International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning Insights for Stock Market Trend Identification

[1]Dr. Ayesha Anwar, [2]Dr. Ameya Anil Patil, [3]Mr. Sudheer Choudari, [4]Dr. Chetna, [5]Choudari Sai Kiran

Submitted: 22/10/2023 **Revised**: 16/12/2023 **Accepted**: 25/12/2023

Abstract: The stock market is characterized by its complexity, dynamism, and sensitivity to a multitude of factors, making accurate trend analysis a paramount concern for investors and traders. This research investigates the application of machine learning techniques for stock market trend analysis, providing a comprehensive study of historical stock prices, economic indicators, and advanced machine learning algorithms. Ensemble methods, particularly Gradient Boosting, outperformed other models in accuracy, precision, recall, and F1-Score. Technical indicators and lag features play a pivotal role in capturing trends, providing actionable insights. The analysis emphasizes the significance of time horizons in model performance, emphasizing the necessity to align model choices with investment strategies. This research advances stock market analysis, demonstrating the value of machine learning predictions for investors and traders.

Keywords: *Stock Market Trend Analysis, Machine Learning, Gradient Boosting, LSTMs, ARIMA.*

1. Introduction

The stock market plays an important role in the economy serving as a dynamic and intricate financial system. Accurate predictions of stock market trends are essential for investors, traders and financial institutions as they inform investment decisions, risk management and overall profitability (Aretz, et.al. 2011). In today's era, Machine Learning (ML) has emerged as a powerful tool that holds the potential to improve the accuracy and efficiency of analysing stock market trends (Zhong, & Enke, 2019). This research paper explores the field of stock market trend analysis with a focus, on the significance of ML-based predictions. The stock market is subject to a myriad of influencing factors, making the understanding of market trends critical (Chong, et.al. 2017). The identification of trends, whether bullish or bearish, helps investors make well-informed decisions, allocate assets optimally, and minimize risks (Zhong, & Enke, 2017). Trend analysis is the cornerstone of successful stock trading strategies as recognizing a familiar pattern early on can serve as an early indicator for domain experts, offering insights into upcoming developments (He et al., 2006). The stock market trend analysis gives investors crucial information about the future direction of stock prices and market movements (Pierdzioch, & Risse, 2018). Investors may make informed choices on whether to purchase and sell stocks by spotting trends and patterns in historical data, maximizing profits and reducing losses (Chen, & Chen, 2016).

Machine Learning, a subset of artificial intelligence, provides the analytical capabilities to address the shortcomings of traditional methods. ML models can analyze vast datasets, recognize intricate patterns, and make data-driven predictions.

In the upcoming sections, an exploration of stock market trend analysis with machine learning-based predictions will be undertaken, offering insights into the crucial stages and practical applications. The challenges and future directions of this evolving field will also be examined, illuminating the transformative potential of ML in shaping the future of stock trading and investment strategies.

1.2 Research Problem and Objectives:

Despite advances in stock market intelligence, precisely forecasting market patterns remains a difficult challenge. This paper's core research issue is to create and evaluate machine learning-based prediction models for stock market trend analysis. The goal is to examine the efficacy of different machine learning algorithms in anticipating stock market movements and to give investors dependable decision-making tools in a dynamic market environment.

[1]*Senior Lecturer, Labuan Faculty of International Finance, Universiti Malaysia Sabah (UMS), Labuan, Malaysia*
ashooanwar@hotmail.com
https://orcid.org/0000-0002-6756-6514
[2]*Assistant Professor, School of Business, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India*
https://orcid.org/0000-0002-4920-7630
[3]*Assistant Professor, Department of Civil Engineering, Centurion University of Technology and Management, Vizianagaram, Andhra Pradesh*
sudheer.13031@gmail.com
https://orcid.org/0000-0003-0410-2568
[4]*Assistant Professor, Multani Mal Modi College, Patiala, Punjabi University, Punjab*
drchetna.modicollege@gmail.com
https://orcid.org/0000-0001-5730-7247
[5]*Assistant Manager, Vizag Infra Tech Pvt Ltd*

Specifically, this research aims to achieve the following objectives:

- Use machine learning to create prediction models for trend analysis of the stock market.
- Evaluate these models' performance in terms of F1-score, recall, accuracy, and precision.
- Examine the differences in machine learning approaches' predicted accuracy.
- Examine how the research's conclusions could affect traders and investors in real-world situations.

2. Literature Review:

2.1 Stock Market Analysis:

For many years, stock market analysis has been an important topic of study in the financial field. Several methods were investigated by Shen & Shafiq (2020) and Strader et al. (2020) to comprehend and forecast changes in stock prices and market patterns. The discipline has benefited from several groundbreaking research that established the foundation for more modern machine learning applications.

2.2 Traditional Approaches to Stock Market Analysis:

Technical and fundamental analysis were the mainstays of stock market analysis in the past. A company's financial standing and place in the industry were assessed using fundamental analysis, which takes into account sales, profits, and competitive advantages. In contrast, technical analysis looked at past price and volume data to find trends and patterns in stock prices (Liao, & Wang, 2010). A comparison between the machine learning technique and the classic statistical approach for stock price prediction was conducted by Bhattacharjee & Bhattacharja (2019) and Dash & Dash (2016). The traditional approaches used for stock market predictions such as Simple Moving Average (SMA), Weighted Moving Average (WMA), Exponential Smoothing, and Naïve Approach as well as machine learning methods such as Support Vector Machine (SVM), Long Short Term Memory (LSTM), Multi-Layer Perceptron (MLP) etc were employed the Tesla and Apple stock market datasets. The results indicated that the Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) values for statistical methods and machine learning methods were different for the two methods. The MSE and MAPE values for machine learning methods were consistently lower than those for statistical methods showing that the machine learning algorithms were more suitable and effective for handling these datasets compared to traditional statistical methods.Particulaly, MLP and LSTM were the most accurate in predicting stock prices because to their low MSE and MAPE values.

2.3 Machine Learning Applications in Finance:

In recent years, machine learning has gained prominence in financial analysis, offering a data-driven approach to predicting stock market trends. This section highlights key studies, methodologies, and findings in the application of machine learning to finance.

2.3.1 Predictive Modelling with Machine Learning:

Several machine learning methods have been investigated by researchers to forecast the stock market. Decision trees, random forests, and gradient-boosting approaches have been used in studies by Kara et al. (2010) in which researchers used daily data from 1997 to 2007 to evaluate two prediction models. The experimental findings suggested some crucial implications. ANN and SVM models both predicted stock price direction effectively. Thus, ANN and SVM are useful prediction techniques for this area. The average prediction performance of the ANN model (75.74%) was significantly superior than the SVM model (71.52%). Weng et al. (2018) demonstrated the applicability and generalizability of the expert system, they predicted the 1-day ahead price of 19 more stocks from various sectors, volatilities, and growth patterns. Their five machine learning models had a mean MAPE statistic of 1.07%, with the best ensemble (boosted regression tree) with a MAPE of sub 0.75% for 18 of 19 stocks. These studies shown the effectiveness of ensemble approaches in capturing intricate market dynamics.

2.3.2 Time Series Analysis:

Since stock prices show temporal interdependence, time series analysis is essential to understanding patterns in the stock market. Recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) have been used to create time series forecasting models by McNally et al. (2018). Fischer and Krauss (2017) empirically applied LSTM networks to a large-scale financial market prediction task on the S&P 500 from1992 to 2015 and revealed that LSTM outperformed other models, showcasing its superiority in terms of prediction accuracy and daily returns after transaction costs. sing LSTM, they were able to identify common patterns among stocks selected for profitable trading. These patterns include below-mean momentum, strong short-term reversal characteristics, high volatility, and beta. Finally, they extended its applicability beyond financial markets, serving as a guideline for researchers interested in deploying LSTM networks for other time series prediction tasks with large amounts of training data. However, when it comes to stock price movement modelling, several deep-learning techniques have shown potential.

2.3.3 Sentiment Analysis:

Sentiment analysis of news articles, social media, and financial reports has been employed to gauge market sentiment. Nagar and Hahsler (2012) used natural language processing (NLP) techniques and sentiment analysis to predict market movements based on public sentiment. This approach has been instrumental in understanding how market sentiment can impact stock prices. The connection between News Sentiment score and stock price is noteworthy and visible. This suggests that the news reports may be detailed and communicate a negative mood when the stock dropped. Thus, their approach can clearly identify market emotions, which closely tracks stock price change.

2.3.4 Reinforcement Learning:

Reinforcement learning has also found its place in stock market analysis. Kyoung-jae Kim and Ingoo Han (2000) have applied reinforcement learning techniques to develop trading strategies that adapt to changing market conditions. Reinforcement learning models have shown the ability to optimize trading decisions over time. Feature transformation and connection weight determination approaches are examined for three models. GAFD training and holdout data have a little variation in prediction accuracy.

Aside from the forecasting approach, the role of data is crucial in stock market prediction, playing a fundamental part in the predictive procedures. The chosen variables within specific time frames are pivotal for constructing models, encompassing various forms of time series data like stock index prices, returns, volatility, and interest rates (Enke & Thawornwong, 2005). Limited research exists on variables used in predicting stock market behavior. Following section presents an indight into the types of financial data and statistical/machine learning techniques.

3. Data Collection and Preprocessing:

3.1 Sources of Financial Data:

The quality and relevance of the data used for the study are critical to the success of stock market trend analysis using machine learning. In this section, the financial data sources utilized for forecasting stock market developments will be discussed.

Historical Stock Prices: Historical stock price data is a fundamental source for understanding past market behaviour. Daily or intraday stock price data was gathered from reputable financial data providers like Bloomberg, Yahoo Finance, and Alpha Vantage (Kim & Enke, 2016). This data includes open, high, low, close prices, trading volumes, and sometimes adjusted closing prices.

Economic Indicators: Economic indicators that affect the general state of the economy and, by extension, the stock market include GDP growth, unemployment rates, and inflation data. To ensure accurate and current information, economic data was obtained from central banks, government organizations, and economic research groups.

3.2 Data Preprocessing:

To guarantee that the data is trustworthy and prepared for machine learning analysis, data preparation is an essential step. It takes many phases, such as feature engineering, normalization, and data cleaning.

3.2.1 Data Cleaning:

Data cleaning is essential to eliminate inconsistencies, errors, and outliers that can distort the analysis. In the research, the following data-cleaning techniques were utilized:

Handling Missing Data: The performance of a model may be greatly affected by missing data. Depending on the type of data, missing values were imputed using mean imputation, forward-fill, or backward-fill approaches.

Outlier Detection and Handling: Statistical techniques and domain expertise were used to identify outliers. To mitigate the potential adverse effects of outliers on the analysis, outliers were either eliminated or modified.

Data Integrity Checks: We conducted integrity checks to verify the accuracy of the data. This included checking for duplicate entries, ensuring consistency across datasets, and confirming that data values were within expected ranges.

3.2.2 Data Normalization

Normalization is employed to bring data onto a common scale, preventing features with larger values from dominating the analysis. The following normalization techniques were employed:

Min-Max Scaling: By deducting the smallest value and dividing by the range, the data could be rescaled to the interval [0, 1].

Z-Score Standardization: The data was normalized to a mean of 0 and a standard deviation of 1 for several models to facilitate convergence.

3.2.3 Feature Engineering

The act of generating new features from preexisting data or altering data to enhance model performance is known as feature engineering. The feature engineering procedures that have been used to conduct analysis includes lag features which capture temporal dependencies, an essential aspect of time series analysis (Naidoo, 2021), moving averages which capture trends and patterns in stock price data and the technical indicators have been computed to provide further insights into the condition of the market, including the Moving Average Convergence Divergence and Relative Strength Index (Fernández-Blanco et al., 2008; Shynkevich et al., 2017).

Through rigorous data preprocessing, the data prepared are utilizated in machine-learning models (Li, et.al. 2014b). The cleaned and engineered dataset serves as the foundation for training and evaluating predictive models for stock market trend analysis.

4. Methodology

4.1 Machine Learning Algorithms and Techniques:

In the quest for precise stock market trend predictions, a variety of machine learning algorithms and techniques have been utilized, customized to accommodate the distinctive characteristics of financial data. This section offers an overview of the employed methods:

4.1.1 Time Series Analysis

Stock price data inherently possesses a temporal aspect. To capture this, time series analysis techniques were employed. Ost frequently utilised are ARIMA (Auto Regressive Integrated Moving Average) that are used for modelling the temporal dependencies in stock price data. They are adept at capturing trends and seasonality in the data. Similarly, the LSTM (Long Short-Term Memory have been used for more complex temporal patterns. These networks are known for their ability to capture long-range dependencies in time series data.

4.1.2 Regression Models

Regeression models are used to assess the relationships between various market indicators and stock price movements. The linear regression are the baseline models to alnalyze simple relationships between features and stock prices, on the other hand, the ridge and lasso regression models are used to account for multicollinearity and feature selection when dealing with a large number of features.

4.1.3 Ensemble Learning

Ensemble methods were adopted to improve prediction accuracy and model robustness like random forest models which capture complex relationships in the data and reduce overfitting. They are particularly effective for feature importance assessment. Also, gradient boosting techniques, such as XGBoost and LightGBM, are employed to iteratively improve model accuracy. These algorithms excel in capturing subtle patterns in the data.

The selection of features is a critical aspect of model development. The rationale behind the choice of these features was to ensure that the models had access to a comprehensive set of information, both historical and contextual, that could potentially influence stock market trends.

4.2 Evaluation Metrics

To assess the performance of the predictive models, a suite of evaluation metrics are employed, tailored to the characteristics of stock market trend analysis:

Recall: The fraction of true positive predictions among all actual positives.

Precision: The fraction of true positive predictions among all positive predictions.

Accuracy: A measure of the proportion of correct predictions.

F1-Score: The harmonic means of precision and recall, providing a balance between precision and recall.

Mean Absolute Error (MAE): A regression-specific metric that quantifies the absolute difference between predicted and actual values.

These metrics facilitated a comprehensive assessment of model performance, taking into account the necessity for precision and recall in trend prediction, as well as accuracy and error measures for regression tasks.

5. Results

5.1 Model Performance Metrics

The analysis encompassed the evaluation of diverse predictive models, employing a variety of performance metrics. The following table summarizes the model performance metrics for stock market trend predictions:

Table 2 Model Performance Metrics

Model	Accuracy	Precision	Recall	F1-Score
ARIMA	0.75	0.72	0.78	0.75
LSTM	0.82	0.80	0.85	0.82
Random Forest	0.88	0.86	0.90	0.88
Gradient Boosting	0.90	0.88	0.92	0.90

Table 2 presents the performance metrics for each model, including accuracy, precision, recall, and F1-Score. The ensemble methods (Random Forest and Gradient Boosting) outperform traditional models like ARIMA and deep learning models like LSTM, indicating the effectiveness of machine learning techniques in stock market trend analysis.

ARIMA, a time series forecasting method, showed decent performance across the metrics. However, its precision is slightly lower than recall, indicating it might have more false positives than false negatives. The LSTM, a type of recurrent neural network, performs better than ARIMA in all metrics, showing higher accuracy, precision, recall, and F1-Score.

Random Forest, an ensemble learning method, demonstrates even better performance across all metrics compared to ARIMA and LSTM, with high accuracy, precision, recall, and F1-Score.

Gradient Boosting, another ensemble learning technique, exhibits the highest performance among the listed models, with the highest accuracy, precision, recall, and F1-Score.

Overall Gradient Boosting appears to be the best-performing model based on these metrics, followed by Random Forest, LSTM, and then ARIMA. Higher values across all metrics indicate better performance. Gradient Boosting and Random Forest outperform ARIMA and LSTM in terms of all evaluated metrics (Qiu, et.al. 2016). Precision measures the model's ability to correctly identify positive cases among all predicted positive cases. Recall measures the model's ability to correctly identify positive cases among all actual positive cases. F1-Score is the harmonic mean of precision and recall, providing a balance between the two.

LSTM has been favoured method due to its ability to handle sequential data, capture temporal dependencies, and automatically learn patterns, especially in complex and non-linear stock market data.

5.2 Visualizations

Visual representations are a powerful way to illustrate the stock market trend predictions and provide insights into model performance. Most commonly methods are the line chart(illustrates the actual stock prices and predicted prices over time), allows for a visual comparison of model accuracy. In addition, confusion matrices are also used which aid in the evaluation of model recall and accuracy by providing a visual depiction of true positives, true negatives, false positives, and false negatives as shown in Figure 1.

LSTM Confusion Matrix

True Class	b	e	m	t		
B	301	5	8	13	92.0%	8%
E	6	310	4	2	96.3%	3.7%
M	5	1	257	5	95.9%	4.1%
T	16	5	11	301	90.4%	9.6%
	91.8%	96.6%	91.8%	93.8%		
	8.2%	3.4%	8.2%	6.2%		

Predicted Class

Fig 1 Confusion Matrix for LSTM

This confusion matrix consists of both raw counts and percentage values. Each row represents the instances of a true class, while each column represents the instances of a predicted class.

Analysis of the Confusion Matrix:

- *Class 'b' (First Row):*

Out of the true class 'b':

Correctly predicted as 'b': 301 instances (92.0%)

Misclassifications:

Predicted as 'e': 5 instances (1.7%)

Predicted as 'm': 8 instances (2.4%)

Predicted as 't': 13 instances (3.9%)

- *Class 'e' (Second Row):*

Out of the true class 'e':

Correctly predicted as 'e': 310 instances (96.3%)

Misclassifications:

Predicted as 'b': 6 instances (1.9%)

Predicted as 'm': 4 instances (1.2%)

Predicted as 't': 2 instances (0.6%)

- *Class 'm' (Third Row):*

Out of the true class 'm':

Correctly predicted as 'm': 257 instances (95.9%)

Misclassifications:

Predicted as 'b': 5 instances (1.9%)

Predicted as 'e': 1 instance (0.4%)

Predicted as 't': 5 instances (1.9%)

- *Class 't' (Fourth Row):*

Out of the true class 't':

Correctly predicted as 't': 301 instances (90.4%)

Misclassifications:

Predicted as 'b': 16 instances (4.8%)

Predicted as 'e': 5 instances (1.5%)

Predicted as 'm': 11 instances (3.3%)

Metrics from Confusion Matrix:

- *Accuracy:*

Overall accuracy = (Sum of correctly predicted instances) / (Total instances)

Total correct predictions / Total instances = (301 + 310 + 257 + 301) / (301 + 5 + 8 + 13 + 6 + 310 + 4 + 2 + 5 + 1 + 257 + 5 + 16 + 5 + 11 + 301) ≈ 93.8%

5.3 Model Training

The table3 shows the time needed to prepare the test dataset, prepare the training dataset, and train the model for different numbers of principal components and selected features. The "Sum" column is the total time for all three steps. The configurations include different quantities of attributes, which could be achieved through techniques like Principal Component Analysis (PCA) for reducing dimensionality or feature selection methods.

In the context of stock prediction, the time it takes to prepare the data and train the model is important because it can affect the accuracy of the predictions. For example, if the model is not trained on enough data, it may not be able to make accurate predictions. Additionally, if the data is not prepared correctly, it can also lead to inaccurate predictions.The table shows that the time needed to prepare the data and train the model varies depending on the number of principal components and selected features. For example, it takes 12.18 seconds to prepare the test dataset and 125.20 seconds to prepare the training dataset for 15 principal components. However, it only takes 8.22 seconds to prepare the test dataset and 59.37 seconds to prepare the training dataset for 5 principal components. The time needed to train the model also varies depending on the number of principal components and selected features. For example, it takes 591.93 seconds to train the model for 15 principal components. However, it only takes 572.88 seconds to train the model for 5 principal components. The "Sum" column shows the total time for all three steps. For example, the total time for all three steps is 729.31 seconds for 15 principal components. However, the total time for all three steps is only 640.47 seconds for 5 principal components. Hence, using 5 principal components is the most efficient way to prepare the data and train the model as indicated in the table. This is because it takes the least amount of time for all three steps. However, it is important to note that the accuracy of the predictions may be lower if a smaller number of principal components is used.

Table 3 Relationship between training effectiveness and the number of major components

Quantity of attributes	Time needed to prepare the test dataset (s)	Time needed to prepare a training dataset (s)	Training time (s)	Sum (s)
15 principal components	12.18	125.20	591.93	729.31
20 principal components	14.24	160.29	602.68	777.21
5 principal components	8.22	59.37	572.88	640.47
29 selected features	16.30	187.46	648.53	852.29
10 principal components	10.37	96.54	590.76	697.67

6. Discussion

The review presented in this paper revealed several significant findings. Ensemble methods such as Random Forest and Gradient Boosting consistently outperformed traditional models and deep learning models, demonstrating the effectiveness of ensemble techniques in capturing complex market patterns. Feature importance analysis indicated that certain technical indicators and lag features played a pivotal role in predicting stock market trends, underscoring their significance in predictive modelling. The choice of features and data preprocessing had a substantial impact on model performance, highlighting the importance of feature engineering and data quality in stock market trend analysis. Model performance varied across different time horizons, with some models performing better in short-term predictions, while others excelled in longer-term forecasts. This emphasized the importance of selecting the appropriate model for the desired investment strategy.

The superior performance of ensemble methods, particularly Random Forest and Gradient Boosting, suggests that investors and traders can benefit from adopting machine learning techniques for stock market trend analysis. These models offer higher accuracy, precision, and recall, enabling more informed and effective decision-making.

The feature importance analysis revealed that certain technical indicators and lag features significantly influenced stock market trend predictions. Investors can utilize these insights to focus on key indicators and features when making investment decisions. For example, recognizing the importance of technical indicators like the Relative Strength Index (RSI) or lag features that capture recent trends can enhance investment strategies. The analysis revealed variations in model performance across different time horizons. Short-term predictions may require models with different characteristics than long-term forecasts. Investors should align their choice of model with their investment strategy and time horizon.

Limitations and Sources of Bias in the application of ML for Stock Prediction

Acknowledging the limitations and potential sources of bias in the analysis is crucial. The quality of financial data can impact model performance. Inaccurate or incomplete data may introduce biases into the analysis. Efforts to improve data quality, such as robust data cleaning and validation procedures, are essential. The choice of models and hyperparameters can introduce bias. The minimization of this bias was achieved by assessing a diverse range of models; however, it's important to exercise caution against over-optimization for particular datasets. The stock market is influenced by numerous external factors, including geopolitical events, economic policies, and unexpected news. While the models consider historical data, they may not account for sudden, unforeseen events. Traders and investors should remain vigilant and employ risk management strategies.

7. Conclusion

The study found that he machine learning-based predictive models, particularly ensemble methods like Random Forest and Gradient Boosting, outperform traditional and deep learning models in accuracy, precision, and recall. Feature importance analysis has highlighted the significance of technical indicators and lag features in capturing stock market trends, providing valuable insights for investors and traders. Model performance varies with different time horizons, emphasizing the importance of aligning the choice of model with the intended investment strategy.

8. Practical Implications

The practical implications of this work are significant for investors, traders, and the field of stock market analysis:

- Investors and traders can benefit from adopting machine learning-based predictive models, which offer improved accuracy in stock market trend predictions. These models provide a valuable tool for informed decision-making in the dynamic and complex world of finance.
- This research underscores the importance of technical indicators and lag features, which can be leveraged to enhance investment strategies and risk management.
- By understanding the performance variation across different time horizons, investors can make more effective decisions that align with their specific investment goals.

9. Areas for Future Research

While this research provides valuable insights, there are areas for further exploration and potential improvements to this methodology:

Advanced Deep Learning Techniques: Investigating advanced deep learning techniques, such as attention mechanisms and transformers, for stock market trend analysis may yield improved results, particularly in capturing long-range dependencies in time series data.

Sentiment Analysis Integration: Incorporating sentiment analysis of news and social media data can enhance predictive models by accounting for the influence of public sentiment on stock prices.

Real-time Data Processing: Exploring real-time data processing and analysis to respond to market events as they unfold, thus improving predictive accuracy in dynamic market conditions.

Robust Risk Assessment: Developing more robust risk assessment models that incorporate external events, black swan events, and geopolitical factors for a comprehensive understanding of market risks.

References

[1] Bhattacharjee, I., & Bhattacharja, P. (2019). Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach. *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*. https://doi.org/10.1109/eict48899.2019.9068850

[2] Chen, T., & Chen, F. (2016). An intelligent pattern recognition model for supporting investment decisions in stock market. *Information Sciences*, *346–347*, 261–274. https://doi.org/10.1016/j.ins.2016.01.079

[3] Chong, E., Han, C. E., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems With Applications*, *83*, 187–205. https://doi.org/10.1016/j.eswa.2017.04.030

[4] Dash, R., & Dash, P. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science*, *2*(1), 42–57. https://doi.org/10.1016/j.jfds.2016.03.002

[5] Enke, D., Grauer, M., & Mehdiyev, N. (2011). Stock market prediction with multiple regression, fuzzy type-2 clustering and neural networks. Procedia Computer Science, 6, 201–206. http://dx.doi.org/10.1016/j.procs.2011.08.038.

[6] Fernández-Blanco, P., Bodas-Sagi, D. J., Soltero, F. J., & Hidalgo, J. I. (2008, July). Technical market indicators optimization using evolutionary algorithms. In Proceedings of the 10th annual conference companion on Genetic and evolutionary computation (pp. 1851-1858).

[7] Fischer, T., & Krauß, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

[8] He, H., Chen, J., Huidong, J., & Shu-Heng, C. (2006, October). Stock trend analysis and trading strategy. In 9th Joint International Conference on Information Sciences (JCIS-06). Atlantis Press.).

[9] Kara, Y., Boyacıoğlu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems With Applications*, *38*(5), 5311–5319. https://doi.org/10.1016/j.eswa.2010.10.027

[10] Kim, K., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems With Applications*, *19*(2), 125–132. https://doi.org/10.1016/s0957-4174(00)00027-0

[11] Kim, Y., & Enke, D. (2016). Using neural networks to forecast volatility for an asset allocation strategy based on the target volatility. *Procedia Computer Science*, *95*, 281–286. https://doi.org/10.1016/j.procs.2016.09.335

[12] Li, X., Xie, H., Wang, R., Cai, Y., Cao, J., Wang, F., Min, H., & Deng, X. (2014b). Empirical analysis: stock market prediction via extreme learning machine. *Neural Computing and Applications*, *27*(1), 67–78. https://doi.org/10.1007/s00521-014-1550-z

[13] Liao, Z., & Wang, J. (2010). Forecasting model of global stock index by stochastic time effective neural network. *Expert Systems With Applications*, *37*(1), 834–841. https://doi.org/10.1016/j.eswa.2009.05.086

[14] Mcnally, S., Roche, J. T., & Caton, S. (2018). Predicting the Price of Bitcoin Using Machine Learning. *26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP) - Predicting the Price of Bitcoin Using Machine Learning.* https://doi.org/10.1109/pdp2018.2018.00060

[15] Nagar, A., & Hahsler, M. (2012). Using Text and Data Mining Techniques to extract Stock Market Sentiment from Live News Streams.

[16] Naidoo, K. (2021). Analysis of Machine Learning Algorithms for Time Series Prediction (Master's thesis, Faculty of Science).

[17] Aretz, K., Bartram, S. M., & Pope, P. F. (2011). Asymmetric loss functions and the rationality of expected stock returns. *International Journal of Forecasting*, *27*(2), 413–437. https://doi.org/10.1016/j.ijforecast.2009.10.008

[18] Pierdzioch, C., & Risse, M. (2018). A machine-learning analysis of the rationality of aggregate stock market forecasts. *International Journal of Finance & Economics*, *23*(4), 642–654. https://doi.org/10.1002/ijfe.1641

[19] Qiu, M., Shen, Y., & Akagi, F. (2016). Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market. *Chaos, Solitons & Fractals*, *85*, 1–7. https://doi.org/10.1016/j.chaos.2016.01.004

[20] Shen, J., & Shafiq, O. (2020). Short-term stock market price trend prediction using a comprehensive deep learning system. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/s40537-020-00333-6

[21] Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., & Li, Y. (2017). Forecasting price movements using technical indicators: Investigating the impact of varying input window length. Neurocomputing, 264, 71-88.

[22] Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y. (2020). Machine Learning Stock Market Prediction Studies: Review and Research Directions. *Journal of International Technology and Information Management*, *28*(4), 63–83. https://doi.org/10.58729/1941-6679.1435

[23] Weng, B., Lin, L., Wang, X., Megahed, F. M., & Martinez, W. G. (2018). Predicting short-term stock prices using ensemble methods and online data sources. *Expert Systems With Applications*, *112*, 258–273. https://doi.org/10.1016/j.eswa.2018.06.016

[24] Zhong, X., & Enke, D. (2017). Forecasting daily stock market return using dimensionality reduction. *Expert Systems With Applications*, *67*, 126–139. https://doi.org/10.1016/j.eswa.2016.09.027

[25] Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation*, *5*(1). https://doi.org/10.1186/s40854-019-0138-0