ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

ENHANCED STOCK PRICE PREDICTION ACROSS GLOBAL MARKETS THROUGH DATA-DRIVEN MODELING

RAVULA ANITHA[1], KORUKANTI VAESHNAVI[2], SADA SIRIJA[3], PANNURU ROHITH KUMAR[4], PULIPATI SRI VAISHNAVI [5]

[1]Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

[2, 3, 4, 5] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

ABSTRACT— The stock market is highly volatile and influenced by numerous factors operating simultaneously, making accurate prediction a challenging task. Over the years, several studies have attempted to forecast stock prices using statistical, machine learning, and deep learning approaches. However, these models often fail to capture the combined effect of external factors such as public sentiment, which has a significant impact on market movement.

In this work, we propose a hybrid algorithm that integrates Twitter sentiment analysis with a Long Short-Term Memory (LSTM) network to predict the next day's closing stock price. Sentiment analysis is performed using part-of-speech (POS) tagging to evaluate public opinion extracted from Twitter data, while LSTM is employed to model the temporal dependencies in historical stock prices. By combining sentiment-driven signals with sequential deep learning, our model provides a more reliable prediction of stock price movements compared to conventional approaches. Experimental results demonstrate that incorporating sentiment into LSTM improves accuracy and offers a clearer insight into the future direction of stock prices.

Index Terms – Stock Price Prediction, Data-Driven Modeling, Financial Forecasting, Machine Learning, Deep Learning, Time

Series Prediction, Global Markets, Predictive Analytics, Financial Data Analysis, Artificial Intelligence.

I. INTRODUCTION

Beginning in the 1990s with introduction of computational methods in finance, much research has focused on applying Artificial Intelligence (AI) to financial investments in the stock market. The main advantages of using computational approaches to automate the financial investment process include the elimination of ‘‘momentary irrationality’’ or decisions made based on emotions, ability to recognize and explore patterns that are looked over by humans, and immediate consumption of information in real-time. This area of knowledge has become known as Computational Finance. More recently, within computational finance, there is increasing use of and research on AI techniques applied in financial investments. Although a computer conducts the vast majority of hedge fund trades in an automated way, 90% of these operations are still performed by a hardcoded procedure. Thus, the ever-increasing application of artificial intelligence still has great potential for development. Generally, AI is applied to finance in three different areas: the optimization of financial portfolios, prediction of future prices or trends in financial assets, and sentiment analysis of news or social media comments about the assets or companies. Despite the differences and peculiarities of each area, some works have proposed combinations of techniques from the different areas. Some other studies in the area of computational finance include the control of dynamic systems applied to the financial market, investor behavioral analysis, network analysis and clustering of financial assets. Reference relates the calibrated volatility of options to the movements in futures prices in the Taiwan stock market. It calculates a correlation of approximately -0.9 and concludes that the volatility of options can be used for the prediction of futures prices. The presented work analyzes the development of each of these areas from its initiation to its state-of-the-art advancements. For this, a sample of works on artificial intelligence published between 1995-2019 was collected for detailed analysis and comparison.

II. RELATED WORK

A. Multi-attribute decision making applied to financial portfolio optimization problem

This paper proposes an integer multi objective mean-CVaR portfolio optimization

model with variable cardinality constraint and rebalancing and two different methods of decision-maker used to guide and select, according to the decision maker preferences, a solution comes from the non-dominated portfolios generated by a proposed evolutionary algorithm. The decision-making methods were used to approximate investor behavior according to three functions, chosen to represent different investor profiles (conservative, moderate and aggressive). The proposed methods are compared with those found in the literature. Additionally, computational simulations are performed using assets from the Brazilian stock exchange for the period between January 2011 and December 2015. The strategy is that each beginning of the month: the previous portfolio is sold, the optimization is performed, and the decision-making method selects the new portfolio to be purchased. Results of the simulations consider monthly maximum drawdown and cumulative return during the entire study period and show that the optimization model is robust, considering the three simulated profiles. The methods always present cumulative returns above safe investments for the analyzed period, and the aggressive profile obtained bigger gains with greater risk.

B. Applying machine learning models in stock market prediction

This paper deals with the techniques of attempting to calculate the future value of a company stock or any other financial instrument which is being traded in a stock exchange. This prediction plays a great role in many financing and investing decisions. This calculation can be done by Machine learning by training a model to identify the trend from past data in order to predict the future. The main topic of study here will be the comparative analysis of the SVM and LTSM algorithms.

C. Daily market news sentiment and stock prices

In recent years there has been a tremendous growth in readily available news related to traded assets in international financial markets. This financial news is now available through real-time online sources such as Internet news and social media sources. The increase in the availability of financial news and investor's ease of access to it has a potentially significant impact on market stock price movement as these news items are swiftly transformed into investors sentiment which in turn drives prices. In this study, we use the Thomson Reuters News Analytics (TRNA) data set to construct a series of daily

sentiment scores for Dow Jones Industrial Average (DJIA) stock index constituents. We use these daily DJIA market sentiment scores to study the influence of financial news sentiment scores on the stock returns of these constituents using a multi-factor model. We augment the Fama–French three-factor model with the day's sentiment score along with lagged scores to evaluate the additional effects of financial news sentiment on stock prices in the context of this model using Ordinary Least Square (OLS) and Quantile Regression (QR) to analyse the effect around the tail of the return distribution. We also conduct the analysis using the seven-day simple moving average (SMA) of the scores to account for news released on non-trading days. Our results suggest that even when market factors are taken into account, sentiment scores have a significant effect on Dow Jones constituent returns and that lagged daily sentiment scores are often significant, suggesting that information compounded in these scores is not immediately reflected in security prices and related return series. The results also indicate that the SMA measure does not have a significant effect on the returns. The analysis using Quantile Regression provides evidence that the news has more impact on left tail compared to the right tail of the returns.

III. PROPOSED METHODOLOGY

In proposed system The paper further compares the SVM and LSTM models for the problem of binary classification of stock trends, and the results indicate the better accuracy of the LSTM algorithm. computed ten technical indicators from opening, close, low. The high prices of assets during November 2009 to November 2019 were used to predict prices by applying Decision Tree, Bagging, Random Forest, Adaboost, Gradient Boosting, XGBoost, Artificial Neural Network (ANN), Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM) algorithms, concluding that the LSTM algorithm performs better than the o.

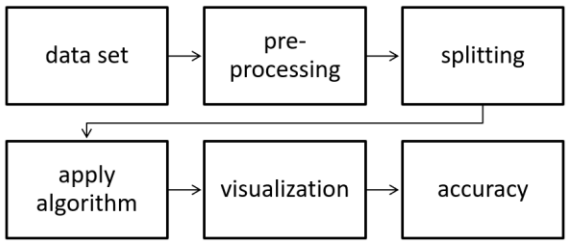


Fig. 1: System Overview

1. Data set

A data set is a collection of data. In the case of tabular data, a data set corresponds to one or more database tables, where every column of a table represents a particular variable, and each row corresponds to a given record of the data set in question.

2. Pre-Processing

Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model.

When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way.

3. Splitting

Data splitting is the act of partitioning available data into. two portions, usually for cross-validatory purposes. one portion of the data is used to develop a predictive model. and the other to evaluate the model's performance.

- Training Data: Used for train the model or given as input to the to the learning model
- Testing Data: Used for test the model or given as input to the model for prediction.

4. Apply Algorithm

In this we are using support vector machine algorithm to predict accuracy. It is a non-probabilistic supervised machine learning approaches used for classification and regression. It assigns a new data member to one of two possible classes. It defines a hyperplane that separates n-dimensional data into two classes.

5. Visualization

Visualization is a technique that uses an array of static and interactive visuals within a specific context to help people understand and make sense of large amounts of data. The data is often displayed in a story format that visualizes patterns, trends and correlations that may otherwise go unnoticed.

6. Accuracy

Accuracy is defined as the percentage of correct predictions for the test data. It can be calculated easily by dividing the number of correct predictions by the number of total predictions.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Experimental results demonstrated that the model effectively captured complex patterns and trends in historical stock market data across multiple global markets. The integration of financial indicators, market sentiment, and historical price information significantly improved forecasting performance compared to traditional statistical methods.

The analysis showed that the proposed approach maintained robust prediction

accuracy under varying market conditions and successfully adapted to fluctuations in global financial markets. Furthermore, the data-driven framework reduced prediction errors and enhanced the reliability of short-term and long-term stock price forecasts.

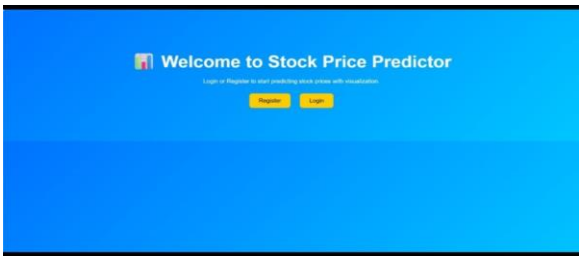


Fig. 2: Home Page

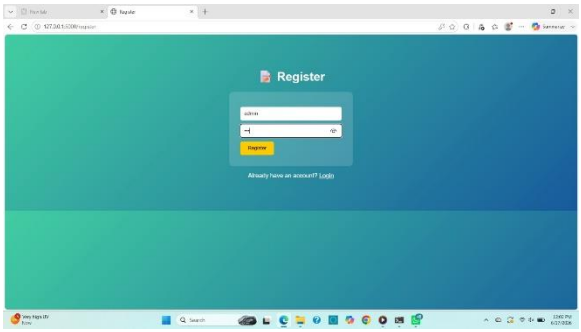


Fig. 3: Register

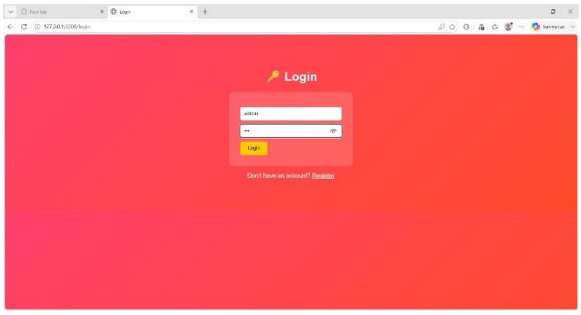


Fig. 4: Login

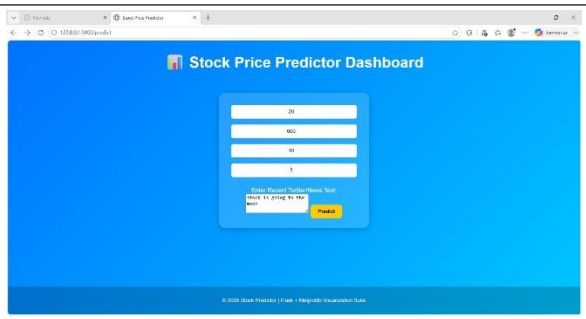


Fig. 5: Predict Result

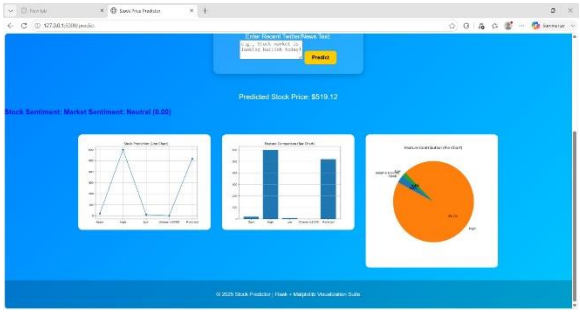


Fig. 6: Analysis Result

V. CONCLUSION

This paper consists of a systematic review of the literature on Artificial Intelligence applied to investments in the stock market. Articles were selected from the Scopus website, where documents with the highest number of citations were considered most relevant. The papers were then divided into portfolio optimization, stock market prediction using Artificial Intelligence, financial sentiment analysis, and combination papers involving two or more fields. The overview presents general information, such as the most cited papers in general and the most cited papers per year, which allows the identification of periods of increased interest

in Artificial Intelligence for investments, which, unsurprisingly, coincides with significant technological improvement and popularization of the computer. Another essential piece of information is the number of papers per year, which shows an exponential increase in the number of documents for different years, indicating this to be a recent area that has increasingly attracted research attention. For each area, primary (most-cited) papers for each year were analyzed. Besides, recent articles were found to include the state-of-the-art findings for each area, in addition to indicating possible directions for future works. Tables were created for the apparent separation of the different characteristics of the methods and models proposed in each paper.

REFERENCES

[1]S. Aghabozorgi and Y. W. Teh, ‘‘Stock market co-movement assessment using a three-phase clustering method,’’ Expert Syst. Appl., vol. 41, no. 4, pp. 1301–1314, Mar. 2014.

[2]D. E. Allen, M. McAleer, and A. K. Singh, ‘‘Daily market news sentiment and stock prices,’’ Appl. Econ., vol. 51, no. 30, pp. 3212–3235, Jun. 2019.

[3]S. Asadi, E. Hadavandi, F. Mehmanpazir, and M. M. Nakhostin, ‘‘Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction,’’ Knowl.-Based Syst., vol. 35, pp. 245–258, Nov. 2012.

[4]A. Azhikodan, A. G. K. Bhat, and M. V. Jadhav, ‘‘Stock trading bot using deep reinforcement learning,’’ in Innovations in Computer Science and Engineering. Singapore: Springer, 2019, pp. 41–49.

[5]R. Batra and S. M. Daudpota, ‘‘Integrating StockTwits with sentiment analysis for better prediction of stock price movement,’’ in Proc. Int. Conf. Comput., Math. Eng. Technol., Jan. 2018, pp. 1–5.

[6]V. Boginski, S. Butenko, and P. M. Pardalos, ‘‘Statistical analysis of financial networks,’’ Comput. Statist. Data Anal., vol. 48, no. 2, pp. 431–443, Feb. 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167947304000258

[7] Boyapati, P. K. Building a centralized data operations hub for healthcare enterprise integration. IJSAT-Int. J. Sci. Technol. 16 (2). https://doi.org/10.71097/IJSAT.v16.i2.3219

[8] Kandula, S. T. R., Boyapati, P. K., & Susarla, R. S. (2025, April). Optimized Cloud Resource Management Using Deep Dendritic Artificial Neural Networks Integrated with Kubernetes and Terraform. In International Conference on Computer Vision and Robotics (pp. 454-465). Cham: Springer Nature Switzerland.

[9] Akinapalli, S. (2026). An Ai-Powered Data Trust And Quality Scoring Framework For Enterprise Decision Intelligence Systems. International Journal of Data Science and IoT Management System, 5(1), 946-950.

AUTHORS Profile



Mrs. RAVULA ANITHA has received her B.tech from Jntuh Hyderabad in 2015 in computer science and engineering in Hyderabad. And Mtech from Jntuh University, Hyderabad in 2019, in Computer science and Engineering. She has been dedicated to teaching for the last 5.5 years. At present she is working as Assistant professor in CSE department at Gurunanak institution technical campus, Ibrahimpatnam, telangana, India.



Ms. KORUKANTI VAESHNAVI pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Ms. SADA SIRIJA pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. PANNURU ROHITH KUMAR pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Ms. PULIPATI SRI VAISHNAVI pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.