ISSN 2277-2685

IJESR/April. 2026/ Vol-16/Issue-2s1/53-58

Syed Shah Mohiuddin Zohaib *et. al.*, / International Journal of Engineering & Science Research

Full Length Article

Enhanced Stock Price Prediction Across Global Markets Through Data-Driven Modeling

Syed Shah Mohiuddin Zohaib[1], Mohd Saif Ur Rahman Khan[2], Mohd Faisal Ahmed[3], Dr. Syed Asadullah Hussaini[4]

[4]Assistant Professor; Department Of Computer Science Engineering, ISL Engineering College Hyderabad India

[1,2,3]B.E Students Department Of Computer Science Engineering ISL Engineering College Hyderabad India

Mail Id;160522733070@islec.edu.in,160522733117sk@gmail.com,160522733090@islec.edu.in

Accepted 26-04-2026

Author(s) Retains the Copyrights of This Article

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.

Keywords- *Stock Price Prediction, Sentiment Analysis, Twitter Data, Long Short-Term Memory (LSTM), Deep Learning, Time Series Forecasting, Natural Language Processing (NLP)*

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.

Objective

The overview presents general information, such as

the most cited papers in general and the most cited paper 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 from 1995 to 2019, 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, for the period between 1995 and 2019.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.

EXISTING SYSTEM:

In Existing system Deep Neural Network (DNN) to classify future trends of asset prices (considering two price directions of the next price). The data set consists of 60 attributes (including returns and technical indicators) of assets belonging to the SPDR S&P 500 ETF between June 2003 and May 2013, with daily frequency. Compared to an Artificial Neural Network (ANN), results show that, although the DNN presents higher accuracy, the ANN provides greater returns and lower risks (variance) in a Stock Market trading simulation. Reference collected open, close, low, and high prices of Yahoo and Microsoft assets from January 2011 to December 2015 and computed five technical indicators as features: momentum, volatility, index momentum, index volatility, stock momentum, stock price volatility.

Existing System Disadvantages:

➢ Less accuracy.

➢ Assurance of proper network structure

➢ The difficulty of showing the problem to the network.

Proposed System

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.

Literature Review:

Title: Multi-attribute decision making applied to financial portfolio optimization problem

Author: G.H.M. Mendonça, F.G.D.C. Ferreira, R.T.N. Cardoso, and F.V.C. Martins.

Year: 2020

Description: This paper proposes an integer multi objective mean-CVaR portfolio optimization model with variable cardinality constraint and rebalancing and two different methods of decisionmaker 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.

Title: Applying machine learning models in stock market prediction

Author: C. K. Vignesh.

Year: 2020

Description: 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.

Title: Daily market news sentiment and stock prices

Author: D. E. Allen, M. McAleer, and A. K. Singh.

Year: 2019

Description: 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.

Title: Stock trading bot using deep reinforcement learning

Author: A. Azhikodan, A. G. K. Bhat, and M. V. Jadhav.

Year: 2019

Description: This paper proposes automating swing trading using deep reinforcement learning. The deep deterministic policy gradient-based neural network model trains to choose an action to sell, buy, or hold the stocks to maximize the gain in asset value. The paper also acknowledges the need for a system that predicts the trend in stock value to work along with the reinforcement learning algorithm. We implement a sentiment analysis model using a recurrent convolutional neural network to predict the stock trend from the financial news. The objective of this paper is not to build a better trading bot, but to prove that reinforcement learning is capable of learning the tricks of stock trading.

Methodology

Module Description

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.

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.

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.

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.

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.

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.

Implementation

Algorithm Used

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

GENERAL

Stock market prediction or forecasting using historical time series has become a technique widely used by researchers and investors to obtain financial profits in stock trading. These predictions, initially carried out by statistical methods, have been increasingly performed by Artificial Intelligence algorithms.

Testing

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner.

There are various types of test. Each test type addresses a specific testing requirement.

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program input produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

System Test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration-oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

Performance Test

The Performance test ensures that the output be produced within the time limits, and the time taken by the system for compiling, giving response to the users and request being send to the system for to retrieve the results.

Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects. The task of the integration test is to check that components or software applications, e.g., components in a software system or – one step up – software applications at the company level – interact without error.

Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

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.

Future Scope

In our future work, It has higher precision, recall, and accuracy. The result of the attrition prediction will be helpful for an organization to reduce the attrition rate of their company.

References

[1] S. Aghabozorgi and Y. W. Teh, ‘‘Stock market co-movement assessment using a threephase 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] V. Boginski, S. Butenko, and P. M. Pardalos, ‘‘Mining market data: A network approach,’’ Comput. Oper. Res., vol. 33, no. 11, pp. 3171–3184, Nov. 2006.

[8] P. J. Bolland and J. T. Connor, ‘‘A constrained neural network Kalman filter for price estimation in high frequency financial data,’’ Int. J. Neural Syst., vol. 8, no. 4, pp. 399–415, Aug. 1997.

[9] J. Bollen, H. Mao, and X. Zeng, ‘‘Twitter mood predicts the stock market,’’ J. Comput. Sci., vol. 2, no. 1, pp. 1–8, Mar. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187775031100007X

[10] A. Booth, E. Gerding, and F. McGroarty, ‘‘Automated trading with performance weighted random forests and seasonality,’’ Expert Syst. Appl., vol. 41, no. 8, pp. 3651–3661, Jun. 2014.

[11] R. C. Cavalcante, R. C. Brasileiro, V. L. F. Souza, J. P. Nobrega, and A. L. I. Oliveira, ‘‘Computational intelligence and financial markets: A survey and future directions,’’ Expert Syst. Appl., vol. 55, pp. 194–211, Aug. 2016.

[12] M. Chang. (2018). How A.I. Traders Will Dominate Hedge Fund Industry.[Online]. Available: https://www.youtube.com/watch?v=lzaBbQKUtAA.

[13] T.-J. Chang, N. Meade, J. E. Beasley, and Y. M. Sharaiha, ‘‘Heuris-tics for cardinality constrained portfolio optimisation,’’ Comput. Oper. Res., vol. 27, no. 13, pp. 1271–1302, Nov. 2000.

[14] T.-J. Chang, S.-C. Yang, and K.-J. Chang, ‘‘Portfolio optimization problems in different risk measures using genetic algorithm,’’ Expert Syst. Appl., vol. 36, no. 7, pp. 10529–10537, Sep. 2009.

[15] C. Chen and Y. Wei, ‘‘Robust multiobjective portfolio optimization: A set order relations approach,’’ J. Combinat. Optim., vol. 38, no. 1, pp. 21–49, Jul. 2019, doi: 10.1007/s10878-018-0364-9.

[16] W. Chen, ‘‘Artificial bee colony algorithm for constrained possibilistic portfolio optimization problem,’’ Phys. A, Stat. Mech. Appl., vol. 429, pp. 125–139, Jul. 2015.

[17] E. Chong, C. Han, and F. C. Park, ‘‘Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies,’’ Expert Syst. Appl., vol. 83, pp. 187–205, Oct. 2017.

[18] Y. Crama and M. Schyns, ‘‘Simulated annealing for complex port-folio selection problems,’’ Eur. J. Oper. Res., vol. 150, no. 3, pp. 546–571, Nov. 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377221702007841

[19] F. D. Paiva, R. T. N. Cardoso, G. P. Hanaoka, and W. M. Duarte, ‘‘Decision-making for financial trading: A fusion approach of machine learning and portfolio selection,’’ Expert Syst.

Appl., vol. 115, pp. 635–655, Jan. 2018.
[20] K. Doerner, W. J. Gutjahr, R. F. Hartl, C. Strauss, and C. Stummer, ‘‘Pareto ant colony optimization: A Metaheuristic approach to multiobjective portfolio selection,’’ Ann. Oper. Res., vol. 131, nos. 1–4, pp. 79–99, Oct. 2004, doi: 10.1023/B:ANOR.0000039513.99038.c6.
[21] M. Ehrgott, K. Klamroth, and C. Schwehm, ‘‘An MCDM approach to portfolio optimization,’’ Eur. J. Oper. Res., vol. 155, no. 3, pp. 752–770, Jun. 2004.
[22] D. Enke and S. Thawornwong, ‘‘The use of data mining and neural networks for forecasting stock market returns,’’ Expert Syst. Appl., vol. 29, no. 4, pp. 927–940, Nov. 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417405001156
[23] O. Ertenlice and C. B. Kalayci, ‘‘A survey of swarm intelligence for portfolio optimization: Algorithms and applications,’’ Swarm Evol. Comput., vol. 39, pp. 36–52, Apr. 2018.
[24] A. Fernández and S. Gómez, ‘‘Portfolio selection using neural networks,’’ Comput. Oper. Res., vol. 34, no. 4, pp. 1177–1191, Apr. 2007.[Online]. Available: http://www.sciencedirect.com/science/article/pii/S0305054805002042
[25] G. H. M. Mendonça, F. G. D. C. Ferreira, R. T. N. Cardoso, and F. V. C. Martins, ‘‘Multiattribute decision making applied to financial portfolio optimization problem,’’ Expert Syst. Appl., vol. 158, Nov. 2020, Art. no. 113527. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417420303511
[26] T. Fischer and C. Krauss, ‘‘Deep learning with long short-term memory networks for financial market predictions,’’ Eur. J. Oper. Res., vol. 270, no. 2, pp. 654–669, Oct. 2018.
[27] T. Geva and J. Zahavi, ‘‘Empirical evaluation of an automated intraday stock recommendation system incorporating both market data and textual news,’’ Decis. Support Syst., vol. 57, no. 1, pp. 212–223, Jan. 2014.
[28] R. H. Golan and W. Ziarko, ‘‘A methodology for stock market analysis utilizing rough set theory,’’ in Proc. Conf. Comput. Intell. Financial Eng., Apr. 1995, pp. 32–40.
[29] E. Hadavandi, H. Shavandi, and A. Ghanbari, ‘‘Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting,’’ Knowl.-Based Syst., vol. 23, no. 8, pp. 800–808, Dec. 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0950705110000857
[30] M. Hagenau, M. Liebmann, and D. Neumann, ‘‘Automated news reading: Stock price prediction based on financial news using context-capturing features,’’ Decis. Support Syst., vol. 55, no. 3, pp. 685–697, Jun. 2013.
[31] G. Hassan and C. D. Clack, ‘‘Multiobjective robustness for portfolio optimization in volatile environments,’’ in Proc. 10th Annu. Conf. Genetic Evol. Comput., 2008, pp. 1507–1514.
[32] C.-M. Hsu, ‘‘A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming,’’ Expert Syst. Appl., vol. 38, no. 11, pp. 14026–14036, May 2011.
[33] C.-M. Hsu, ‘‘A hybrid procedure with feature selection for resolving stock/futures price forecasting problems,’’ Neural Comput. Appl., vol. 22, nos. 3–4, pp. 651–671, Mar. 2013.
[34] C. Huang, D. Yang, and Y. Chuang, ‘‘Application of wrapper approach and composite classifier to the stock trend prediction,’’ Expert Syst. Appl., vol. 34, no. 4, pp. 2870–2878, May 2008.