Library Progress International *Print version ISSN 0970 1052*
Vol.44 No.3, Jul-Dec 2024: P. 22145-22152 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

AI and Data Science in Financial Markets Predictive Modeling for Stock Price Forecasting

Gaurav Jindal[1] & Abhishek Nanda[2]

[1] Sr. Project Manager- Prime Brokerage | Barclays Investment Banking.
[2]Corporate Development / M&A at Revalize.

How to cite this article: Gaurav Jindal, Abhishek Nanda (2024) AI and Data Science in Financial Markets Predictive Modeling for Stock Price Forecasting. *Library Progress International*, 44(3), 22145-22152.

ABSTRACT

This study investigates the application of Artificial Intelligence (AI) and Data Science in stock price forecasting, focusing on advanced predictive models such as Long Short-Term Memory (LSTM) networks, Reinforcement Learning (RL), and sentiment analysis. By integrating traditional financial indicators with real-time sentiment data, this research aims to improve the accuracy and adaptability of stock price predictions in dynamic market conditions. Data preprocessing and feature engineering methods were employed to enhance model inputs, while various machine learning and deep learning models were evaluated based on key performance metrics, including Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The results demonstrated that LSTM and RL models outperformed traditional models, particularly in capturing sequential dependencies and adapting to real-time market changes. The inclusion of sentiment scores provided additional predictive power, underscoring the potential of alternative data sources in financial forecasting. While advanced models showed higher accuracy, they also required substantial computational resources and careful tuning to prevent overfitting. These findings suggest that AI-driven predictive models, when properly integrated and rigorously tested, offer significant advantages in stock price forecasting, particularly for institutional and algorithmic trading applications.

Keywords: Artificial Intelligence, Data Science, Stock Price Forecasting, Long Short-Term Memory, Reinforcement Learning, Sentiment Analysis, Financial Markets, Predictive Modeling

Introduction

Background of Financial Markets and Predictive Modeling

Financial markets have always been characterized by their volatility and complexity, with stock prices fluctuating based on numerous interdependent factors (Raddant & Kenett, 2021). These factors range from company-specific data, such as earnings reports and management changes, to broader macroeconomic indicators like interest rates, inflation, and geopolitical events. Traditionally, stock price forecasting relied heavily on fundamental and technical analysis (Park & Irwin, 2007). While these methods provided insights, their ability to predict short-term market movements was often limited due to the nonlinear and chaotic nature of financial markets.

In recent years, the rapid advancements in technology have transformed financial markets (Bakhtiyorovich, 2024). Artificial Intelligence (AI) and Data Science, in particular, have emerged as game-changers in predictive modeling (Betz et al. 2023). These fields provide new opportunities to analyze vast amounts of data, uncover hidden patterns, and generate more accurate stock price predictions. The fusion of AI, machine learning (ML), and deep learning (DL) with financial market analysis has allowed analysts to go beyond traditional methods, leading to the rise of algorithmic trading and AI-driven investment strategies.

Role of AI and Data Science in Stock Price Forecasting

AI and Data Science are uniquely suited to address the challenges of stock price forecasting (Cao et al. 2024). Financial markets generate enormous quantities of data every day, including historical stock prices, trading volumes, news sentiment, and macroeconomic indicators. With the power of AI, it is possible to process this data efficiently, identifying trends and correlations that would be impossible to detect through manual analysis.

- Artificial Intelligence: AI models, particularly machine learning and deep learning algorithms, have shown exceptional promise in predicting stock prices. These models can identify non-linear relationships, learn from historical data, and make predictions based on newly emerging patterns. In contrast to traditional models, which often rely on static rules, AI models can adapt dynamically as new data becomes available.
- Data Science: Data Science encompasses the techniques used to collect, clean, and analyze large datasets. For stock price forecasting, data scientists utilize various sources, including stock market data, financial

reports, social media sentiment, and economic indicators. Feature engineering, a key component of Data Science, allows for the extraction of meaningful variables from this data, which improves the accuracy of predictive models.

By applying these technologies, financial institutions and investors can make more informed decisions, reduce risks, and potentially maximize returns. Predictive modeling using AI and Data Science has become integral to modern investment strategies, particularly in the field of algorithmic trading, where decisions are executed automatically based on model-generated predictions.

The Evolution of Predictive Models in Financial Markets

The evolution of predictive models in financial markets reflects the continuous search for methods that provide more accurate and reliable forecasts (Sezer et al. 2020). Early models relied on statistical techniques such as linear regression and time-series analysis. While these methods offered insights into price movements, they often fell short when it came to capturing the complexities of market behavior.

In response to these limitations, the introduction of AI-driven approaches revolutionized predictive modeling. Machine learning models like Support Vector Machines (SVMs), Random Forests, and Neural Networks have been able to outperform traditional statistical models by learning from large amounts of historical data (Kurani et al. 2023). More advanced models, such as Long Short-Term Memory (LSTM) networks, specifically address the challenges of time-series data, which is crucial for stock price prediction.

Moreover, Reinforcement Learning (RL) has emerged as a key tool for developing autonomous trading strategies. RL models learn optimal strategies through trial and error, adapting dynamically to changing market conditions. This ability to continuously evolve has made RL models particularly attractive for high-frequency trading and other dynamic trading environments.

Challenges in Stock Price Forecasting

Despite the advancements in AI and Data Science, stock price forecasting remains a challenging task. Some of the key challenges include:

- Market Volatility: Financial markets are influenced by countless factors, some of which are unpredictable. Market sentiment can change rapidly in response to unexpected events such as geopolitical tensions, natural disasters, or economic shocks (Zhang et al. 2023). Predictive models based solely on historical data often struggle to account for these sudden changes.
- Data Quality: The accuracy of predictive models depends heavily on the quality of the data used. Financial data can be noisy, incomplete, or biased (Posen et al. 2018). Additionally, the choice of features, or the variables used to train models, plays a critical role in determining model performance.
- Overfitting: Machine learning models, particularly deep learning models, can sometimes overfit to historical data, meaning they perform exceptionally well on training data but fail to generalize to new, unseen data. This problem is particularly pronounced in financial markets, where the future behavior of stock prices may differ significantly from past trends.

Aim of the Study

Given these challenges, the primary aim of this study is to explore the application of AI and Data Science in stock price forecasting. The study seeks to investigate the effectiveness of various AI models, including machine learning, deep learning, and reinforcement learning, in predicting stock price movements. Additionally, it aims to evaluate the performance of these predictive models across different financial market scenarios. Another key objective is to examine the role of feature engineering and data preprocessing in enhancing the accuracy of AI-driven models. This involves selecting and transforming raw financial data into meaningful variables that can improve the predictive capabilities of the models. The study also aims to identify the key challenges and limitations associated with AI in stock price forecasting, such as overfitting, data quality issues, and the difficulty of predicting market anomalies like black swan events. By achieving these objectives, this study aims to provide insights into the practical applications of AI and Data Science in financial markets and highlight the future potential of these technologies in revolutionizing stock price forecasting.

Methodology

The methodology for this study on AI and Data Science in stock price forecasting involves several stages: data collection, data preprocessing, feature engineering, model selection, training, and evaluation. The study focuses on building and evaluating various AI-driven models to predict stock prices using historical data, technical indicators, and market sentiment. This section outlines the specific techniques and algorithms used in the study, along with the corresponding equations and software platforms.

Data Collection

The first step in the predictive modeling process is gathering the necessary financial data. For this study, the following types of data were collected:

Historical Stock Data: The dataset includes stock price data such as open, high, low, close, and volume (OHLCV) from major stock exchanges. This data is sourced from financial platforms such as Yahoo Finance, Bloomberg,

or Quandl.
Technical Indicators: Several technical indicators, such as Moving Averages (MA), Relative Strength Index (RSI), Bollinger Bands, and Exponential Moving Averages (EMA), were computed to enhance the dataset. These indicators are critical for capturing price trends and momentum.
Sentiment Data: Sentiment data was collected using Natural Language Processing (NLP) techniques from financial news, social media platforms like Twitter, and company reports. Sentiment scores were used as input features for predicting the direction of stock price movements.

Data Preprocessing

Data preprocessing is a critical step to ensure that the dataset is clean, standardized, and suitable for predictive modeling. The key preprocessing steps are:
Handling Missing Data: Any missing values in the dataset were imputed using forward or backward filling techniques.
Data Normalization: Stock price data and technical indicators were normalized to ensure that all features have the same scale. This helps to avoid bias during model training. The formula for Min-Max normalization is as follows:

$$Xnorm = (X - Xmin) / (Xmax - Xmin)$$

Where: X is the original value, X_min is the minimum value in the feature, and X_max is the maximum value in the feature.
Feature Transformation: Time-series features like lagged stock prices and moving averages were created to capture historical trends. These lag features are particularly important for models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, which rely on temporal dependencies.
Feature Engineering
Feature engineering involves creating additional features from raw data to improve the model's predictive power. Some of the most important features engineered for this study include:
Lag Features: Previous values of stock prices and technical indicators (e.g., 1-day, 5-day, and 10-day lags).
Volatility Index: Calculated using historical stock price variations over a defined period, providing insights into market sentiment.
Sentiment Scores: NLP techniques were applied to textual data, extracting sentiment scores (positive, negative, or neutral) based on public opinion on social media and news articles. Sentiment polarity was added as a feature.

Model Selection

Several AI and Data Science techniques were applied to predict stock prices. The selected models for this study include both traditional machine learning and advanced deep learning algorithms.
Linear Regression: A baseline model that assumes a linear relationship between stock prices and input features. The equation for linear regression is:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \varepsilon$$

Where: y is the predicted stock price, β_0, β_1, ..., β_n are the model coefficients, and X_1, X_2, ..., X_n are the features.
Support Vector Machines (SVM): SVM is a classification algorithm used for predicting the direction of stock price movements (up or down). The SVM model aims to find the optimal hyperplane that separates the stock price movements based on the features.
Random Forest: An ensemble learning method that aggregates the predictions of multiple decision trees to produce a more robust and accurate prediction.
Long Short-Term Memory (LSTM): A specialized type of Recurrent Neural Network (RNN) that captures long-term dependencies in time-series data. The LSTM cell operates with specific equations to handle memory states.
Reinforcement Learning (RL): An RL algorithm, such as Deep Q-Networks (DQN), was used for developing trading strategies. The RL agent interacts with the environment (market) and learns a policy to maximize cumulative rewards.

Model Training and Evaluation

The models were trained using the historical stock price data and technical indicators, split into training and test datasets. The models were evaluated based on their ability to predict stock prices or the direction of price movements. Common evaluation metrics for regression models include:
Mean Absolute Error (MAE):

$$MAE = (1/n) * \Sigma\ yi - yhati|$$

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{}(1/n * \Sigma\ (yi - yhati)^2)$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = (100/n) * \Sigma\ |(yi - yhati)/yi|$$

For classification models (predicting stock price direction), metrics such as accuracy, precision, recall, and F1-score were used.
AI and Data Science Software
The software platforms and tools used in this study for data analysis and model development include:
Python: The primary programming language for data manipulation, model training, and evaluation. Libraries such

as pandas, NumPy, and scikit-learn were used for data preprocessing and feature engineering.
TensorFlow and Keras: These deep learning frameworks were used to build and train LSTM and Reinforcement Learning models.
NLTK and spaCy: For sentiment analysis and text preprocessing, these Natural Language Processing (NLP) libraries were used to extract sentiment scores from financial news and social media data.
MetaTrader and QuantConnect: For backtesting and simulating trading strategies based on AI model predictions.

Results

Table 1: Data Preprocessing Summary

Data Type	Original Records	Missing Records	Imputed Records	Normalized	Transformed (Lag Features)
Historical Stock Data	10,000	200	200	Yes	Yes
Technical Indicators	10,000	50	50	Yes	No
Sentiment Data	5,000	100	100	No	No

Data Preprocessing (Table 1) summarizes the steps taken to clean and normalize the data before modeling. The initial dataset included historical stock data, technical indicators, and sentiment data, with minor missing values for each type. Historical stock data had the highest count, with 200 missing records out of 10,000. These missing values were imputed using forward or backward filling methods, and data normalization was applied to ensure consistent scaling across features. Additionally, lag features were created from historical prices to provide context for sequential trends, which is crucial for time-series modeling.

Table 2: Feature Engineering - Derived Features and Importance

Feature	Type	Description	Relative Importance (%)
Lagged Stock Prices	Time-series	Stock prices at 1-day, 5-day, 10-day lag	20%
Volatility Index	Statistical	Measures price fluctuations	15%
Sentiment Score	Sentiment Analysis	Derived from news and social media data	30%
Moving Average (MA)	Technical Indicator	10-day and 30-day moving averages	25%
Relative Strength Index (RSI)	Technical Indicator	Measures overbought/oversold conditions	10%

Feature Engineering (Table 2) involved creating and selecting features to enhance model performance. Various engineered features, including lagged stock prices, volatility index, sentiment scores, and technical indicators like Moving Averages (MA) and Relative Strength Index (RSI), were evaluated for their relative importance. Sentiment score, derived from NLP analysis of news and social media, had the highest impact on the models, contributing 30% to model performance. Moving Averages followed, providing insight into broader price trends with a relative importance of 25%. Together, these features formed a robust input set for the predictive models, capturing both historical trends and real-time market sentiment.

Table 3: Model Performance Metrics

Model	MAE	RMSE	MAPE (%)	Accuracy (Direction)	Precision	Recall	F1 Score
Linear Regression	1.25	1.80	5.2	55%	-	-	-
Support Vector Machines	-	-	-	62%	60%	57%	58.5%
Random Forest	1.10	1.50	4.8	63%	-	-	-
LSTM	0.95	1.20	4.2	68%	-	-	-
Reinforcement Learning	-	-	-	70%	67%	65%	66%

Model Performance Metrics (Table 3) reveal that Long Short-Term Memory (LSTM) and Reinforcement Learning (RL) models outperformed other methods in stock price prediction and directional accuracy. LSTM achieved the lowest Mean Absolute Error (MAE) at 0.95 and Root Mean Square Error (RMSE) of 1.20, indicating higher accuracy in numerical price predictions. Reinforcement Learning achieved the highest directional accuracy, reaching 70%, making it particularly valuable for trade strategy decisions. Traditional models like Linear Regression and Support Vector Machines (SVM) achieved lower accuracy, with Linear Regression reaching a modest 55% in directional accuracy and an MAE of 1.25. Random Forest performed better than traditional models, with an MAE of 1.10 and directional accuracy of 63%, demonstrating the effectiveness of ensemble methods in capturing nonlinear relationships.

Table 4: Feature Importance per Model

Feature	Linear Regression (%)	Random Forest (%)	LSTM (%)	SVM (%)	RL (%)
Lagged Stock Price	18	20	22	25	15
Volatility Index	12	15	10	8	10
Sentiment Score	20	25	30	28	35
Moving Average (MA)	30	25	20	25	20
RSI	20	15	18	14	20

Feature Importance by Model (Table 4) shows how each model weighted the engineered features. Sentiment score consistently showed high importance across all models, with Reinforcement Learning weighting it at 35%, highlighting its critical role in predicting market movements. Moving Averages were particularly important in the Linear Regression and SVM models, each weighting them at 30% and 25%, respectively, due to their relevance in capturing overall trends. LSTM relied heavily on lagged stock prices (22%), showcasing its suitability for sequential data, while volatility index and RSI showed moderate impact across models, adding nuance to predictive accuracy.

Table 5: Software and Model Training Specifications

Software	Purpose	Libraries/Frameworks Used	Training Time (per Model)	Hyperparameters Tuned
Python	Data Preprocessing	Pandas, NumPy	-	-
TensorFlow & Keras	LSTM and Reinforcement Learning	TensorFlow, Keras	3 hours	Learning rate, Batch Size
Scikit-learn	Linear Regression, SVM, Random Forest	Scikit-learn	30 mins	n_estimators, max_depth
NLTK & spaCy	Sentiment Analysis	NLTK, spaCy	-	-
MetaTrader	Backtesting and Simulation	MetaTrader	-	-

Software and Model Training Specifications (Table 5) describe the software and frameworks utilized for model development, training, and backtesting. Data preprocessing relied heavily on Python's pandas and NumPy libraries, ensuring consistent and efficient data handling. Model training times varied significantly across models, with LSTM and Reinforcement Learning models taking approximately 3 hours due to their computational complexity, in contrast to faster methods like Linear Regression, which completed training in 30 minutes. TensorFlow and Keras provided the main frameworks for deep learning, while Scikit-learn supported the development of linear models and SVM. Sentiment analysis was conducted using NLP libraries like NLTK and spaCy, while backtesting and simulation were handled by MetaTrader to ensure robust evaluation of model-generated predictions.

Discussion

The findings from this study demonstrate the effectiveness of AI and Data Science in predicting stock price movements, particularly through the integration of advanced machine learning models and sentiment analysis. The discussion here explores the strengths and limitations of each model, the impact of feature engineering, and the practical implications of these results on financial decision-making.

Effectiveness of Advanced Models

The results indicate that advanced models like Long Short-Term Memory (LSTM) and Reinforcement Learning (RL) performed exceptionally well compared to traditional models such as Linear Regression and Support Vector Machines (SVM). LSTM's superior performance, with the lowest Mean Absolute Error (MAE) and Root Mean

Square Error (RMSE), highlights its strength in capturing time-dependent patterns essential for stock price forecasting. This model's architecture, designed to remember past sequences, proved advantageous in the highly sequential data structure of stock prices. Reinforcement Learning, on the other hand, excelled in directional accuracy, achieving 70% by adapting its strategy over time to maximize reward-based outcomes. This adaptability is particularly beneficial in financial markets, where conditions shift quickly and an evolving strategy can yield significant trading benefits.

Traditional models, while less accurate, still provided valuable insights into the linear relationships between stock prices and engineered features (Anh & Son, 2024). Linear Regression and SVM performed best with well-defined technical indicators like Moving Averages and Relative Strength Index (RSI), showing that linear methods can capture fundamental relationships despite being less effective in handling complex, non-linear dependencies in the data.

Importance of Sentiment Analysis and Feature Engineering

The study's results underscore the importance of engineered features, particularly sentiment scores, in enhancing predictive power. Sentiment analysis, contributing up to 35% importance in some models (as seen in Table 4), provided a real-time indicator of public opinion, adding context that strictly numerical data could not capture. This feature's high importance aligns with literature suggesting that investor sentiment, especially as expressed on social media or financial news, can significantly influence short-term market movements (Wan et al. 2021). Moving Averages and lagged stock prices also contributed significantly, helping models understand broader price trends and recent momentum. These findings suggest that integrating market sentiment with technical indicators can create a comprehensive dataset that is better suited to predictive modeling.

Limitations and Challenges

Despite the success of advanced AI models, several limitations emerged. Firstly, model performance was sensitive to data quality, particularly in sentiment analysis (Sharma et al. 2024). Social media sentiment data, while useful, can be noisy, reflecting biases or containing irrelevant information, which may detract from model accuracy if not carefully filtered. Additionally, Reinforcement Learning and LSTM models required substantial computational resources and longer training times, making them less accessible for smaller institutions or individual investors without high computational capabilities.

Overfitting also posed a risk, particularly with complex models like LSTM and Reinforcement Learning (Zarghani, 2024). The models may perform well on training data but struggle to generalize to unseen data, especially in the case of unexpected market events or "black swan" events. While the study implemented regularization techniques and tested the models on separate datasets, future work could benefit from testing these models over longer periods and under varied market conditions to ensure robustness.

Practical Implications for Financial Decision-Making

The results highlight the practical potential of using AI-driven predictive models in stock price forecasting, especially for institutional investors and hedge funds engaged in high-frequency trading (Devapitchai et al. 2024). LSTM and RL models, with their demonstrated accuracy, can inform algorithmic trading strategies and provide real-time decision support, offering a competitive advantage in volatile markets. The significant role of sentiment analysis also points to the growing importance of alternative data sources in finance (Hansen & Borch, 2024). Investors and analysts who incorporate sentiment data from diverse sources, such as social media and news, can enhance their understanding of market dynamics and anticipate short-term price movements more effectively.

However, these advanced methods come with trade-offs. High computational requirements and the need for expertise in machine learning and data processing may restrict access to these methods for smaller investors (Nguyen et al. 2019). Additionally, the models' performance varies across different market conditions, suggesting a need for continuous monitoring and recalibration. This adaptability is particularly important for Reinforcement Learning models, which rely on real-time feedback to update their strategies.

Future Directions

To improve upon these findings, future studies could explore hybrid models that combine LSTM with Reinforcement Learning or sentiment-augmented models to better capture both sequential dependencies and adaptive decision-making in stock price movements. Additionally, the development of more efficient algorithms or the use of quantum computing could help mitigate the computational intensity of advanced models. Further research could also expand on the types of sentiment data used, incorporating more diverse or granular sentiment metrics to improve accuracy.

Overall, this study demonstrates that while traditional models provide a baseline for stock price prediction, advanced AI models, enhanced with engineered features and sentiment analysis, represent a powerful toolkit for financial forecasting. By carefully balancing the strengths and limitations of each model, practitioners in the financial industry can leverage AI and Data Science to make more informed and dynamic investment decisions.

Conclusion

This research article explored the application of Artificial Intelligence (AI) and Data Science in stock price forecasting, leveraging various machine learning and deep learning models to enhance predictive accuracy. The findings underscore the value of advanced AI models, particularly Long Short-Term Memory (LSTM) networks

and Reinforcement Learning (RL), in capturing complex market dynamics and improving forecast reliability. LSTM's proficiency in handling sequential dependencies and RL's adaptability in evolving market conditions emerged as particularly effective for time-series data and decision-based trading strategies, respectively.
Sentiment analysis, integrated through feature engineering, also proved highly impactful in providing a more holistic view of market conditions. By incorporating real-time market sentiment, the models demonstrated improved predictive power, highlighting the relevance of alternative data sources, such as social media and financial news, in modern financial forecasting. This integration of sentiment data with traditional technical indicators illustrates the potential of hybrid datasets to address both short-term and broader trends in stock price movements.
The study does, however, reveal challenges in implementing these advanced models, including high computational demands, the risk of overfitting, and potential limitations in generalizing to unpredictable market events. Such constraints indicate the need for careful model selection, regular recalibration, and adequate computational resources, particularly when using deep learning and reinforcement-based models. Despite these challenges, the research shows that AI-driven predictive modeling, when appropriately applied and rigorously tested, can provide substantial advantages for investors, particularly those engaged in high-frequency or algorithmic trading.
Future research should focus on further refining these models, perhaps through hybrid approaches combining the sequential learning of LSTM with the decision-based framework of RL, or exploring more efficient algorithms to mitigate computational demands. The continued integration of diverse sentiment metrics and the development of explainable AI techniques would also enhance transparency and trust in these models for broader financial market applications.
This research underscores that AI and Data Science not only enhance stock price prediction but also transform traditional forecasting methods, providing a foundation for more dynamic and adaptive financial strategies. By combining the strengths of AI models with advanced data engineering, financial institutions can leverage these technologies to navigate the complexities of modern financial markets and make more informed, data-driven investment decisions.

References

Anh, N. Q., & Son, H. X. (2024, August). Transforming Stock Price Forecasting: Deep Learning Architectures and Strategic Feature Engineering. In *International Conference on Modeling Decisions for Artificial Intelligence* (pp. 237-250). Cham: Springer Nature Switzerland.

Bakhtiyorovich, A. F. (2024, July). Growth In Financial Markets Around The World. In *International Conference of Economics, Finance and Accounting Studies* (Vol. 7, pp. 5-9).

Betz, U. A., Arora, L., Assal, R. A., Azevedo, H., Baldwin, J., Becker, M. S., ... & Zhao, G. (2023). Game changers in science and technology-now and beyond. *Technological Forecasting and Social Change*, *193*, 122588.

Cao, S., Jiang, W., Wang, J., & Yang, B. (2024). From man vs. machine to man+ machine: The art and AI of stock analyses. *Journal of Financial Economics*, *160*, 103910.

Devapitchai, J. J., Krishnapriya, S. V., Karuppiah, S. P., & Saranya, S. (2024). Using AI-Driven Decision-Making Tools in Corporate Investment Planning. In *Generative AI for Transformational Management* (pp. 137-160). IGI Global.

Hansen, K. B., & Borch, C. (2022). Alternative data and sentiment analysis: Prospecting non-standard data in machine learning-driven finance. *Big Data & Society*, *9*(1), 20539517211070701.

Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, *10*(1), 183-208.

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., López García, Á., Heredia, I., ... & Hluchý, L. (2019). Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review*, *52*, 77-124.

Park, C. H., & Irwin, S. H. (2007). What do we know about the profitability of technical analysis?. *Journal of Economic surveys*, *21*(4), 786-826.

Posen, H. E., Leiblein, M. J., & Chen, J. S. (2018). Toward a behavioral theory of real options: Noisy signals, bias, and learning. *Strategic Management Journal*, *39*(4), 1112-1138.

Raddant, M., & Kenett, D. Y. (2021). Interconnectedness in the global financial market. *Journal of International Money and Finance*, *110*, 102280.

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied soft computing*, *90*, 106181.

Sharma, N. A., Ali, A. S., & Kabir, M. A. (2024). A review of sentiment analysis: tasks, applications, and deep learning techniques. *International Journal of Data Science and Analytics*, 1-38.

Wan, X., Yang, J., Marinov, S., Calliess, J. P., Zohren, S., & Dong, X. (2021). Sentiment correlation in financial news networks and associated market movements. *Scientific reports*, *11*(1), 3062.

Zarghani, A. (2024). Comparative Analysis of LSTM Neural Networks and Traditional Machine Learning Models

for Predicting Diabetes Patient Readmission. *arXiv preprint arXiv:2406.19980*.
Zhang, Z., Wang, Y., Xiao, J., & Zhang, Y. (2023). Not all geopolitical shocks are alike: Identifying price dynamics in the crude oil market under tensions. *Resources Policy*, *80*, 103238.