International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 16 (2024) pp. 293–308

Article

Predicting Stock Market Prices and Provide Recommendations

Krishna Arora [1,*], Akshima Aggarwal [2] and Kamal Kumar Gola [3]

1 Department of Information Technology, College of Engineering Roorkee, Roorkee 247667, Uttarakhand, India
2 Department of Computer Science and Engineering, Indian Institute of Information Technology, Una 177209, Himachal Pradesh, India; aggarwalakshima@gmail.com
3 Department of Computer Science and Engineering, College of Smart Computing, COER University, Roorkee 247667, Uttarakhand, India; kkgolaa1503@gmail.com
* Correspondence: krishnaarora747@gmail.com

Received date: 1 March 2024; Accepted date: 21 March 2024; Published online: 10 July 2024

Abstract: The project "Stock Market Price Prediction and Recommendation System" endeavors to address the intricate challenge of accurately predicting stock prices in the dynamic landscape of the global market. Utilizing a combination of traditional analysis techniques and advanced machine learning algorithms, the project aims to develop predictive models that can forecast stock price movements with heightened accuracy. Python programming emerges as a vital tool in facilitating data analysis, model development, and evaluation. Through a comprehensive exploration of historical data, incorporation of alternative data sources, and optimization of hyperparameters, the project seeks to enhance the predictive capabilities of stock market analysis. Furthermore, ethical considerations surrounding algorithmic trading are carefully examined, emphasizing transparency, fairness, and accountability in decision-making processes. Ultimately, the project aspires to empower investors with actionable insights and recommendations, enabling informed decision-making in the complex realm of stock market investments.

Keywords: stock market; price prediction; recommendation system; Python programming; machine learning; data analysis; ethical considerations

1. Introduction

In the sprawling landscape of the global economy, the stock market stands as a towering colossus, its movements dictating the fortunes of nations and individuals alike. Within its labyrinthine corridors, investors tread a precarious path, seeking to decipher the enigmatic dance of supply and demand that shapes the ebb and flow of stock prices. Yet, for all their efforts, the task of accurately predicting these movements remains an elusive pursuit, fraught with uncertainty and complexity [1]. It is within this crucible of market dynamics that the project "Stock Market Price Prediction and Recommendation System" takes root, driven by a relentless quest to unlock the secrets of the stock market and empower investors with the tools they need to navigate its treacherous waters [2]. This challenge, fraught with complexities and uncertainties, forms the backdrop against which the project "Stock Market Price Prediction and Recommendation System" [3] unfolds, driven by an unwavering commitment to unraveling the mysteries of market dynamics and empowering investors with actionable insights [4]. In the grand tapestry of global commerce, the stock market emerges as a pulsating nexus, where capital flows, dreams are realized, and fortunes are made. It serves as the epicenter of economic activity, a realm where investors, both seasoned professionals and eager novices, converge in pursuit of prosperity [5]. Yet, amidst the bustling chaos of buy and sell orders, lies a fundamental challenge: the enigma of predicting stock price movements with precision and foresight. This challenge, fraught with complexities and uncertainties, forms the backdrop against which the project "Stock Market Price Prediction and Recommendation System" [6] unfolds, driven by an unwavering commitment to unraveling the mysteries of market dynamics and empowering investors with actionable insights [7]. Its core, the project is driven by a profound recognition of the multifaceted nature of the stock market. Far from being a monolithic entity, the market is a complex ecosystem shaped by a myriad of factors, both tangible and intangible. In this vast and intricate landscape, the ability to discern meaningful patterns from the cacophony of data is a skill coveted by investors and analysts alike [8]. It is within this context that the project sets its sights, seeking to harness the power of advanced analytics and machine learning to distill actionable intelligence from the deluge of market information [9]. At the heart of the project lies a recognition of the transformative potential of technology, particularly Python programming, in revolutionizing the field of stock market analysis.

1.1. Technology to Expore Dataset

Python's versatility and robust ecosystem of libraries make it the tool of choice for data scientists and financial analysts seeking to extract insights from complex datasets [10]. From data wrangling and preprocessing to model development and evaluation, Python streamlines the analytical process, allowing researchers to focus their efforts on extracting meaningful signals from the noise of market data [11]. Central to the project's mission is the integration of traditional analysis techniques with cutting-edge machine-learning algorithms, all powered by the versatile capabilities of Python programming. In this symbiotic fusion of art and science, historical data is transformed into predictive models that seek to illuminate the future trajectory of stock prices [12]. By harnessing the power of Python's rich ecosystem of libraries and tools, researchers can explore vast troves of data, develop sophisticated models, and deploy them with ease, paving the way for a new era of data-driven investing.

But the project is not merely an exercise in technical proficiency; it is a quest for deeper understanding and insight into the inner workings of the stock market [13]. It is a journey of exploration and discovery, guided by a series of research questions aimed at uncovering the underlying patterns and dynamics driving stock price movements. From the role of Python in facilitating technical and fundamental analysis to the optimization of machine learning algorithms for predictive modeling [14], the project seeks to push the boundaries of knowledge and innovation in the field of quantitative finance. However, the project's ambitions extend beyond mere prediction; it aspires to empower investors with actionable insights and recommendations, guiding them towards informed decision-making in an uncertain world [15]. Through the lens of machine learning, patterns emerge, trends reveal themselves, and opportunities beckon to those with the foresight to seize them. Yet, amidst the promise of profit lies a responsibility: to navigate the ethical and moral considerations that accompany algorithmic trading. It is a delicate balance, one that the project endeavors to strike through rigorous research and transparent practices [16]. As the project unfolds, it embarks on a journey of discovery, probing the depths of the stock market's complexities and uncovering the hidden truths that lie beneath the surface. Research questions abound, from the role of Python in facilitating analysis to the optimization of hyperparameters, from the exploration of alternative data sources to the ethical implications of algorithmic decision-making [17]. Through rigorous inquiry and collaborative effort, the project seeks not only to advance the frontiers of stock market prediction but also to empower investors with the knowledge and tools they need to navigate this ever-evolving landscape with confidence and clarity. Moreover, Ref. [18] the project is not conducted in isolation; it is situated within the broader context of ethical considerations and societal impacts. As algorithms increasingly shape decision-making processes in the financial markets, questions of transparency, fairness, and accountability loom large. The project thus grapples with these issues, striving to develop methodologies and practices that uphold the highest standards of integrity and responsibility [19].

1.2. Intricacies of the Stock Market

The stock market serves as a vital component of the global economy, offering individuals and institutions the opportunity to invest in a wide range of financial instruments. In this literature review, we delve into the multifaceted realm of stock market analysis and recommendation, Ref. [20] exploring the methodologies, approaches, and emerging trends that guide decision-making in this intricate landscape [21]. The synthesis of historical data, predictive modeling, and real-time information sources has led to innovative strategies for enhancing stock market analysis and assisting investors in making informed choices [22]. The stock market serves as a cornerstone of the global economy, exerting profound influence over the financial landscape and shaping the fortunes of investors worldwide [23]. Predicting stock market movements accurately is a perpetual challenge, necessitating the development of robust analytical tools and methodologies. For a considerable amount of time, traditional methods of stock market analysis—like fundamental and technical analysis—have been used to interpret market dynamics and predict price changes [24]. By closely examining a company's financial statements, industry trends, and macroeconomic factors, fundamental analysis assesses a company's competitive position, growth potential, and financial health [25].

Technical analysis, on the other hand, is centered on past market data and examines patterns and trends in trading volumes and stock prices to predict future movements. However conventional approaches [26] have their limits, especially when it comes to understanding the complex interactions between economic indicators, geopolitical developments, and investor emotions that affect stock prices. To improve the accuracy of stock market predictions, researchers have consequently been using machine learning (ML) and artificial intelligence (AI) approaches more and more [27]. ML algorithms, such as decision trees, support vector machines (SVM), and random forests, leverage historical data to identify patterns and relationships, offering valuable insights into market behavior [28]. Convolutional neural networks (CNN) and recurrent neural networks (RNN), two examples of deep learning models, have become extremely effective tools for evaluating sequential data and identifying temporal correlations in financial time series [29]. These models are excellent at deciphering intricate patterns and handling enormous volumes of data, which makes them suitable for tasks involving stock market prediction [30]. CNN-LSTM models are an example of a hybrid model that combines several machine learning and deep learning approaches [31]. These models perform better than individual algorithms and provide more stable and accurate predictions.

Overall, the integration in Figure 1, regarding the generic scheme of Machine Learning Techniques, deep learning, and sentiment analysis techniques represents a significant advancement in stock market prediction [32]. These innovative approaches offer investors valuable insights into market trends and dynamics, empowering them to make more informed decisions and navigate the complexities of the stock market with greater confidence [33].

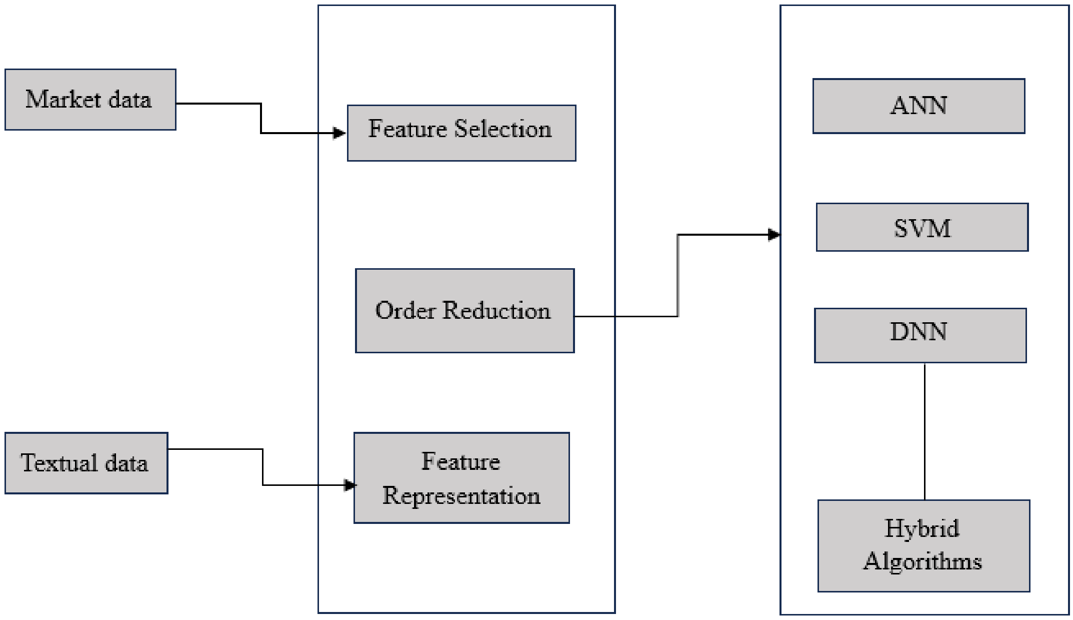

Figure 1. Generic Scheme for Evaluation.

1.3. Historical Trends and Traditional Analysis Methods

Over the years, analysts and researchers have sought to develop tools and models that can decipher the intricate dynamics of the stock market, ultimately leading to more accurate predictions. Technical analysis involves the study of historical price charts and patterns to forecast future price movements, leveraging tools such as moving averages, relative strength indicators, and candlestick patterns by using different pre-processing methods [34]. Fundamental analysis, on the other hand, delves into a company's financial health, macroeconomic factors, and industry trends to uncover the intrinsic value of a stock. While these traditional methods have been instrumental in guiding investment decisions, they often fall short of capturing the complexity of market dynamics.

1.4. Machine Learning and Predictive Modeling

The potential of machine learning approaches to improve stock market analysis and prediction has attracted a lot of attention in recent years. Through the use of machine learning techniques, researchers have been able to evaluate historical stock data and produce precise predictions. This is especially true of recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks. LSTM networks, in particular, have demonstrated exceptional capabilities in sequence prediction problems due to their memory retention properties. By capturing sequential patterns inherent in historical stock price data, LSTM networks offer valuable insights into future market movements.

In Figure 2, it is one kind of RNN (recurrent neural network) that can handle and analyze sequential input is an LSTM (Long Short-Term Memory) network [35]. An LSTM network is made up of a sequence of LSTM cells, each of which has a set of input, output, and forget gates to regulate the information entering and leaving the cell. The LSTM can preserve long-term dependencies in the input data by using the gates to forget or keep information from earlier time steps selectively. Additionally, the LSTM cell has a memory cell that uses data from earlier time steps to affect the cell's output at the present step. The following cell in the network receives the output from each LSTM cell, which enables the LSTM to process and interpret consecutive data over many time steps. An input gate, an output gate, a forget gate, and a cell make up an LSTM unit [36]. The three gates control the information flow into and out of the cell, and the cell retains values for arbitrarily long periods. Forget gates use a value between 0 and 1 to indicate which information from a prior state should be discarded about the present input. To retain the information, a (rounded) value of 1 is indicated, and to discard it, a value of 0. Using the same mechanism as forget gates, input gates determine which new pieces of information to store in the existing state [36, 37]. To retain valuable long-term dependencies for prediction-making in both present and future time steps, the LSTM network selectively outputs pertinent information from the current state.

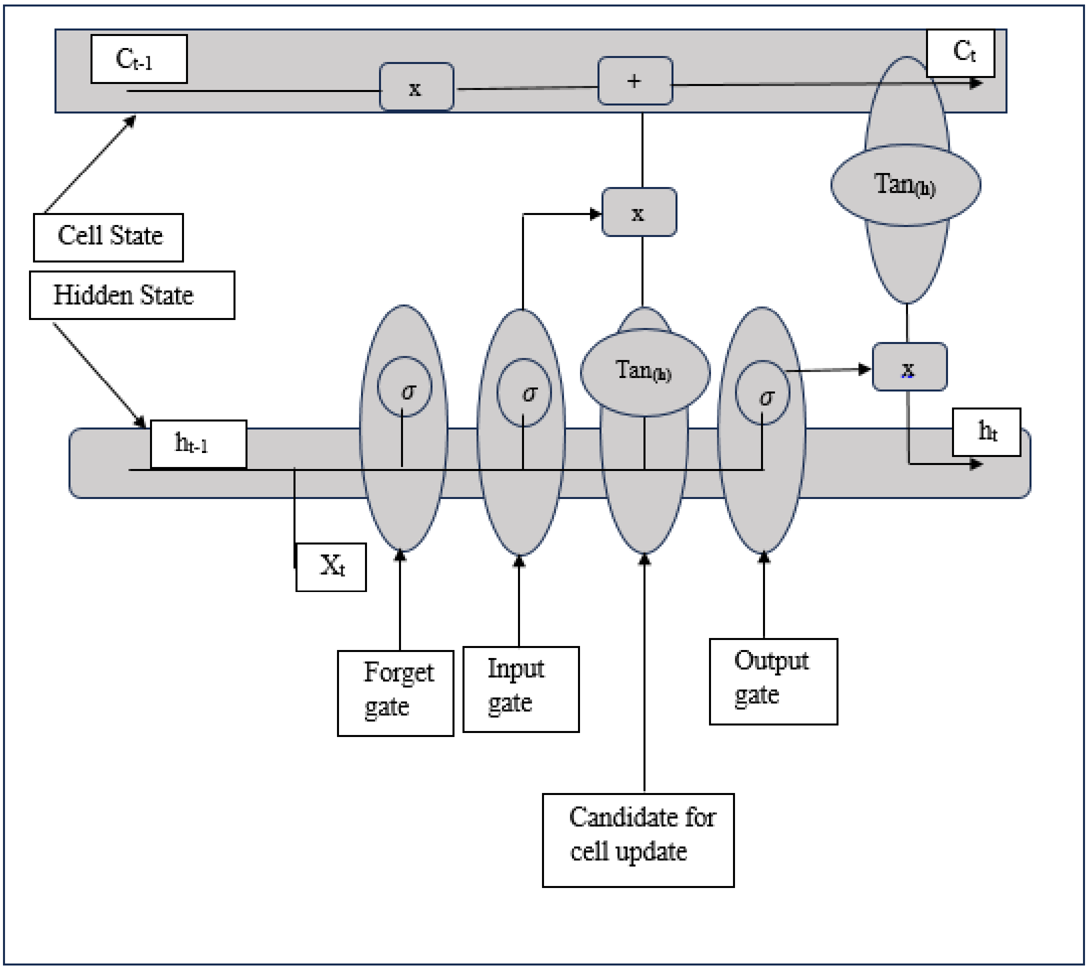

Figure 2. LSTM (Long Short-Term Memory) Cell.

LSTM is made up of Gates:
In LSTM we will have 3 gates:

1) Input Gate.
2) Forget Gate.
3) Output Gate

1.5. Need for Accurate Stock Market Forecasting

Forecasting the stock market accurately is crucial for investors looking to minimize risks and maximize returns. Investors can use cutting-edge tools and software to watch market movements and make wise judgments with the aid of machine learning techniques. Market dynamics are greatly impacted by variables such as projected future income, earnings releases, and management changes [37]. As a result, accurate stock market trend forecasting can enable investors to take advantage of new possibilities and make wiser selections.

1.5.1. Impact of Stock Market on Economic Growth

A major factor in the economic growth of developing nations like India is the stock market. On the other hand, stock market volatility may have unfavorable consequences that hinder economic expansion [38]. Thus, to reduce risks and promote economic stability, authorities and investors must comprehend and correctly forecast stock market developments.

1.5.2. Role of Social Networking in Stock Market Prediction

Social media sites such as Facebook, Google+, and Twitter are excellent resources for obtaining popular thoughts and attitudes on a range of subjects, including changes in the stock market [38]. These platforms offer public opinion insights that can be used to forecast movements in stock prices.

1.5.3. Impact of Financial News on Stock Market Prediction

Because they offer insights into corporate performance, market trends, and economic considerations, financial news items have a big influence on stock market prediction. Forecasting changes in stock prices and directing trading tactics can be aided by analyzing news stories about particular industries.

1.5.4. Role of Gross Domestic Product (GDP) in Stock Market Movements

GDP is a measure of economic progress that affects consumer spending power and business earnings. GDP changes affect sales turnover and the manufacturing of consumer products, which in turn affects stock prices.

1.5.5. Impact of Market on Stock Price Determinants

The basic concepts of supply and demand in economics have an impact on stock prices. A stock's price rises in response to strong demand, and vice versa. Furthermore, several macroeconomic variables, including exchange rates, GDP, inflation, and oil prices have a big impact on stock price swings.

2. Evaluation of the Proposed Model

The proposed model is evaluated for accuracy and effectiveness in predicting stock price movements. Through extensive experimentation and analysis, the model demonstrates its ability to produce accurate predictions and outperform traditional machine-learning approaches.

The future of stock market analysis [39, 40] holds promising prospects as artificial intelligence and machine learning algorithms continue to mature. The integration of real-time sentiment analysis, big data analytics, and ethical considerations will shape the landscape of stock market analysis and recommendation in the years to come. However, it is essential to approach these advancements with caution, maintaining a balance between data-driven decision-making and understanding the broader economic, political, and social contexts for successful investment outcomes. In Table 1, compares the terminologies used in the different Research Papers to understand the concept that has already been used to predict the stock Prediction.

Table 1. Comparative Analysis.

REF.	Dataset	ALGORITHM	LIMITATION
[1]	NYSE Dataset	Multilayer Perceptron	Limited historical data, market volatility
[2]	Financial Market Data	Random Forest + LSTM	Dependency on model fusion, computational complexity
[3]	Various financial datasets	Decision Trees, Random Forest	Overfitting, data quality issues

[4]	Grouped Time-Series Data	LSTM, GRU	Data grouping challenges, model complexity
[5]	Stock Market Data	Random Forest	Model interpretability, market unpredictability
[6]	Stock Market Data	LSTM	Data noise, model hyperparameters
[7]	Stock Market Data	LSTM, RNN, CNN	Model complexity, data preprocessing
[8]	Stock Market Data	LSTM	Overfitting, data preprocessing
[9]	Stock Market Data	Decision Trees	Limited to historical data, market unpredictability
[10]	Financial News Corpus, Sentiment Analysis Data	LSTM	Sentiment analysis accuracy, market noise
[11]	Twitter and Stock Twits Data	NLP	Social media bias, sentiment accuracy
[12]	Stock Market Data	Support Vector Machines	Model interpretability, feature selection
[13]	Stock Market Data	Random Forest, Gradient Boosting	Model complexity, data preprocessing
[14]	Financial Market Data	ARIMA	Dependency on model accuracy, market dynamics
[15]	Stock Market Data	Gradient Boosting	Model interpretability, market noise
[16]	Stock Market Data	ARMA	Limited to pattern recognition, market noise
[17]	Stock Market Data	K-Nearest Neighbors	Model interpretability, computational complexity
[18]	Stock Market Data	ARIMA, SARIMA	Model performance, feature selection
[19]	Stock Market Data	AdaBoost	Model interpretability, data preprocessing
[20]	Stock Market Data	Literature review	Dependency on dataset variety, model selection
[21]	Stock Market Data	XGBoost	Model interpretability, data preprocessing
[22]	Stock Market Data	Random Forest	Model complexity, data preprocessing
[23]	Stock Market Data	Gradient Boosting	Model performance, data preprocessing
[24]	Stock Market Data	Logistic Regression	Model interpretability, feature selection
[25]	Stock Market Data	Linear Regression	Model interpretability, feature selection
[26]	Financial News Corpus, Sentiment Analysis Data of stock market	NLP	Sentiment analysis accuracy, data noise
[27]	Stock Market Data	Decision Trees	Dependency on historical data, model assumptions
[28]	Stock Market Data	Decision Trees	Model performance, feature selection
[29]	Social Media Data, Stock Market Data	CNN	Social media bias, sentiment accuracy
[30]	Social Media Data, Stock Market Data	Reinforcement Learning	Data integration challenges, model complexity
[31]	Stock Market Data	k-Means Clustering	Model interpretability, feature selection

[32]	Indian Stock Market Data, Sentiment Analysis Data of stock market	SVM	Sentiment analysis accuracy, data noise
[33]	Stock Market Data	Gradient Boosting	Model interpretability, data preprocessing
[34]	Stock Market Data	Bagging	Model performance evaluation, feature selection
[35]	Stock Market Data	Random Forest	Model complexity, data preprocessing
[36]	Stock Market Data	LSTM	Model interpretability, hyperparameter tuning
[37]	Stock Market Data	RNN, LSTM	Model interpretability, data preprocessing
[38]	Stock Market Data	LSTM, RNN, CNN	Dependency on feature selection, model hyperparameters
[39]	Stock Market Data	Linear Regression, Support Vector Machine	Sensitivity to outliers, limited performance in non-linear relationships
[40]	Various Data Sources	Decision Trees, Random Forest, K-Means	Data privacy concerns, scalability issues

3. Methodology

ALGORITHM: Stock Price Prediction with LSTM

Chart 1, describes the steps which are to be followed during the prediction. The code begins by importing necessary libraries such as NumPy, Pandas, Matplotlib, and scikit-learn's MinMaxScaler for data pre-processing as well as keras for building the LSTM model. The dataset, named "RELIANCE (telecom).xlsx"," meta(media).xlsx", "TATAMOTORS(Automobile).xlsx" is read into a Pandas DataFrame named 'stock_data'.

Null values in the dataset are checked and addressed. Rows containing null values are dropped or filled with appropriate values. The null values in the 'Open', 'High', 'Low', 'Close', Adj Close', and 'Volume' columns are handled by filling them with either zeros or the mean of the respective columns.

Features such as 'Open', 'High', 'Low', and 'Volume' are selected for modelling. The features are scaled using Min-Max scaling to ensure all the features have the same scale, which is essential for training neural networks. The dataset is split into training and test sets using Timeseries Split from scikit-learn to maintain temporal order in the data. The training and test data are reshaped to fit the input requirements of the LSTM model.

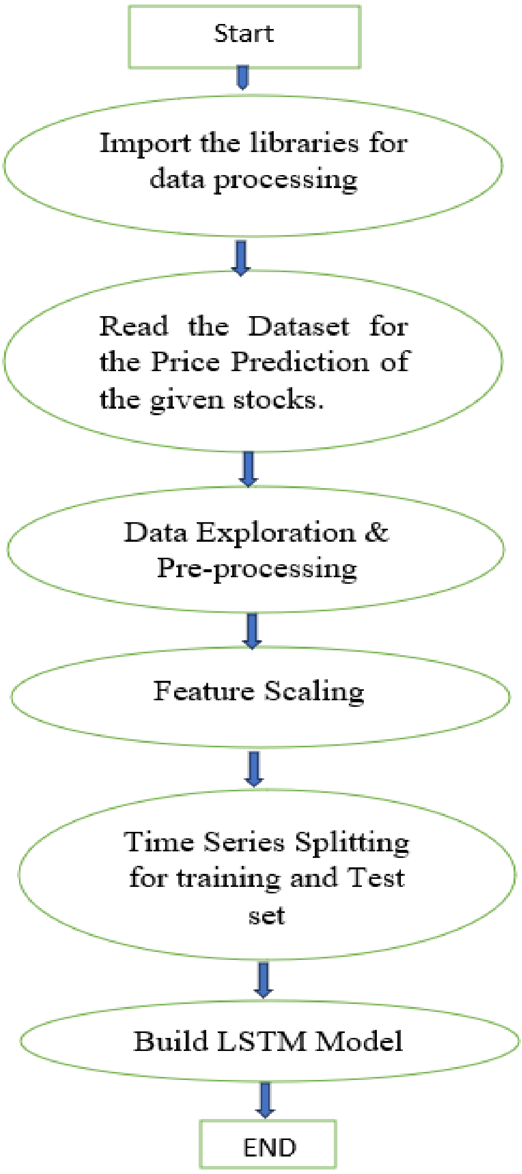

Chart 1. Stock Price Prediction with LSTM.

A sequential model is initialized. An LSTM layer with 32 units and a ReLU activation function is added. The input shape is defined based on the features. A Dense layer with one output neuron is added. The model is compiled using mean squared error loss and the Adam optimizer.

The LSTM model is trained on the training data for 100 epochs with a batch size of 8. Training history is stored for later analysis. The trained LSTM model is saved to a file named "RELIANCE (telecom)", "meta(media).xlsx", "TATAMOTORS(Automobile).xlsx". The saved model is loaded back for making predictions. Predictions are made on the test data using the trained LSTM model. Both the loaded model and the original model predictions are stored for comparison.

The true values and LSTM predicted values are plotted against the time scale. The plot visually compares the true and predicted Adj Close values, showcasing the performance of the LSTM model.

3.1. Graphical Representation of Adjusted Closing Prices

The adjusted closing prices of the stock under investigation are depicted graphically in Figure 3. The X-axis represents the temporal dimension, denoting discrete time intervals, while the Y-axis signifies the adjusted closing prices of the stock.

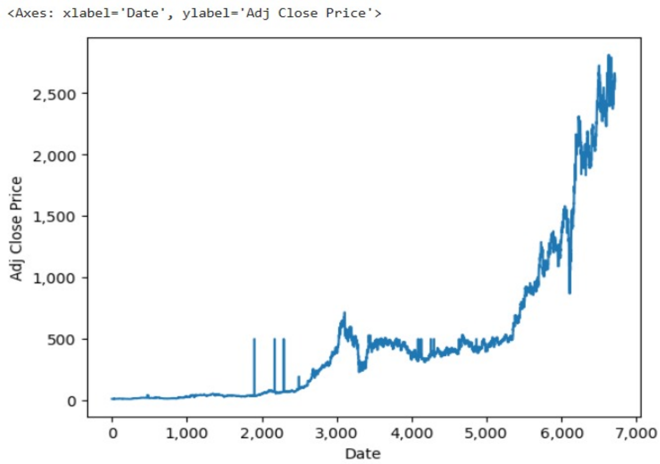

Figure 3. Adjusted Closing Prices.

The graph illustrates the historical trajectory of the stock's adjusted closing prices spanning the observed period. Each data point on the graph corresponds to the adjusted closing price of the stock at a specific point in time, reflecting its market valuation at the close of trading sessions.

In Conclusion, the graphical representation of adjusted closing prices serves as a valuable tool for analyzing the historical performance of the stock and deriving actionable insights for investment strategies.

3.2. Comparative Analysis of LSTM Predictions and True Values

The graph presented in Figure 4 just opposes the true values of the stock's adjusted prices against the predictions generated by the LSTM model.

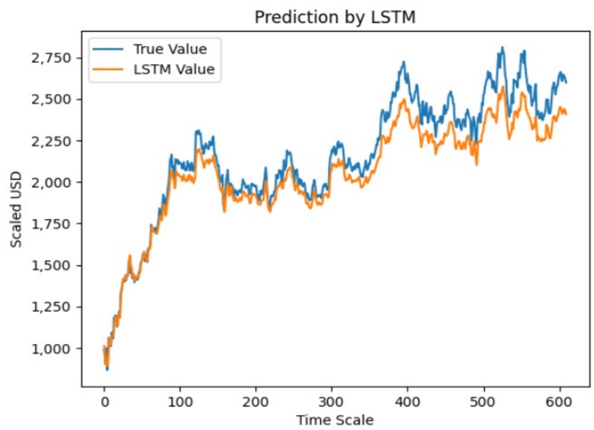

Figure 4. Comparative Analysis of True and LSTM.

The X-axis denotes the temporal dimension, representing discrete time intervals, while the Y-axis signifies the scaled values in USD.

The graph provides a visual assessment of the performance of the LSTM model in predicting the stock's adjusted closing prices over the observed time period. Two distinct lines are depicted on the graph: True Value Line: This line represents the actual adjusted closing prices of the stock, obtained from the dataset's test set. Each data point on this line corresponds to the true value of the stock's adjusted closing price at a specific point in time.

LSTM Value Line: This line illustrates the predicted values of the stock's adjusted closing prices generated by the LSTM model. The model's predictions are based on the input features provided to the model during the testing phase.

In conclusion, the graphical representation facilitates a comparative analysis of the true values and LSTM predictions, offering valuable insights into the performance and efficacy of the LSTM model in forecasting the stock's adjusted closing prices. The observations derived from the graph contribute to a comprehensive evaluation of the model's predictive accuracy and reliability, thereby informing decision-making processes in investment and financial analysis.

3.3. Dataset: RELIANCE (TELECOM).XLSX

The RELIANCE (TELECOM) dataset contains financial information related to Reliance Communications' stock performance. The dataset includes data on daily stock prices (open, high, low, and close), adjusted close prices, and trading volume.

The final output demonstrates the results in Figure 5 after applying a Long Short-Term Memory (LSTM) model to predict stock prices based on the RELIANCE (TELECOM) dataset. The LSTM model was trained on historical stock price data to forecast future prices.

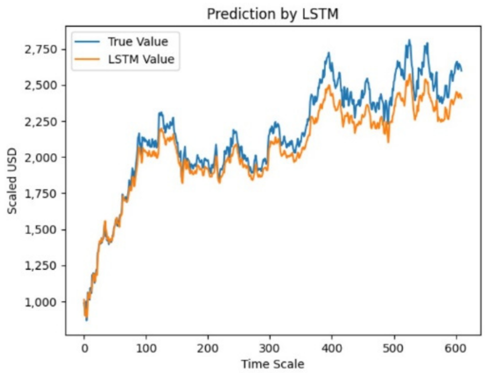

Figure 5. Prediction of Stock prices after LSTM algorithm.

3.4. Dataset: META (MEDIA).XLSX

The META dataset in contains financial information related to Meta Platforms, Inc. (formerly known as Facebook, California, United States). It includes data on daily stock prices (open, high, low, and close), adjusted close prices, and trading volume.

The final output demonstrates the results in Figure 6 of applying a Long Short-Term Memory (LSTM) model to predict stock prices based on the META (MEDIA) dataset. The LSTM model was trained on historical stock price data to forecast future prices.

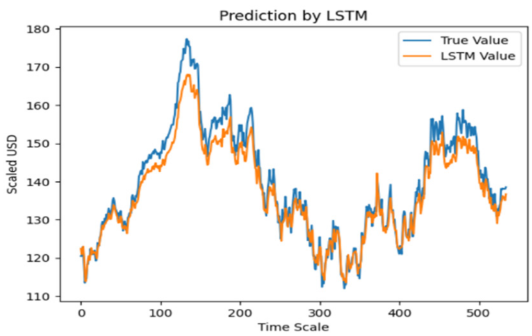

Figure 6. Prediction of Stock prices after LSTM algorithm.

3.5. Dataset: TATAMOTORS (AUTOMOBILE).XLSX

The TATAMOTORS (AUTOMOBILE) dataset contains financial information related to Tata Motors' stock performance. The dataset includes data on daily stock prices (open, high, low, and close), adjusted close prices, and trading volume. The final output demonstrates the results of Figure 7.

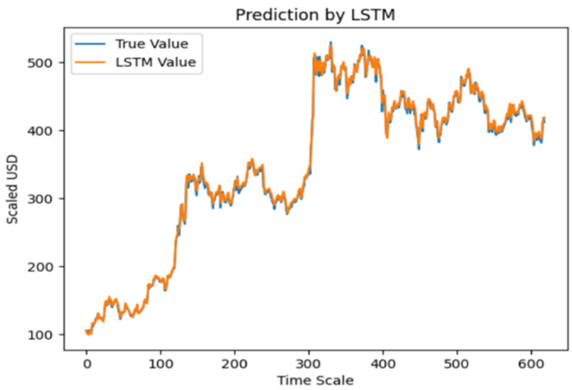

Figure 7. Prediction of Stock prices after LSTM algorithm.

After applying a Long Short-Term Memory (LSTM) model to predict stock prices based on the TATA MOTORS (AUTOMOBILE) dataset. The LSTM model was trained on historical stock price data to forecast future prices.

4. Algorithm: Stock Recommendation System

The code begins by importing necessary libraries including Pandas for data manipulation, scikit-learn for TF-IDF vectorization and cosine similarity computation, and NumPy for numerical operations in Chart 2. The dataset containing stock information is read into a Pandas DataFrame named 'stock_data’. The dataset likely includes attributes such as ‘Company Name’, ‘Symbol’, ‘Industry’, ‘Last Traded Price’, and ‘365 Day Percentage Change’.

Null values in the dataset are checked and addressed. Rows containing null values are either dropped or filled with appropriate values to ensure data integrity. Features relevant for recommendation, such as ‘Company Name’, ‘Symbol’, ‘Industry’, ‘Last Traded Price’, and ‘365 Day Percentage Change’, are selected from the dataset.

Numeric features like ‘Last Traded Price’ and ‘365 Day Percentage Change’ are converted to strings for TF-IDF vectorization. All selected features are combined into a single column to create documents for TF-IDF vectorization. TF-IDF vectorization is performed on the combined features using scikit-learn’s TfidVectorizer. This process converts text data into numerical vectors, capturing the importance of each word int the documents.

Cosine similarity scores between documents are computed using the linear_kernel function from scikit-learn. Cosine similarity measures the cosine of the angle between two vectors and is used to determine the similarity between stocks based on their features.

A function named ‘get_top_companies’ is defined to retrieve the top N similar companies for a given stock based on cosine similarity scores. The function takes the stock name, cosine similarity matrix, and the desired number of recommendations as inputs.

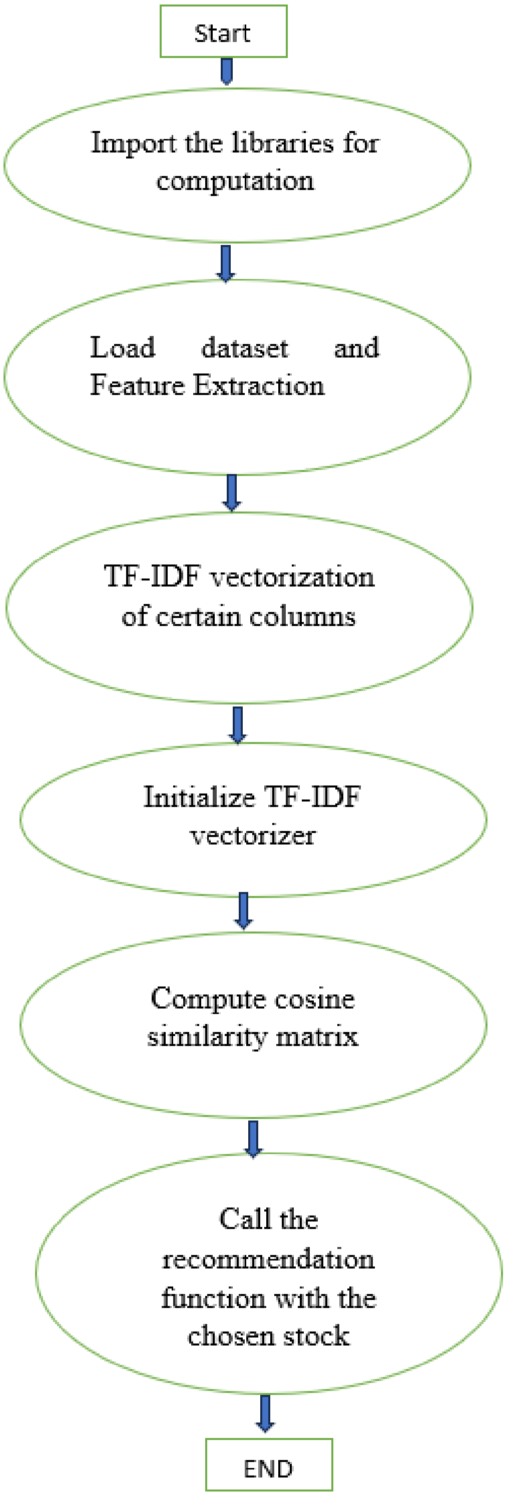

Chart 2. Algorithm for Recommendation System.

5. Final Output Analysis

The recommendation system successfully generated personalized recommendations for stocks based on their features. Using the example provided in the code for illustration with the help of graph as described in the Figure 8, the top 3 companies similar to ‘Gujarat Fluorochemicals Ltd. Vadodara, India’ Were recommended.

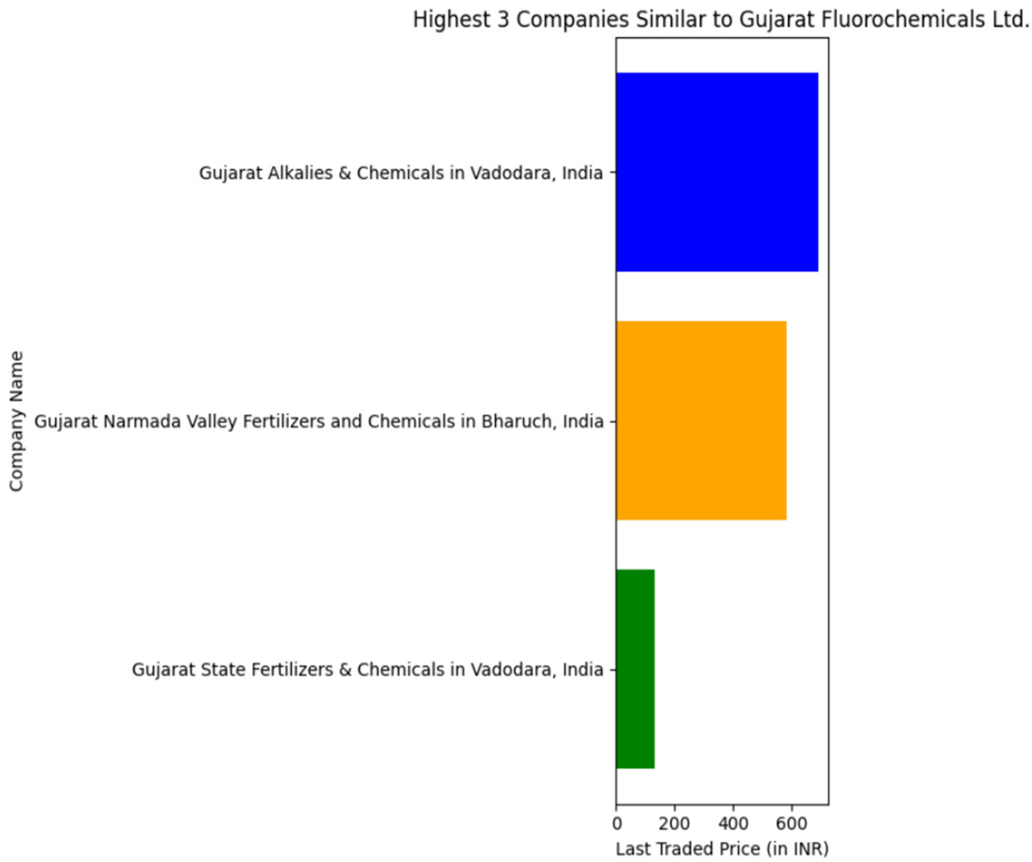

Figure 8. Recommendation based on the base company.

These recommendations in Figures 8 and 9 are based on the similarity of features such as company name, industry, last traded price, and 365-day percentage change. The recommendations aim to assist investors in identifying stocks that exhibit similar characteristics to the provided stock, Gujarat Fluorochemicals Ltd. Vadodara, India, thereby aiding in portfolio diversification and investment decision-making.

```
Highest 3 companies similar to Gujarat Fluorochemicals Ltd. based on features:
                                          Company Name     Symbol  \
178                 Gujarat Alkalies & Chemicals Ltd.  GUJALKALI
176          Gujarat State Fertilizers & Chemicals Ltd.       GSFC
163  Gujarat Narmada Valley Fertilizers and Chemica...       GNFC

    Last Traded Price 365 Day Percentage Change
178             689.0                     71.35
176             130.6                     14.09
163             580.5                     51.59
```

Figure 9. Recommendation of Stocks.

6. Conclusions

In conclusion, this project represents a concerted effort to revolutionize stock market analysis and recommendation by integrating traditional techniques with advanced machine learning methodologies. By combining historical stock data with Python programming, we aim to provide investors with actionable insights to navigate the complexities of the stock trading landscape. The integration of technical and fundamental analysis forms the cornerstone of our approach, enabling a deeper understanding of market dynamics and individual stock performance. Moreover, the incorporation of recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks enhances the predictive capabilities of our model, allowing for more accurate forecasting of stock price movements. Our project also introduces a practical recommendation system, providing investors with actionable guidance based on the outputs of our predictive model. Overall, our goal is to contribute to the advancement of stock market analysis and empower investors with the tools necessary to make informed decisions in today's dynamic financial environment.

Author Contributions
Conceptualization, K.A. and A.A.; methodology, K.A.; software, K.A.; validation, K.A., and A.A; formal analysis, K.A.; writing—original draft preparation, K.A.; writing—review and editing, K.A.; and K.K.G.; All authors have read and agreed to the published version of the manuscript. Authorship must be limited to those who have contributed substantially to the work reported.

Funding
This research received no external funding.

Conflict of Interest Statement
The authors declare no conflicts of interest.

Data Availability Statement
No new data were created.

References

1. W. Weina and M. Krishn. (2018). A novel stock trading prediction and recommendation system. *Multimedia Tools and Applications* 77. 1-13. 10.1007/s11042-017-4587-z.
2. S. Vikram; R. Manik and M. Gauri. (2023). Hybrid Approaches for Stocks Prediction and Recommendation System. *E3S Web of Conferences* 453. 10.1051/e3sconf/202345301047.
3. M. Latrisha; H. Jeta; A. Callista; A. Said & K. Aditya. (2023). Machine learning approaches in stock market prediction: A systematic literature review. *Procedia Computer Science* 216. 96-102. 10.1016/j.procs.2022.12.115.
4. L. Armin; M. Hendra and A. Supri. (2022). Implementation of Long Short-Term Memory and Gated Recurrent Units on grouped time-series data to predict stock prices accurately. *Journal of Big Data* 9. 10.1186/s40537-022-00597-0.
5. S. Dhruhi and S. Manan. (2023). Predicting stock market using machine learning: best and accurate way to know future stock prices. *International Journal of System Assurance Engineering and Management* 14. 10.1007/s13198-022-01811-1.
6. M. Adil and H. Mhamed. (2020). Stock Market Prediction Using LSTM Recurrent Neural Network. *Procedia Computer Science* 170. 1168-1173. 10.1016/j.procs.2020.03.049.
7. S. Sreelekshmy; R. Vinayakumar; Gopalakrishnan, E. A & Menon, Vijay and K. Soman. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 1643–1647. 10.1109/ICACCI.2017.8126078.
8. C. Shounak; B. Saptarshee; R. Saumik and D. Amit. (2023). Long-Short Term Memory Based Stock Market Analysis. 1–5. 10.1109/IEMENTech60402.2023.10423465.
9. A. Aggarwal, K. K. Gola and R. Kanauzia, "Digital Twin Technology with Machine Learning Algorithms for Reducing Wastage: A Comprehensive Review," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 2627-2632, doi: 10.1109/IC3I59117.2023.10398005.
10. M. Ganesh & K. Vaishali. (2023). Stock prediction based on sentiment analysis using long short term memory. *AIP Conference Proceedings* 020039. 10.1063/5.0176613.A. Bonnaccorsi. "On the Relationship between Firm Size and Export Intensity", *Journal of International Business Studies*, XXIII (4), pp. 605-635, 1992.
11. R. Caves. *Multinational Enterprise and Economic Analysis*, Cambridge University Press, Cambridge, 1982.
12. M. Clerc, M. "The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization". In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pp. 1951-1957, 1999.
13. H.H. Crokell. Specialization and International Competitiveness, in "*Managing the Multinational Subsidiary",* H. Etemad and L. S, Sulude (Eds.), Croom-Helm, London, 1986.
14. K. Deb, S. Agrawal, A. Pratab, T. Meyarivan. "A Fast Elitist Non-dominated Sorting Genetic Algorithms for Multiobjective Optimization: NSGA II". *KanGAL report 200001*, Indian Institute of Technology, Kanpur, India, 2000.
15. G. Arman; B. Debajyoti; K. Priyansh and S. Neelamani. (2023). Stock Market Price Prediction. *International Journal for Research in Applied Science and Engineering Technology* 11. 591-595. 10.22214/ijraset.2023.57184.
16. K. Saroj and S. Naman. (2009). Pattern Prediction in Stock Market. In "*AI 2009: Advances in Artificial Intelligence: 22nd Australasian Joint Conference, Melbourne, Australia, December 1-4, 2009. Proceedings 22",* 5866. 81–90. 10.1007/978-3-642-10439-8_9.
17. P. Jayesh and M. Rejo. (2020). Machine Learning Algorithms in Stock Market Prediction. In "*Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI-2018)",* 10.1007/978-3-030-24643-3_23.
18. T. Drashti; P. Dr and P. Prof. (2022). Stock Market Prediction Using LSTM Technique. *International Journal for Research in Applied Science and Engineering Technology* 10. 1820-1828. 10.22214/ijraset.2022.43976.
19. S. Jingyi and M. Omair Shafiq. (2020) Short-term stock market price trend prediction using a comprehensive deep learning system. *J Big Data* 7, 66. https://doi.org/10.1186/s40537-020-00333-6
20. K. Siddhesh; K. Anvit; K. Shubham and P. Deepali. (2022). Study and Analysis of Stock Market Prediction Techniques. *ITM Web of Conferences* 44, 03033. 10.1051/itmconf/20224403033.
21. R. Nusrat; M. Majid; A. Tasleem; S. Sparsh; S. Saurabh; A. Satyabrata; K. Hee-Cheol. (2021). Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions. *Electronics* 10. 2717. 10.3390/electronics10212717.

22. B. Prakash; P. Chinthan; B. Saleena & S. Harini. (2024). A systematic literature survey on recent trends in stock market prediction. *PeerJ Computer Science* 10. e1700. 10.7717/peerj-cs.1700.
23. S. Gurjeet. (2022). Machine Learning Models in Stock Market Prediction. *International Journal of Innovative Technology and Exploring Engineering* 11. 18-28. 10.35940/ijitee.C9733.0111322.
24. S. Troy; R. John; R. Thomas; H. Yu-Hsiang. (2020). Machine Learning Stock Market Prediction Studies: Review and Research Directions. *Journal of International Technology and Information Management* 28. 63-83. 10.58729/1941-6679.1435.
25. D. Manju. (2022). Machine learning techniques and methodology analysis for stock market price prediction. *International Journal of Computing and Artificial Intelligence* 3. 13-18. 10.33545/27076571.2022.v3.i2a.50.
26. P. Ramkrishna; C. Vikas; S. Deepika; S. Ashutosh. (2021). *Review of Stock Prediction Using Machine Learning Techniques* 10.1109/ICOEI51242.2021.9453099.
27. V. Mehar; C. Deeksha; T. Vinay; K. Arun. (2020). Stock Closing Price Prediction using Machine Learning Techniques. *Procedia Computer Science* 167. 599-606. 10.1016/j.procs.2020.03.326.
28. P. Shivani; V. Shweta; S. Smriti and T. Ms. (2024). Stock Price Prediction Using Sentiment Analysis. *International Journal for Research in Applied Science and Engineering Technology* 12. 1301-1305. 10.22214/ijraset.2024.58167.
29. A. Sonali; K. Saloni; K. Anushka; K. Rohan; K. Prathamesh; K. Dhawal and K. Raj. (2023). Stock Price Prediction Website Using Linear Regression - A Machine Learning Algorithm. *ITM Web of Conferences* 56. 10.1051/itmconf/20235605016.
30. P. Chinthan; R. Mishra; B. Prakash and B. Saleena. (2023). Stock Market Prediction Using Machine Learning Techniques: A Comparative Study. 10.1007/978-981-99-1051-9_7.
31. M. Pooja; P. Sharnil; K. Ketan. (2021). Harvesting social media sentiment analysis to enhance stock market prediction using deep learning. *PeerJ Computer Science* 7. e476. 10.7717/peerj-cs.476.
32. L. Mei and Z. Ye. (2023). Integrating Social Media Data and Historical Stock Prices for Predictive Analysis: A Reinforcement Learning Approach. *International Journal of Advanced Computer Science and Applications* 14. 10.14569/IJACSA.2023.0141203.
33. P. Ashwini. (2020). Study of Machine learning Algorithms for Stock Market Prediction. *International Journal of Engineering Research and Technology* V9. 10.17577/IJERTV9IS060064.
34. H. Sepp and S. Jürgen. (1997). Long Short-term Memory. *Neural Computation* 9. 1735-80. 10.1162/neco.1997.9.8.1735.
35. D. Stéphane; J. Thibault and N. Alexandra. (2022). Performance attribution of machine learning methods for stock returns prediction. *The Journal of Finance and Data Science* 8. 10.1016/j.jfds.2022.04.002.
36. O. Mehtabhorn; T. Nongnuch; J. Watthanasak and Wills, Gary. (2020). A Survey on Machine Learning for Stock Price Prediction: Algorithms and Techniques. 10.5220/0009340700630071.
37. Dash, A.S. & Mishra, U.. (2023). Sentiment Analysis using Machine learning for forecasting Indian stock Trend: A brief Survey. *Finance: Theory and Practice* 27. 136-147. 10.26794/2587-5671-2023-27-6-136-147.
38. S. Yang. (2023). Research on the Stock Price Prediction Using Machine Learning. Advances in Economics, Management and Political Sciences. 22. 174-179. 10.54254/2754-1169/22/20230307.
39. P. Bhawna; D. Gaurav; J. Prashant; Y. Sudeept and G. Nitin. (2021). Stock Market Prediction Using Linear Regression and SVM. In "*2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)",* 629-631. 10.1109/ICACITE51222.2021.9404733.
40. A. Aggarwal, V. Kumar and R. Gupta, "Object Detection Based Approaches in Image Classification: A Brief Overview," 2023 IEEE Guwahati Subsection Conference (GCON), Guwahati, India, 2023, pp. 1-6, doi: 10.1109/GCON58516.2023.10183609.

Author Biography

Krishna Arora has done matrix and intermediate from Dayanand Anglo Vedic (D.A.V) School in Haridwar which is affiliated to the Central Board of Secondary Education (C.B.S.E) with an aggregate of 86% and 74% respectively. Currently, he is pursuing a Bachelor in Technology (B-TECH) in the branch of Information Technology (IT) from the College of Engineering Roorkee which is approved by AICTE New Delhi and affiliated to Veer Madho Singh Bhandari Uttarakhand Technical University. He is proficient in programming, teamwork, and communication. His technical skills include Python programming, C programming, and some basic knowledge of Data Structure, and C++ and his strong interpersonal skills have allowed him to connect with different people.

Akshima Aggarwal, pursuing PHD in Computer Science Engineering from IIIT Una and post-graduated in the same field from Guru Nanak Dev University. She worked in TCS for two years as a System Engineer after completing her graduation. She published many research papers in SCOPUS and IEEE and Book Chapters in Springer. She has also presented and published paper in a UGC sponsored National Seminar. She has filed multiple Indian patents. Studying and working at universities for a long time has given her a lot of valuable insights into learning and teaching. Advisable teaching attracts students' interest in the subject. With a strong background in Computer Science and various teachings and project experiences, she can teach a broad set of courses

Kamal Kumar Gola is an Assistant Professor at COER University, Roorkee, India. He completed his B.Tech. Degree in Computer Science and Engineering from Moradabad Institute of Technology and his M.Tech. Degree in Computer Science and Engineering from Uttarakhand Technical University. He is pursuing a Ph.D. in Computer Science and Engineering from Indian Institute of Technology, Jodhpur. With over 14 years of university teaching experience and six months of industrial experience, Kamal Kumar Gola is a seasoned educator with a wealth of knowledge in his field. He has published more than fifty papers in international journals and has participated in various international and national conferences where he has presented his research. He kept himself updated by attending various professional development programs, workshops, and training courses organized by esteemed institutions such as IIT, NIT, IIIT, and others. His main research interests lie in Underwater Acoustic Sensor Networks, Algorithms and Security.