IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 2, December 2024

# Prediction of the Financial Stock Market: A Comprehensive Analysis of Artificial Intelligence

**MD Shadman Soumik**

MS Student, Department of Information Technology (MSIT)

Washington University of Science and Technology (WUST), 2900 Eisenhower Ave, Alexandria, VA 22314

msoumik.student@wust.edu

**Abstract**: *Since the inception of stock trading, scholars and investors have searched for reliable methods to forecast the course of stock values the next day. Since there are several variables that might influence the stock values of the next day, forecasting stock prices is a challenging undertaking. Stock Market Forecasting (SMF) is a forward-looking process anticipating future stock values, allowing to make sound financial decisions. In order to create predictions, academics and investors have started using machine learning approaches in conjunction with technical indicator analysis. However, the precision of the predictions is lacking. One of the progress in applying ML, particularly LSTM networks, to stock market forecasting lies in automating this process. Human bias implies that the same predictions can be misleading and contribute to the fact that they need to use ML and AI technology. The data used was fetched from finance.yahoo.com, and for confidence in the data, it took steps such as lemmatisation, null value management and deletion of duplicates. A total of four different ML prediction methods were utilised: LSTM is also being used ANN, CNN, K-Nearest Neighbour and many other algorithms. The model's performance was evaluated using measures including F1-score(Fs), recall(Rc), accuracy(Acc), and precision(Pr). Outcomes showed that the models were not all equally successful; however, the LSTM model had the best accuracy at 93%. Future attempts might consider other categorisation strategies and improving preprocessing methods to improve model performance and forecast Acc*

**Keywords:** Stock Market, lemmatising, yahoo dataset, Machine Learning, deep learning

## I. INTRODUCTION

A stock market is a venue for the conversion, distribution, and sale of shares. It comes out that large organisations may use it as a valuable network to expand their assets from investors[1]. However, once the stocks are issued, a large number of money enters the stock market, increasing the structure of commercial assets by promoting cautious investing, which has a major positive impact on the expansion of the product economy[2]. On the other hand, growth in investment is effectively encouraged, and money is shared via the flow of stocks. As a result, it is commonly accepted that the stock market is a trustworthy gauge of the financial and economic health of a nation or region. In particular, buying and selling prices of stocks often function as an indicator of the quantity and value of stocks due to the correlation between supply and demand they reveal[3]. Since its inception, investors have endeavoured to forecast the stock market. On a daily basis, the exchange facilitates the trading of billions of dollars, with each dollar being backed by an investor with any hope of profit. Profitability and influence are alluring prospects for an investor who can effectively forecast market movements. Consequently, it is unsurprising that whenever the stock market misbehaves, the public is stirred to attention regarding the challenges it entails.


DOI: 10.48175/IJARSCT-22786

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, December 2024

Figure 1: Financial stock market

This share market is a collection of various individuals purchasing and selling shares. Commonly referred to as stock (stake), the term "generally" denotes ownership claims made by an individual or group of individuals regarding a business[4][5]. The process by which prospective stock market prices are valued is referred to as the "stock market estimate." Strong, accurate, and effective performance is anticipated. The functionality and integration of the system should be congruent with real-life situations. Typically, the sentiments of thousands of investors dictate the fluctuation of the stock market[6]. Getting historical stock market data is not hard, but the disadvantage of this approach is that important news events are often overlooked. It has been shown that a number of events may impact the stock price, including mergers and acquisitions, profit fluctuations, financial statement upgrades or downgrades, and changes to corporate processes or top executives. In order to analyse the vast quantities of data at their disposal, organisations are progressively placing greater reliance on fast computer processing. This is done to develop decision support systems that predict future trends and assist investors in making more informed choices. Price trend forecasting has been attributed to numerous systems that utilise financial news articles as their foundation [7].

The conduct of the market is subject to various assessments. Considerable attention has been recently devoted to subordinates, destinies, and alternatives, among other things. For risk management purposes, anticipating these subordinates is inconsequential. The prediction of financial exchange rates is a challenging expectation that is time-dependent. Market conduct is not entirely arbitrary, notwithstanding the complex and multidimensional character of market cost developments. Despite this, research using ML approaches for market prediction has proliferated in academic literature, and several successful technical assessments exist in the financial industry[8].

### A. Motivation and Contribution of this paper

This study aims to cope with the challenges of accurately forecasting stock market trends by leveraging machine learning techniques. Traditional methods often struggle with the complexity and unpredictability of financial markets, while human-labeled data is time-consuming and costly to obtain. This work aims to provide a scalable, efficient, and reliable approach to stock market prediction by utilising automated data collection and advanced preprocessing techniques. The study seeks to enhance decision-making for investors by comparing the effectiveness of different ML models, ensuring robust and practical solutions for financial forecasting. This study aims to compare multiple ML and DL approaches for the stock market system. The following summarises this work's main contribution:

- The study demonstrates the ability to forecast stock market trends without the need for human-labelled data by leveraging automated data collection and annotation methods from financial news articles, enhancing scalability and reducing manual intervention
- It introduces a robust preprocessing pipeline, including lemmatisation, handling missing and duplicate values, and feature selection, ensuring high-quality input data for machine learning models.
- Several machine learning models, including CNN, KNN, ANN, and LSTM, are assessed in the study, and performance metrics, including Acc, Pr, Fs and Rc, are used to compare each model's efficacy in depth.
- Lastly, assess each model and determine which one has the best accuracy performance.
- To improve stock market prediction, the work integrates conventional and contemporary machine learning techniques, offering insightful information on which algorithms are most suited for financial applications.

**B. Structure of paper**

The following sections make up the structure of this study: Section II examines the literature relevant to our inquiry. Methodological analysis of the study: In Section III, the research approach that was employed is presented. In Section IV, the results and evaluations of the research effort are discussed. Section V consists of the conclusions of our research study and our future plans.

## II. LITERATURE REVIEW

This section presents the relevant research in the domain of stock market prediction utilising various approaches and strategies.

In this research, Almusawi et al. (2023) forecast the market utilising the LSTM with the IARO optimisation algorithm. To maximise the precision of stock market price forecasts, the hyperparameters of the LSTM approach are optimised using the IARO algorithm. The suggested approach outperforms current approaches with lower values for MSE (0.43), MAE (0.37), MAPE (0.21), and R2 (0.17) [9].

In this study, Nikhil et al. (2023) examine four distinct models—LSTM, CNN, LSTM-CNN, and GA-LSTM-CNN—to predict stock prices. The dataset employed in this study consists of a specific company's daily stock prices for over a decade. Experimental results demonstrate that a GA-LSTM-CNN model outperformed the competition, with a training RMSE of 0.0064 and a testing RMSE of 0.03. The results show that stakeholders may benefit greatly from using the GA-LSTM-CNN model to forecast stock prices in the future[10].

In this research, Hotasi and Satya (2023) work lays out a two-step process: first, using a NB classifier to identify the bias in financial news; second, using a KNN classifier to forecast stock market movements based on the bias and historical stock price data. Prediction accuracy ranged from 92.3% to 96% during validation and from 91.9% to 97.3% during training and testing, all related to the first stage of this study's methodology—identifying the polarity of financial news. In the second stage, which included forecasting future stock market movements, the prediction accuracy ranged from 72.2% to 90% during testing and training and from 43.5% to 64.6% during validation[11].

This study presents, Ivanenko's (2023) study used an experimental design to compare two ARIMAX models: one that was baseline and one that was boosted with social media data. The prediction accuracy of the improved model was 65.7%, which is a considerable increase above the baseline model's accuracy rate of 60.1%. To further evaluate the usefulness of the suggested technique, future studies might include expanding the dataset and investigating additional social media sites[12].

In this research, Zouaghia, Kodia Aouina and Ben Said (2023) The authors provide a system that uses five ML classifiers—GNB, RF, GB, SVM, and kNN—to forecast the trends in closing prices. Using stock price history as input, technical indicators are computed and applied. The Quotations (NASDAQ) stock data used in these classifiers spans the years 2018 through 2023. The RFC model produced the greatest results, with an accuracy of 61%[13].

Table I offers a comparative analysis of the financial stock market through the use of multiple methodologies.

Table 1: Comparative study on the financial stock market using various techniques

| Author | Methodology | Key Findings | Performance | Limitations & Future Work |
|---|---|---|---|---|
| Almusawi et al., | LSTM with IARO algorithm | Superior accuracy compared to existing methods. | MSE: 0.43, MAE: 0.37, MAPE: 0.21, R2: 0.17 | Lack of comparison with other optimisation algorithms; Further exploration of the IARO algorithm's applicability to other domains. |
| Nikhil et al., | LSTM, CNN, LSTM-CNN, GA-LSTM-CNN | GA-LSTM-CNN outperforms other models. | Training RMSE: 0.0064, Testing RMSE: 0.03 | Limited explanation of how the genetic algorithm assists in hyperparameter optimisation and generalisation to different datasets. |

| | | | | |
|---|---|---|---|---|
| Hotasi and Satya, | Naive Bayes classifier, KNN classifier | High prediction accuracy in news polarity determination. Moderate accuracy in stock trend prediction. | Training & Testing Accuracy: 72.2% - 90%; Validation Accuracy: 43.5% - 64.6% | Lack of consideration for other sentiment analysis techniques; Further validation on diverse datasets. |
| Ivanenko, | Zero-Shot Learning, ARIMAX model | The enhanced model outperforms the baseline with social media data. | Accuracy: Enhanced model: 65.7%, Baseline model: 60.1% | Limited scope on social media platforms; Expansion of dataset and methodologies. |
| Zouaghia, Kodia Aouina and Ben Said, | ML classifiers (GNB, RF, GB, SVM, kNN) with PCA and Grid Search | Random Forest achieves the highest accuracy. | Accuracy: 61% | Limited exploration of feature selection techniques; Incorporation of additional classifiers and optimisation methods. |

### A. Research gap

Although there are a number of distinct methods available for stock market prediction, good stock market forecasts still need to address certain limitations. This section discusses research gaps and problems in several methodologies used to forecast stock market performance. A research gap in the area of stock market prediction is the lack of integration between sentiment analysis from financial news and social media data and classic quantitative models. Previous research has investigated the efficacy of ML algorithms and advanced DL methods like LSTM and CNN in predicting stock prices. Few research, nonetheless, have looked at how sentiment analysis of social media and financial news could increase the precision of these forecasts. Moreover, integrating qualitative sentiment research with quantitative models has the potential to provide a comprehensive approach to comprehending market dynamics and enhance the Acc of stock market forecasts. Therefore, to improve the predictive capabilities of stock market prediction models, methods that combine sentiment research from several sources with advanced machine learning techniques must be investigated.

## III. METHODOLOGY

This section elaborates the work's methodology and is parted into various subsections.

### A. Methodology

This section provides coverage of stock market forecasts. An important advantage of this method is the capacity to annotate an unlimited amount of data without the need for any human marking. The first stage of this procedure involves gathering data from a Yahoo dataset, which comprises URLs for financial news stories. Subsequently, financial news items obtained by web scraping are used. Figure 2 shows all the phases of methodology also each step discussed.

The study acquired data through the use of the finance.yahoo.com website. Next were the preprocessing stages like lemmatisation, taking care of missing and duplicate values, and feature selection. Data was splitted into training and evaluation sets using a 70:30 ratio to prevent overfitting. The performance of ML prediction models, like CNN, KNN, ANN, and LSTM, was evaluated using Acc, Pr, Rc, and F measures as a comparison measure. The analysis preceded with data processing using both classic as well as modern ML prediction techniques in order to provide the report with relevant models for stock market forecasting.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, December 2024

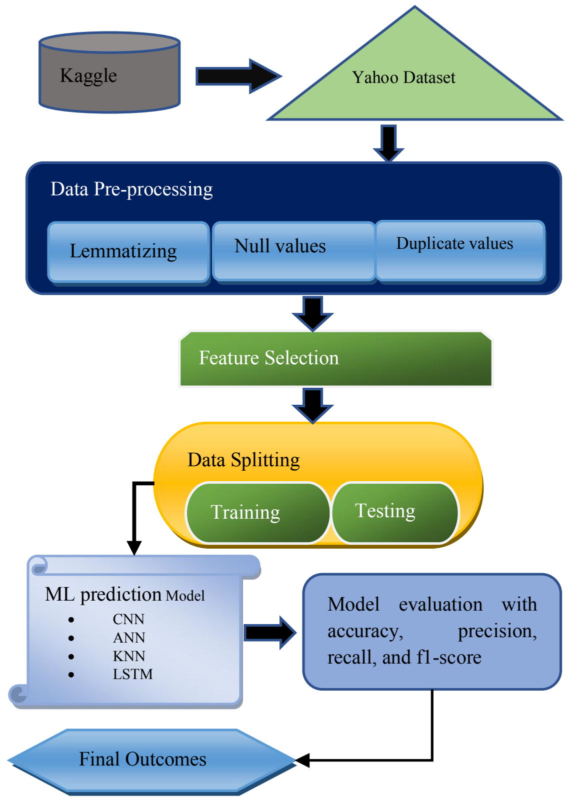


Figure 2: Methodology flowchart for finance stock prediction

The process of methodology is briefly described in the section below.

**1. Data collection**

The goal of data collection is to identify trends and potential solutions to research issues by amassing information from a variety of sources. The Yahoo dataset used in this comparison research was collected from finance. Yahoo. Com, which was found useful for the analysis.

**2. Data preprocessing**

Data processing starts with translating the raw data into a format that can be beneficial for the purpose. Processes include N-gram tokenising, lemmatisation and cleaning the table of unusable columns to ensure accuracy and workability with other data. Meanwhile, this stage is vital for data readying before its use in research for modelling or analysis is deemed appropriate. The following preprocessing steps are as follows:

- **Lemmatising:** The lemmatisation process is when the stem and inflectional variations of the word are reduced into the dictionary form of such word or its base.
- **Null values:** Data preparation cannot be complete without handling null values. Imputation and removal methods are employed for correct data removal. This will guarantee the credibility and quality of the dataset for the following analysis and modelling activities in research processes.
- **Duplicate values:** Detecting and removing redundant figures is a necessary resolution to the entire data preparation procedure. The deduplication techniques are utilised to confirm the uniqueness of each entry, thereby increasing the reliability of the information.

**Feature selection:** There is substantial evidence that feature selection is a fast and effective solution to ML challenges. Improving data mining performance, including predicted accuracy and readability, is one of the goals of feature selection. Another is to construct models that are simpler and easier to understand. Preparing data that is easy to interpret also involves removing irrelevant or redundant information

**3. Data Splitting**

By using the ML approach for data splitting, overfitting may be prevented. A ratio of 70:30 has been calculated using the provided data. They devote 70% of our time to training and 30% to testing.

**4. ML prediction techniques**

ML prediction is an important technique that distinguishes a data point into multiple classes. A comparative algorithm is described in below:

**a. Convolution Neural Network (CNN)**

The CNN model is a type of DL for working with grid-based data, like pictures. It is based on how the visual brain of animals is organised[14] and It is It is intended to learn spatial information ensembles automatically and adaptively, moving from low-level to high-level patterns. Convolutional, pooling, and fully connected layers are used in the mathematical construction of CNNs. A fully connected layer transfers the collected features into the final output, such as ML prediction, whereas the previous two, convolution and pooling layers, execute feature extraction.

**b. K- Nearest Neighbor (KNN)**

Regarding machine learning algorithms, the KNN algorithm is among the most basic: It finds the k "nearest neighbours" of Y and uses that modality to label the profile or feature vector xi. The value of k indicates the number of nearby profiles that should be considered. Assigning the category y of the nearest neighbour to xi when k= 1 and the most frequent category among its nearest neighbours when k> 1 are both done by the method[15].

**c. Artificial Neural Network (ANN)**

Analogous to biological neurons in that it emulates the dendrite, soma, and axon processes, the algorithm for ANN is marginally inspired by biological neurons. Artificial neurons and elementary mathematical operations constitute the inner structure of each ANN. ANNs, which are employed to resolve intricate classification and regression challenges, represent a more advanced iteration of perceptrons. The forward propagation and prediction of a solitary neuron is represented as follows Eq.1:

$$output = b_1 + \sum_{j=1}^{n_x} W_{ij} x_i \qquad (1)$$

Long short-term memory (LSTM)

An RNN subtype with the ability to learn long-term dependencies is the LSTM network. LSTMs are extensively employed and have shown great efficacy in a variety of issues. The network of LSTM is comprised of numerous units of LSTM, each of which is linear and weighted by a self-connection [16]. This is all made possible by a set of gates. There are numerous parts to the LSTM device. The components are made up of cells, input gates, output gates, and forget gates, and they range in size from 2 to 6. The three gates may be seen as regular artificial neurones, similar to those found in a feedforward or multi-layer neural network, and the cellular design is responsible for long-term memory retention. Activation is easily calculated using a weighted sum [15] [17].

**Forget gate:**

$$f_t = \sigma\left(W_f . \left[h_{t-1,} x_t\right] + b_f\right) \qquad (2)$$

**Input gate:**

$$i_t = \sigma\left(W_i . \left[h_{t-1,} x_t\right] + b_i\right) \ (3)$$

**Cell state:**

$$\tilde{C}_t = tanh\left(W_c . \left[h_{t-1,} x_t\right] + b_c\right) \qquad (4)$$

**Output gate:**

$$o_t = \sigma\left(W_o \left[h_{t-1,} x_t\right] + b_c\right) \qquad (5)$$

**Hidden state:**

$$h_t = o_t * \tanh\left(C_t\right) \qquad (6)$$

σ is a sigmoid function, 'W' is a weight variable, and 'b' is a bias variable; the expression [h(t-1), xt] is the sum of the previous hidden state and the current input. A dot product is represented by '.' and element-wise multiplication by '*'.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, December 2024

## IV. RESULTS & DISCUSSIONS

This section presents the outcomes obtained through the models used in this study, including the results, performance metrics, and classifier statistics.

### A. Performance Measures

Among the performance metrics that have been employed include F1-score, recall, accuracy, etc.

#### 1. Confusion Matrix

The confusion matrice shows the number of predicted against actual values being occurred. In statistics, "TN" stands for "True Negative," a measure that displays the proportion of really negative instances that were properly anticipated [18]. True Positives, abbreviated as "TP," indicate the total number of positively identified instances and are conceptually comparable. The False Positive value (abbreviated "FP") is the number of correct negatives that are mistakenly labelled as positive. Conversely, "FN" displays the FN value, which is the quantity of positive cases that were incorrectly classified as negative situations[19].

#### 2. Accuracy

The attribute "val_accuracy" is derived from evaluating the precision of the model where it has been tested on a validation set at the conclusion of every training period. Also, the "accuracy" parameter refers to the accuracy percentage of a prediction on a training set, which is divided once again into random parts after any training period. It is computed as follows in Eq.7:

$$Acuracy = \frac{TP+TN}{TP+TN+FN+FP} \quad (7)$$

#### 3. Precision

Precision defines how precise the model is in making correct predictions. It is a parameter which shows how well the model performs at recognising positive samples, and It is the proportion of TP to TP and FP combined. It offers insight into the percentage of projected positive instances that are genuine positives[20]. The calculation is as follows as Eq.8:

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

#### 4. Recall

The model's capability to capture all truly positive samples is measured by its recall. It determines the ratio of TP to the total number of TP samples and is calculated using Equation 9, which includes both TP and FN:

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

#### 5. F1-score

As a composite measure, the F1-score incorporates both recall and precision into a single numerical result. The evaluation it gives is well-rounded, and it's especially helpful when there's a trade-off between recall and precision[21]. It's computed as Eq.10:

$$F1 - score = 2 * \frac{precision*recall}{precision+recall} \quad (10)$$

Here, provides the comparative analysis of the models according to Fs, Rc, Acc, and Pr.

### B. Experiment results

This section provides the experiment outcomes of a model utilised in this study for stock market prediction.

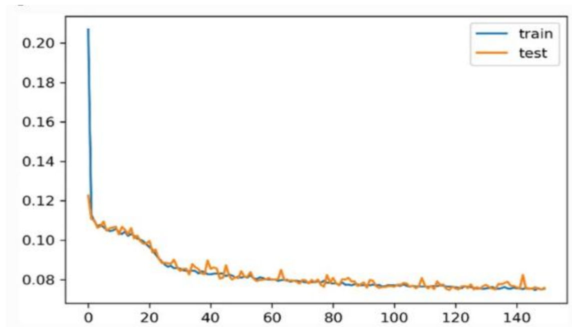


Figure 3: Learning curve of LSTM model

Learning track of LSTM model over epochs is depicted in Figure 3. The evaluation loss is shown by the yellow line in this graph, while the training loss is shown by the blue line. While the testing loss lags little, the training loss drops significantly. The LSTM model decreased the loss for financial stock market forecast by 0.0702%.

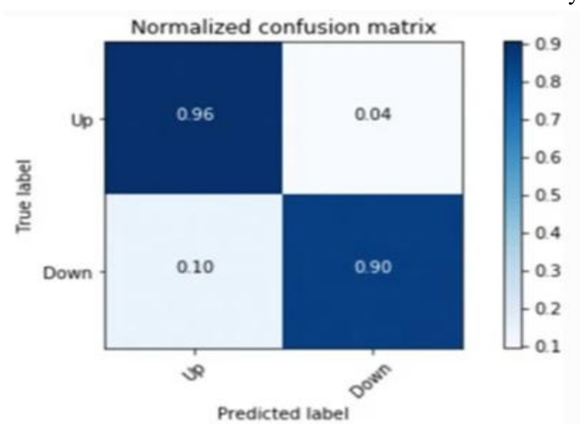


Figure 4: Confusion matrix of the LSTM model

Figure 4 displays a confusion matrix of the LSTM model, which contains two classes: predicted label and true label. In the down class, the values are 0.96 and 0.90, and in the up class, 0.10 and 0.04.

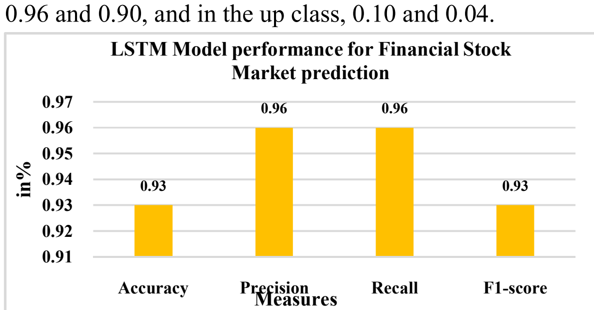


Figure 5: Performance of the LSTM model for financial stock market prediction

The LSTM model's performance on the Yahoo dataset for financial stock market prediction demonstrates high effectiveness across key evaluation metrics, as shown in Figure 5. With an accuracy of 93%, the model shows strong reliability in predicting stock price movements. Additionally, the precision and recall, at 96%, highlight the model's capability to consistently make correct positive predictions and identify true positives effectively. At 93%, the F1-score confirms the model's overall robustness by reflecting a balanced trade-off between recall and accuracy. According to these findings, the LSTM model guarantees good accuracy and dependability when used for financial stock market prediction jobs, including time series forecasting.

**C. Comparative analysis**

This section compares the ML prediction models using Rc, Acc, Pr, and Fs as performance metrics; the results comparison is shown the Table II.

Table 2: Comparative results of model performance

| Models | Acc | Pr | Rc | Fs |
|---|---|---|---|---|
| CNN[22] | 63.0 | 81.6 | 63.1 | 69.7 |
| KNN[23] | 79.15 | 79.50 | 79.50 | 79.12 |
| ANN[23] | 84.45 | 84.50 | 84.50 | 84.50 |
| LSTM | 0.93 | 0.96 | 0.96 | 0.93 |

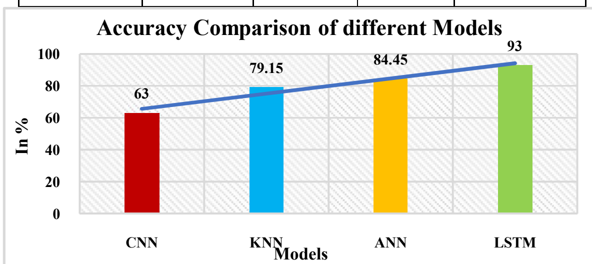


Figure 6: Bar graph of accuracy comparison of models

Figure 6 represents the Accuracy bar graph, which shows the accuracy of different models where LSTM gets the highest score; LSTM achieves 93% accuracy, the highest in all models.

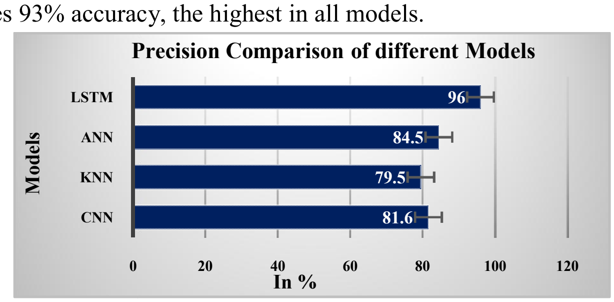


Figure 7: Bar graph of the precision comparison of models

Figure 7 represents the Precision bar graph of models where LSTM gets the highest score of all models; LSTM achieves a 96% precision score, which is the highest in all models.

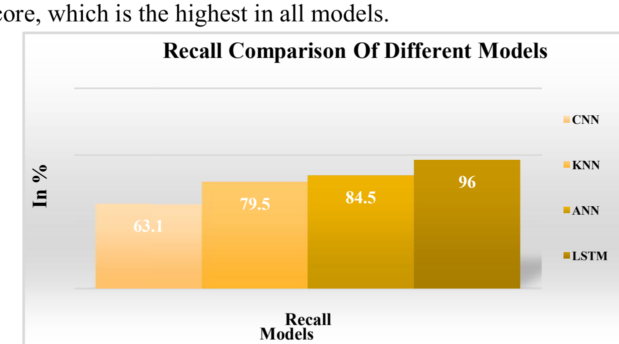


Figure 8: Bar graph of recall comparison of models

Figure 8 represents the recall bar graph, which shows the recall of various models where LSTM gets the highest score of all models; LSTM achieves a 96% recall score, the highest in all models.

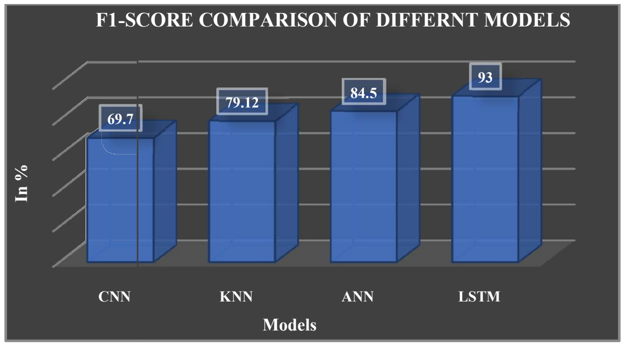


Figure 9: Bar graph of f1-score comparison of models

Figure 9 represents the f1-score bar graph, which shows the f1-score of various models where LSTM gets the highest score; LSTM achieves 93% of the f1-score, which is the highest in all models.

**D. Discussion**

The comparative analysis of ML models for stock market prediction demonstrates that the LSTM model outperforms CNN, KNN, and ANN across all evaluation metrics, achieving the highest accuracy (93%), precision (96%), recall (96%), and F1-score (93%). These results show that because LSTM can precisely capture temporal dependency in time series data, it is a good fit for stock market forecasting. In contrast, ANN and KNN show moderate performance, with ANN achieving an accuracy of 84.45% and KNN at 79.15%. CNN, however, performs relatively poorly, with an accuracy of 63%, indicating its limited suitability for sequential data tasks. While LSTM excels in predictive performance, its computational complexity and longer training time compared to other models may pose limitations in resource-constrained environments. Conversely, simpler models like KNN and ANN are less resource-intensive but fail to match LSTM's prediction accuracy and robustness. Thus, LSTM is advantageous for applications requiring high precision and reliability, albeit at the cost of increased computational demand.

## V. CONCLUSION AND FUTURE WORK

Recent developments in processing power and information management have made it simpler to predict the actions of the stock market. Utilising cutting-edge deep learning techniques, patterns within the data may be examined and discovered. The best prediction models from recent times incorporate two kinds of inputs: textual data such as news headlines and articles. The findings of this study demonstrated how critical it is to use several categorisation models in addition to careful data preparation in order to make accurate predictions. The study compared CNN, KNN, ANN, and LSTM models to help understand their 93% performance across many criteria. Although every model has its advantages and disadvantages, the LSTM model showed the most promise for financial analysis forecasting jobs due to its high accuracy of 93% compared to other models.

Nevertheless, there are several limitations to this research. The models ' performance could differ depending on the dataset's distinctive characteristics and the parameters used. Several potential directions might be pursued by future research in this field. First, new insights into the efficacy of the models might be obtained by investigating other categorisation strategies than the ones used in this work. Additionally, there is an exceptional opportunity to create a holistic prediction system that is trained to utilise a variety of information kinds, including tweets, news, and other text-based data, by fusing the most recent sentiment analysis approaches with feature engineering and deep learning models.

## REFERENCES

[1] R. Tandon, “Prediction of Stock Market Trends Based on Large Language Models,” *J. Emerg. Technol. Innov. Res.*, vol. 11, no. 9, pp. a615–a622, 2024.

[2] H. Sinha, “Benchmarking Predictive Performance Of Machine Learning Approaches For Accurate Prediction Of Boston House Prices : An In-Depth Analysis,” *ternational J. Res. Anal. Rev.*, vol. 11, no. 3, 2024.

[3] M. Ali, D. M. Khan, H. M. Alshanbari, and A. A. A. H. El-Bagoury, “Prediction of Complex Stock Market Data Using an Improved Hybrid EMD-LSTM Model,” *Appl. Sci.*, 2023, doi: 10.3390/app13031429.

[4] S. Bauskar, "BUSINESS ANALYTICS IN ENTERPRISE SYSTEM BASED ON APPLICATION OF ARTIFICIAL INTELLIGENCE," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 04, no. 01, pp. 1861–1870, 2022, doi: DOI : https://www.doi.org/10.56726/IRJMETS18127.

[5] H. Sinha, "Predicting Bitcoin Prices Using Machine Learning Techniques With Historical Data," *Int. J. Creat. Res. Thoughts*, vol. 12, no. 8, 2024, doi: 10.3390/e25050777.

[6] A. Patel, D. Patel, and S. Yadav, "Prediction of Stock Market Using Artificial Intelligence," *SSRN Electron. J.*, 2021, doi: 10.2139/ssrn.3871022.

[7] P. Y. Hao, C. F. Kung, C. Y. Chang, and J. B. Ou, "Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane," *Appl. Soft Comput.*, 2021, doi: 10.1016/j.asoc.2020.106806.

[8] P. Modi, "Big Data Analysis in Stock Market Prediction," vol. 8, no. 10, pp. 384–386, 2019.

[9] M. Almusawi, S. Chakraborty, G. Ravindran, S. Prabu, and Z. A. Almoussawi, "Long Short-Term Memory optimised with Improved Artificial Rabbits optimisation Algorithm for Stock Market Price Prediction," in *International Conference on Integrated Intelligence and Communication Systems, ICIICS 2023*, 2023. doi: 10.1109/ICIICS59993.2023.10420974.

[10] S. Nikhil, R. K. Sah, S. Kumar Parki, T. B. Tamang, D. Somashekhara Reddy, and T. R. Mahesh, "Stock Market Prediction Using Genetic Algorithm Assisted LSTM-CNN Hybrid Model," in *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*, 2023. doi: 10.1109/ICCCNT56998.2023.10306948.

[11] A. Hotasi and D. P. Satya, "Predicting Stock Market Trends with News Sentiment Analysis and Historical Stock Market Data Analysis Using Data Mining Techniques," in *2023 10th International Conference on Advanced Informatics: Concept, Theory and Application, ICAICTA 2023*, 2023. doi: 10.1109/ICAICTA59291.2023.10390037.

[12] V. Ivanenko, "Assessing Social Media's Impact on Stock Market Predictions: Financial Sentiment Embedding Approaches," in *2023 13th International Conference on Dependable Systems, Services and Technologies, DESSERT 2023*, 2023. doi: 10.1109/DESSERT61349.2023.10416463.

[13] Z. Zouaghia, Z. Kodia Aouina, and L. Ben Said, "Stock Movement Prediction Based on Technical Indicators Applying Hybrid Machine Learning Models," in *2023 International Symposium on Networks, Computers and Communications, ISNCC 2023*, 2023. doi: 10.1109/ISNCC58260.2023.10323971.

[14] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol.*, 1968, doi: 10.1113/jphysiol.1968.sp008455.

[15] P. Khare and S. Srivastava, "AI-Powered Fraud Prevention: A Comprehensive Analysis of Machine Learning Applications in Online Transactions," *J. Emerg. Technol. Innov. Res.*, vol. 10, pp. f518–f525, 2023.

[16] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[17] J. Shen and M. O. Shafiq, "Short-term stock market price trend prediction using a comprehensive deep learning system," *J. Big Data*, 2020, doi: 10.1186/s40537-020-00333-6.

[18] A. Kulkarni, D. Chong, and F. A. Batarseh, "5 - Foundations of data imbalance and solutions for a data democracy," in *Data Democracy*, F. A. Batarseh and R. Yang, Eds., Academic Press, 2020, pp. 83–106. doi: https://doi.org/10.1016/B978-0-12-818366-3.00005-8.

[19] K. Ullah *et al.*, "Short-Term Load Forecasting: A Comprehensive Review and Simulation Study with CNN-LSTM Hybrids Approach," *IEEE Access*, vol. 12, no. July, pp. 111858–111881, 2024, doi: 10.1109/ACCESS.2024.3440631.

[20] R. Tandon, "The Machine Learning Based Regression Models Analysis For House Price Prediction," *Int. J. Res. Anal. Rev.*, vol. 11, no. 3, pp. 296–305, 2024.

[21] T. Gulsoy and E. B. Kablan, "FocalNeXt: A ConvNeXt augmented FocalNet architecture for Lung Cancer Classification from CT-scan images," 2024.

[22] D. Nascimento, A. Costa, and R. Bianchi, "Stock Trading Classifier with Multichannel Convolutional Neural Network," 2020. doi: 10.5753/eniac.2020.12136.

[23] A. H. Khan *et al.*, "A performance comparison of machine learning models for stock market prediction with novel investment strategy," *PLoS One*, 2023, doi: 10.1371/journal.pone.0286362.