

## **A STUDY ON ARTIFICIAL INTELLIGENCE FOR STOCK MARKET PREDICTION**

Mr. C. Santhosh Kumar Reddy, MCA \*1, Mr. G. Venkateshwarlu, MCA, MTech \*2  
Mr. Madhavi. S, MSc(CS) \*3

- \*1. Lecturer, Dept. of computer science, Siva Sivani Degree college, Kompally, Sec'Bad-100.
- \*2. Lecturer, Dept. of computer Science, Siva Sivani Degree College, Kompally, Sec'Bad-100.
- \*3. HOD, Dept. of computer Science, Siva Sivani Degree College, Kompally, Sec'Bad-100.

### **ABSTRACT**

Shares of publicly traded corporations can be purchased and sold on the stock market. All sellers and buyers attempt to forecast changes in the stock market price in order to maximise gains and minimise losses. We will discuss and introduce a potential method for making highly accurate stock movement predictions in this model. In order to solve the issue of stock market prediction, artificial intelligence (AI) methodologies are used in this study. Technical and fundamental analyses are the two main types of analysis that can be used to estimate stock market predictions.

Regression machine learning (ML) techniques are used in the technical analysis approach to forecast the trend of the stock price at the conclusion of the business day by using historical price data. On the other hand, the fundamental analysis uses machine learning algorithms for classification to categorise public opinion based on news and social media. Accurate share price forecasting can be a valuable resource for stock market companies and offer practical answers to the problems encountered by individual stock market investors. The aim of this study was to model and forecast a stock market index's future price using artificial intelligence (AI) approaches. Based on past price data, three artificial intelligence techniques—neural networks (NN), support vector machines (SVMs), and neuro-fuzzy systems—are used to anticipate the future price of a stock market index. Techniques from artificial intelligence are employed as financial time series forecasting tools because they can account for the intricacies of the financial system.

### **Keywords**

Stocks, AI, data, Machine learning, time series prediction, technical analysis, sentiment embedding, financial market.

## Introduction

Since the beginning of time, investors have sought to predict the stock market. Each dollar that is traded on the exchange on a daily basis represents the hopes of an investor to make a profit in one way or another. It provides enticing promises of wealth and power, should an investor be able to correctly predict market changes. The process of estimating future stock market prices is known as the stock market estimate. Strong, precise, and efficient are anticipated. The system ought to function in accordance with and be closely related to real-world events. The opinions of thousands of investors typically influence the direction of the stock market. AI has a big potential impact. In numerous research publications, machine learning techniques are devised to assess the accuracy of AI predictions in stock market trading.

The main goals of the ML algorithms used for this purpose are to identify data patterns, calculate investment risk, and forecast future investment returns. The two main theoretical theories developed by this field are the Adaptive Market Hypothesis (AMH) and the Efficient Market Hypothesis (EMH). The path that market prices follow is constantly volatile. According to this theory, there isn't a way to "beat the market." The abundant behavioural finance principles and the evidence EMH are supposed to be correlated, according to the AMH. It should be possible to forecast future market behaviour by using the AMH statement as a guide. Technical analysis bases its trading decisions solely on the historical price of the stock and uses mathematical indicators that are extracted from the stock price. The money flow index (MFI), moving average convergence/divergence (MACD), and relative strength index (RSI) are some of these indicators. These days, both fundamental and technical analyses are seeing a rapid strengthening of ML methods' ability to handle the stock market prediction challenge.

For instance, the F-Score Piotroski model is used to assess the true share value of organisations. Their approach was predicated on nine variables that were extracted from a company's financial records and were categorised into three primary areas: operating efficiency, liquidity, and profitability. This study describes a Machine Learning (ML) technique that will be trained on publicly available stock data in order to gather intelligence and utilise the knowledge gathered to make precise predictions. Upon conducting a thorough analysis of multiple algorithms and their resilience across diverse areas, it was determined that the Artificial Neural Network (ANN) was the most suitable algorithm.

## **SURVEY OF LITERATURE**

The paper "Machine Learning Approach in Stock Market Prediction" was written by authors Raut Sushrut Deepak, Shinde Isha Uday, Dr D. Malathi. According to the article, while estimating share prices, a high degree of accuracy and precision are essential factors to take into account. Many share market institutions and individual investors employ time series analysis, technical analysis, and fundamental study of firms when making predictions.

The paper "Stock Price Prediction Using Technology Analysis and Mechanical Learning" was written by author Jan Ivar Larsen. The study claims that past stock prices can be utilised to forecast changes in stock prices in the future.

## **Existing System**

The different forecasting algorithms in the current system can be categorised into queues (AR, MA, ARIMA, ARMA) and incompatible models (ARCH, GARCH, Neural Network). To advance the gauge model, artificial neural networks (ANN) and Random Forest were utilised, along with Naive Bayes, the closest neighbour k (k-NN), Support Vector Machine (SVM), Linear Regression, and Artificial Neural Network (ANN). Regarding stock values, there is consensus.

## **METHODOLOGY**

The following modules make up the system:

- Information Gathering Data in the .csv format is gathered from a number of sources, including Google Finance and Yahoo Finance. Google News api is used by the news data collection system.
- Data analysis, manipulation, and visualisation to prepare it for use in malicious algorithms and models, the obtained data has been purified and pre-processed. and data visualisation is also done in addition to this.

Construct a Model Machine learning algorithms that can be utilised for predictions are created, built, and trained using the cleaned and pre-processed data.

- Estimate Results Following a successful model build, the next step is to forecast an outcome pattern for a certain stock and assess the forecasts' accuracy.
- Estimate the sum of all algorithmic outcomes. Based on real-time data that is fed to it, the system predicts the output after building models and combining their results for improved prediction and high accuracy.

Mr. C. Santhosh Kumar Reddy / International Journal of Engineering & Science Research  
We hope to use a variety of machine learning techniques in this suggested system to forecast future share values. With the use of various historical data sets, we were able to train and test the machine learning algorithms in this suggested system in order to estimate future share prices. We tested and trained the machine learning models using EOD (End of the Day) data from previous years.

We employed a variety of machine learning frameworks and libraries to achieve the system's goal: math is used to perform mathematical operations on data, Numpy and Pandas are used for data manipulation and visualisation, while Sklearn, Kears, and Tensorflow are used for machine learning models. Old news headlines that were utilised to analyse sentiment were the other dataset.

The primary library was the NumPy library, which is widely used to clean, modify, and transform datasets into formats that are directly usable with machine learning models. The other was the actual computation, estimation, and prediction tool, the sklearn library. We collected historical share market data from a variety of open and public web sources. Depending on the requirements, 80 percent or more of the information was used to train machine learning models, with the remaining portion being used for testing and validation. In order to generate results for the test dataset, supervised machine learning models typically evaluate and retain patterns and correlations in the train and validation datasets.

The system performs feature extraction and other data pre-processing tasks to prepare the data for model training. Pre-processing of the data involved using the pandas Python package to merge many datasets into a single dataframe. It was possible to extract features from this prepared data set. Date, close, open, high, low, volume, delivery percentage, number of trades, turnover, and other derived attributes were among the features of these dataframes. We utilised all features to train a random forest model and forecast the object variable. We also used a subset of features to predict the share values for the next day using the LSTM and SVM algorithms. On both the test dataset and real values, we also assessed the prediction accuracy.

Aspects of research such as pre-processing and data manipulation are included in the method that is being discussed. The Google News API is used by the system to retrieve real-time

Mr. C. Santhosh Kumar Reddy / International Journal of Engineering & Science Research  
stock market news data, and an artificial neural network (ANN) is used to analyse the  
sentiment of news items.

## CONCLUSION

By estimating the accuracy parameters, it was discovered that the best approach for stock market prediction is to combine the predictions of multiple algorithms, such as ANN, Random Forest, SVM, and LSTM, and to take into account all the factors that influence the price of the stock market, including news sentiment, currency fluctuations, commodity prices, and other data from the international stock exchange market. All of this was done with high accuracy. Prediction accuracy is increased when other variables impacting the stock price are taken into account.

Greater profits can be made by trying to anticipate the stock market only when it is more predictable and avoiding it when it is more uncertain. For example, making predictions only when the standard deviation was under the threshold values increased our algorithm's accuracy by 10% or more. Because the algorithms are trained on a huge amount of historical data and then selected to be evaluated on sample data, they are an incredible resource for investors, financial institutions, and share market investments.

## Future Scope

Predictive accuracy can be increased by using more detailed data, such as data collected every minute, every second, or tick by tick. This is because such data contains more insightful information than data collected every day. Future work will include forecasts and suggestions tailored to each person's desire for risk and return as well as portfolio management to help reduce the risk of stock market investing.

Order placement and trade execution processes can be automated to follow recommendations from artificial intelligence (AI) and machine learning (ML) algorithms. This eliminates human mistake and human-generated biases like greed and fear, leading to quicker and more lucrative trading.

## References

- 1) Lufuno Ronald Marwala.  
<https://core.ac.uk/download/pdf/39667613.pdf>
- 2) Sohrab Mokhtari Electrical and Computer Engineering Florida International University Miami, USA, Kang K. Yen Electrical and Computer Engineering Florida

Mr. C. Santhosh Kumar Reddy / International Journal of Engineering & Science Research  
International University Miami, USA and Jin Liu Electrical and Computer  
Engineering Florida International University Miami, USA.

<file:///C:/Users/SSDC%20LAB%2029/Downloads/mokhtari-2021-ijca-9213471.pdf>

- 3) Akash Pate, Devang Pate and Seema Yadav.

<https://deliverypdf.ssrn.com/delivery.php?ID=190081025071014125092006119019065121042037029042091050024117027086002126092080008075126060062111031022046118105081123098011100049015010026004066088126101089068080108023077048127105006016087078101000092125020108020099118123000014015112029097117120126026&EXT=pdf&INDEX=TRUE>

IJESR