

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

EDIA

#### International Journal of Innovative Research in SCIENCE | ENGINEERING | TECHNOLOGY

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

## Impact Factor: 8.118

9940 572 462

6381 907 438

 $\bigcirc$ 



٢



e-ISSN: 2319-8753, p-ISSN: 2320-6710 www.ijirset.com | Impact Factor: 8.118



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105282

### Stock Market Price Trend Prediction using Machine Learning

Rakesh S D<sup>1</sup>, Prof. Md. Irshad Hussain B<sup>2</sup>

P.G. Student, Department of Master Applications, University B. D. T College of Engineering, Davanagere,

Karnataka, India<sup>1</sup>

Assistant Professor, Department of Master Applications, University B. D. T College of Engineering, Davanagere,

Karnataka, India.<sup>2</sup>

**ABSTRACT**: The stock exchange, also known as the stock market, is one of the most complex and sophisticated ways to conduct business. When approached with discipline, the stock market is one of the most effective ways to make large profits. This entity generates revenue and distributes risks for small businesses, brokerage firms, and financial sectors. This project aims to create machine learning models in a Web and Android App to forecast future stock prices for exchange, in order to help make this volatile type of commerce a little more predictable. To avoid the traditional method and obtain results that are entirely based on numbers, a system incorporating a text-based machine learning model and pattern recognition is implemented. The goal is to build a platform for small and inexperienced traders based on existing stock prediction models that take into account variables such as news articles, stock volume, previous close, and so on to forecast future stock market values

**KEYWORDS**:Stock market prediction, machine learning.

#### I. INTRODUCTION

The stock market is one of the oldest methods of trading stocks, making investments, and earning money from companies that sell a portion of themselves on this platform. The stock market facilitates almost all major economic transactions in the world at a dynamic rate known as the stock value, which is based on market equilibrium. If used correctly, this system can be a lucrative investment scheme. Predicting this stock value provides enormous profit opportunities, which is a significant motivator for research in this area. Even a fraction of a second's knowledge of the value of a stock can result in large profits. In the repeated context, a probabilistically correct prediction could be extremely profitable. The allure of finding a solution has prompted researchers in both industry and academia to find a way around issues such as volatility, seasonality, and reliance on time, economics, and the rest of the market. The platform's prices and liquidity, on the other hand, are highly unpredictable, which is where technology can help. One of the more active research areas in the past was stock market forecasting. Small businesses, brokerage firms, and banks all rely on the stock market to generate revenue and spread risk. As a result, forecasting stock prices is a top priority. If we use crowd computing to graph the stock exchange price, we will get an approximation of a real-life graph, but the process will be extremely slow. Because of recent advances in deep learning, trading algorithms can now more accurately predict stock price movements. Unfortunately, there is a significant gap in the real-world implementation of this breakthrough. Furthermore, these cutting-edge technologies are rarely used to help small-scale dealers. The goal of "StockClue" is to create an interactive web and Android app with effective ML models that will assist many small-scale investors in making long-term investments. The app will be a ready-made website with all stock information such as a candlestick graph, open price, close price, volume, relevant news, and so on.

#### II. RELATED WORK

The neural network model used by Lei Shei et al. [1] was designed using news data and Twitter tweets. Each organisation maintains a list of keywords to match each piece of news to the appropriate stocks (e.g., Apple: AAPL, AAPL.O, APPLE, AAPL.N, Apple Inc, etc.). The stock-related tweets were extracted using the Twitter API by matching the firm's hashtags in the tweet content. The model's goal is to forecast a stock price y that is close to the actual stock price y of the company. DeepClue is built with DyNet v1.0, a neural network software library optimised for natural language processing applications. From 2006 to 2015, they looked at S&P 500 stocks in the US stock market.

#### International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

e-ISSN: 2319-8753, p-ISSN: 2320-6710 www.ijirset.com | Impact Factor: 8.118



Volume 11, Issue 5, May 2022

#### DOI:10.15680/IJIRSET.2022.1105282

YangtuoPeng et al. [2] created an input feature vector for DNNs by computing the closing prices of the previous five days. The model searches all financial publications for sentences that mention at least one stock symbol or publicly traded company. Each sentence is divided into samples and labelled with the original article's publication date and stock name. Each example includes a list of sentences that were published on the same day and mentioned the same company or stock. Depending on the closing price the next day, each sample is labelled as positive ("price-up") or negative ("price-down"). DNN is used, which has hidden layers (each with 1024 hidden nodes). The historical price feature serves as a foundation for the addition of financial news-derived elements. The dates of the samples in the test set are used to categorise the DNN outputs. All predictions for unseen stocks are compared to the actual stock movement the following day. This paper relies on financial news data from Reuters and Bloomberg. The CRSP database provides historical stock security data (Center for Research in Security Prices).

Wasiat Khan et al. [3] used Stanford sentimental analysis package of Stanford NLP to analyse financial news Business Insider, awarding positive or negative points for positive or negative words.

Xianghui Yuan et al. [4] proposed a model for forecasting the stock's excess returns for the next month. Over an eightyear period, the financial report, daily opening prices, closing prices, volumes, and other data from the A-share market are used to obtain 60 attributes to be used as input to the model.

When labelling data, JingyiShen et al. [5] created a model to find the price trend by comparing the current closing price to the closing price of n trading days ago. For time-series prediction, they employ LSTM, which ensures that the prediction model can capture both complex hidden patterns and time-series-related patterns. This dataset contains 3558 stocks from the Chinese stock market, as well as data collected via an open-source API and a web-scraping technique to collect data from Sina Finance web pages and the SWS Research website.

Guangyu Ding et al. [6] proposed a long short-term memory network-based associated deep recurrent neural network model with multiple inputs and multiple outputs. The associated network model can predict a stock's opening price, lowest price, and highest price all at once. The experimental data in this paper are actual historical data downloaded from the Internet, Shanghai composite index 000001, and two stocks from the Shanghai stock exchange, PetroChina (stock code 601857) and ZTE (stock code 000063).

The proposed framework by Abdalraouf Hassan et al. [7] combines a joint CNN and RNN framework with a set of feature maps learned by a convolutional layer with long-term dependencies learned via long-short-term memory and use an unsupervised neural language model to train initial word embeddings tuned by deep learning network, then the network's pre-trained parameters are used to initialise the model. The proposed model's performance was evaluated using the Stanford Large Movie Review dataset (IMDB) and the Stanford Sentiment Treebank dataset (SSTb), both of which were derived from Rotten Tomatoes movie reviews. Yash Sharma and colleagues

[8] Applied Glove and discusses its potential application in sentiment analysis. The Glove method word vectors are fed into RNN, and sentiment analysis performs binary classification (Positive and Negative Sentiments).

D V Nagarjana Devi et al. [9] proposed the HARN algorithm, an unsupervised learning method that classifies given sentences using basic sentence structure, domain dictionaries, and predefined polarities.

The author of [10] concludes that applying Machine Learning to stock price prediction can be improved further by delving deeper into deep learning and neural networks. The authors suggest that by switching to Support Vector Machine, they will try to continuously optimise the algorithm for better results. (SVM) and experimenting with various models as well as new and improved features.

According to the authors of [11], some models perform better with historical data than with sentiment data. Fusion algorithms produced more accurate predictions. The GCN-based graph modelled from news co-mentions is the best performing model among those created in [12]. One reason for this could be that the graph is based on causation rather than correlation.

#### III. METHODOLOGY

The proposed method entails creating an interactive online platform (Web and Android App) for stock traders to use in forecasting future stock market values. The App also displays market prices, volume, and related statistics, as well as a prediction for the selected stock. The goal is to build a platform that includes numerous efficient stock market machine learning methods. Finance News and Stock data are two individually learned stock prediction parameters.

- 1. This is an important step in downloading data from the internet. We estimate the financial market value of any stock. As a result, the share value up to the closing date is downloaded from the website.
- 2. The next step is to convert the data value of any stock into a CSV file (Comma Separate Value) so that it can be easily loaded into the algorithm.
- **3.** Login: The Android app has a page for user authentication where users can login to view the current status and forecast of various stocks. If the user is logging in for the first time, he or she can also sign up.

#### International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

よう は IJIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

#### DOI:10.15680/IJIRSET.2022.1105282

- 4. Live Page: A slider with listed stocks, a live stock price display, and predicted values can be found on the live page. Users can use the slider to select the desired stocks, and the corresponding live stock values, predicted graphs, and ML model results will be displayed.
- 5. After selecting the stock dataset file from the folder, the graphs Stock before mapping and Stock after mapping will be displayed.
- 6. In final step algorithm display the predicted value graph of select stock which shows the original value and predicted value of the stock.

Data table	Category	Field
name		
Stock list	Basic data	Stock ID, Stock name, Geographic info, industry, full name, English name,
		Market type, Stock exchange ID, currency. List status, List date, Delist date, if
		the stock is HS constituent.
Trading	Basic data	Stock exchange ID, Calendar date, if the date isopen for trading, previous
Calendar		trading date.
Basic	Basic data	Stock exchange ID, Corporate representative, General manager, Secretary,
information of		Authorized capital, Registration date, Province, city, introduction, Website,
listed		Email, Office address, Number of employees, main business, business scope.
companies		
Renamed	Basic data	Stock Id, Stock name, start date, end date, Announcement date, and Rename
history		reason.
Constituent	Basic data	Stock ID, constituent type, included date, Excluded date, if the stock is new.
stock		
information		
Daily trading	Trading data	Stock ID, Trading date, Opening price, Highest price, Lowest price, closing
data		price, Previous closing price, price change, price change percentage, Volume,
		Amount
Fundamental	Trading data	Stock ID, Trading date, Closing price, Turnover rate, Free turnover rate, Volume
data		ratio, Price to earning ratio, price to earning ratio, TTM, Price-to-book ratio,
		Price to sales ratio, Price to sales TTM, total share capital, Circulating shares,
		Tradable circulating shares, Aggregate market value, circulation market value.

#### Table 1: Dataset overview table with some different categories and subsets of fields

#### **IV. CONCLUSION**

Stock Clue is a platform with many ML models, both text-based and series-based, that help forecast stock price fluctuations. Stock prediction accuracy using machine learning models is more than adequate for real-world implementation. The news vector data indicates stock volatility, allowing users to invest more effectively. The XGBoost model produces more accurate results because it combines news and time series data. Machine learning improves stock prediction accuracy. Small-scale traders with limited time to learn a stock can benefit from this tool at a low cost because most of the systems are open source.

#### References

 DeepClue: Visual Interpretation of Text-based Deep Stock Prediction- Lei Shi, Senior Member, IEEE, ZhiyangTeng, Le Wang, Yue Zhang, and Alexander Binder - IEEE Transactions On Knowledge And Data Engineering.
Leverage Financial News to Predict Stock Price Movements Using Word Embeddings and Deep Neural Networks-YangtuoPeng and Hui Jiang - Department of Electrical Engineering and Computer Science York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada.

[3] Stock market prediction using machine learning classifiers and social media, news-Wasiat Khan, Mustansar Ali Ghazanfar, Muhammad AwaisAzam, Amin Karami, Khaled H. Alyoubi and Ahmed S. Alfakeeh - Journal of Ambient Intelligence and Humanized Computing.

[4] Integrated Long-Term Stock Selection Models Based on Feature Selection and Machine Learning Algorithms for China Stock Market-Xianghui Yuan Jin Yuan ,Tianzhao Jiang , and QuratUl Ain – IEEE.

#### International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

e-ISSN: 2319-8753, p-ISSN: 2320-6710 www.ijirset.com | Impact Factor: 8.118



Volume 11, Issue 5, May 2022

#### DOI:10.15680/IJIRSET.2022.1105282

[5] Short- term stock market price trend prediction using a comprehensive deep learning system – JimgyiShen and M OmairShafq, Shen and Shafiq J Big Data.

[6] Study on the prediction of stock price based on the associated network model of LSTM - Guangyu Ding and Liangxi Qin - International Journal of Machine Learning and Cybernetics (2020).

[7] Convolutional Recurrent Deep Learning Model for Sentence Classification -Abdalraouf Hassan (Member, IEEE) AND AusifMahmood (Senior Member, IEEE) - March 28, 2018.

[8] Vector Representation of Words for Sentiment Analysis Using GloVe - Yash Sharma, GauravAgrawal, Pooja Jain, Member, IEEE, Tapan Kumar Senior Member IEEE Indian Institute of Information Technology, Kota-2017 International Conference on Intelligent Communication and Computational Techniques (ICCT) Manipal University Jaipur, Dec 22-23, 2017.

[9] Sentiment analysis using harn algorithm-D V Nagarjana Devi - Assistant Professor, IIIT, Rgukt,Nuzvid , DR.T.V.RajanikanthProfessor,Snist , Hyderabad.Tg , PantangiRajashekar - UG Student, IIIT,Rgukt,NuzvidGangavarapuAkhil - UG Student , IIIT, Rgukt,Nuzvid.

[10] Pahwa, K. and Agarwal, N., 2019, February. Stock market analysis using supervised machine learning. In 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) (pp. 197-200). IEEE.

[11] Singh, Sukhman, Tarun Kumar Madan, J. Kumar and A. Singh. "Stock Market Forecasting using Machine Learning: Today and Tomorrow." 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) 1 (2019): 738-745.

[12] Patil, P., Wu, C.S.M., Potika, K. and Orang, M., 2020, January. Stock market prediction using an ensemble of graph theory, machine learning and deep learning models. In Proceedings of the 3rd International Conference on Software Engineering and Information Management (pp. 85-92).





Impact Factor: 8.118





## INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com