1-score, καθώς και μέσω confusion matrices και ανάλυσης της σημασίας των χαρακτηριστικών.

Τα αποτελέσματα ανέδειξαν το μοντέλο SVM ως το καλύτερο συνολικά, καταγράφοντας τη μεγαλύτερη ακρίβεια (0.769) και F1-score (0.696), καθώς και την υψηλότερη επίδοση σε recall (0.717). Το Random Forest κατέγραψε σταθερές αποδόσεις σε όλους τους δείκτες, με αξιοσημείωτη ισορροπία μεταξύ precision και recall, ενώ το Logistic Regression υστέρησε. Η προσθήκη του ημερήσιου δείκτη συναισθήματος FinBERT και της εντροπίας Shannon ενίσχυσε σημαντικά την προγνωστική ικανότητα όλων των μοντέλων.

Συμπερασματικά, η εργασία αναδεικνύει την αξία μιας πολυδιάστατης προσέγγισης που συνδυάζει στοιχεία ψυχολογίας, ανάλυσης συναισθήματος και θεωρίας πληροφορίας στην πρόβλεψη αγοραστικών τάσεων. Τονίζεται η ανάγκη για περαιτέρω έρευνα σε υβριδικά μοντέλα που θα ενισχύσουν την προγνωστική ισχύ έναντι της πολυπλοκότητας και της αβεβαιότητας των χρηματοπιστωτικών αγορών.

Λέξεις κλειδιά: Συμπεριφορική Χρηματοοικονομική, Ανάλυση Συναισθήματος, Θεωρία Πληροφορίας, Μηχανική Μάθηση, Πρόβλεψη Αγορών.

Abstract

This thesis explores the contribution of Behavioral Finance and Data Analytics to forecasting financial market trends, with a focus on sentiment analysis and the application of Information Theory. The primary objective is to assess the effectiveness of combined predictive models that incorporate emotional and informational features.

The methodology involves the collection and processing of stock price data and financial news for the period from 2022 to 2024 (training set) and the first half of 2025 (test set). The analysis focuses on six major technology stocks: Apple (AAPL), Microsoft (MSFT), Nvidia (NVDA), Tesla (TSLA), Meta (META), and Amazon (AMZN). Four machine learning models were used: Logistic Regression, Random Forest, XGBoost, and Support Vector Machines (SVM). Evaluation metrics included Accuracy, Precision, Recall, and F1-score, complemented by confusion matrices and feature importance analysis.

The results highlighted the superior performance of the SVM model, which achieved the highest accuracy (0.769), F1-score (0.696), and recall (0.717). Random Forest demonstrated consistent performance across all metrics, showing a good balance between precision and recall, while Logistic Regression lagged. The inclusion of FinBERT-based daily sentiment and Shannon entropy significantly enhanced the predictive accuracy of all models.

In conclusion, this research confirms the value of a multidimensional approach that integrates psychological insights, sentiment analysis, and information-theoretic measures for market trend prediction. Further research into hybrid models is encouraged better to address the inherent complexity and uncertainty of financial markets.

Keywords: Behavioral Finance, Sentiment Analysis, Information Theory, Machine Learning, Market Prediction.