International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 8, pp 44-49, August-2025

https://doi.org/10.47001/IRJIET/2025.908006

Forecasting Future Trends: An LSTM Approach to Identifying High Growth Sectors

[1]Heena J. Patel, [2]Dr. Prashant P. Pittalia

[1]Assistant Professor, M.K. Institute of Computer Studies, Bharuch-392001, Gujarat, India

[2]Professor, P.G. Department of Computer Science and Technology, Sardar Patel University-Vallabh Vidyanagar, Gujarat, India

***Abstract* - Predicting trends in the financial stock market is a challenging task for researchers due to its complex and dynamic nature. The stock market has always been complex due to the market's volatility and non-linearity. Accurate forecasting is difficult, and forecasting errors can result in significant investment risks, even with techniques like diffusion modelling and forecasting. Understanding the pattern of the sector price of the particular company by predicting its financial growth and future development will be highly beneficial. This study optimizes a predictive machine learning model based on long-short term memory (LSTM) neural networks to predict the most performing sector in the Indian sector indices. Using historical data, the LSTM model is used to predict future sector developments. The historical data from the past five years was obtained via Yahoo Finance from January 1, 2019, to December 31, 2023. The proposed method is designed for the ten sectors of the Indian economy. An LSTM model is designed to predict the future sector performance. To predict the three months after, i.e, on April 30, 2024, the actual and predicted returns of each sector are computed. This method is used to choose the most progressive sector based on a ranking system. The proposed model indicates the high accuracy of the LSTM model.**

Keywords: Portfolio optimization, Machine Learning, Sector Analysis, Stock Prediction, Time Series, Long-Short Term Memory(LSTM), Mean Absolute Percentage Error(MAPE), Root Mean Square Error(RMSE), R2.

I. INTRODUCTION

Investing in the stock market involves significant uncertainty. The ability to forecast trends and select profitable sectors and stocks is critical for investors and portfolio managers. Macroeconomic indicators, sector performance, and individual stock results shape the financial landscape, making it complex to develop reliable strategies for stock selection. Portfolio optimization is an analytical process of selecting and allocating a group of assets to maximize expected return and minimize risk. Markowitz introduced a mean-variance model based on Modern Portfolio Theory in 1952 [1][2]. Finding the best mean-variance portfolio, based on Markowitz's foundational work, is known to be an NP-hard computational problem. For this reason of complexity, many researchers have proposed various methods for portfolio optimization and stock price prediction. These approaches aim to improve the efficiency and effectiveness of investment strategies by leveraging advanced algorithms and models to navigate the complexities of the financial market. Among the methods in the literature for stock price prediction, regression, support vector machines, ARIMA, VAR, time series forecasting, and machine learning approaches are popular [3]. Many researchers investigate stock price prediction using deep learning-based LSTM models for selected stocks, but fewer analyze sector performance using machine learning models.

This paper proposes a method for designing an efficient portfolio by choosing the sector from the National Stock Exchange (NSE) of India [4]. Portfolios are designed for selecting the most progressive sector using actual return and predicted return. The past data of these ten sectors for the past five years are extracted using Python from the Yahoo Finance site. To construct the portfolio using LSTM model for predicting future price and future return of the portfolio for different forecast horizons. Three months after the portfolio is constructed with actual return and predicted return by the LSTM model, and compare to evaluate the accuracy of the predicted model. The ten sector studies in the proposed work are auto, information technology, bank, FMCG, metal, media, PSUBANK, reality, consumer durable, and financial services.

The main contribution of the current work is divided into three parts. First, select the ten sectors listed in the NSE sector indices. Second, it proposes an efficient and optimized design of an LSTM model for predicting the future price of the selected ten sectors. Third, evaluate the actual return and predicted return, and calculate the performance matrix.

This paper introduces a systematic approach that utilizes machine learning (ML) to apply time-series forecasting using an LSTM model to predict the future performance of the sector in the selected sector. We aim to provide investors with a reliable framework to identify the best-performing sector for investment purposes.

ISSN (online): 2581-3048
Volume 9, Issue 8, pp 44-49, August-2025
https://doi.org/10.47001/IRJIET/2025.908006

This paper is divided into five sections. Section II defines some existing works on portfolio design, and sector price predictions are discussed briefly. Section III provides the methodology used to choose the most performing sector. Section IV discusses the desired result and analysis. Section V concludes the paper.

II. LITERATURE REVIEW

The literature on constructing resilient stock portfolio systems through advanced predictive models is extensive. Numerous research scholars have proposed a range of sophisticated techniques and approaches to achieve accurate forecasts of future stock price movements.

In Paper [2], the authors propose a novel portfolio construction technique based on a hybrid machine learning model that integrates a Convolutional Neural Network (CNN) with Bidirectional Long Short-Term Memory (Bi-LSTM). The model uses various input features and applies hyperparameters to enhance stock price prediction. The study evaluates the prediction performance using metrics like mean absolute error (MAE), mean square error (MSE), and mean absolute performance error(MAPE). Finally, predicted returns are integrated with the MV model for optimal portfolio construction. This paper presents a comparative study of the predictive performance of LSTM, BiLSTM, and CNN-BiLSTM. The experimental results show that the BiLSTM model outperforms the other predictive models.

Sectoral analysis plays a crucial role in portfolio management. Prior research give emphasized to the various variable used for stock prediction including financial ratios such as Price-to-Earnings (PE), Return on Equity (ROE), Earnings Per Share (EPS), and Price-to-Book (PB) ratio, along with technical indicator and micro economic variable like consumer price index, government spending, gross domestic product(GDP), industrial production, private spending, producer price index. These variables have been combined with various machine learning algorithms, such as ANN, SVM, and LSTM, to predict the sectoral growth. This study concludes that LSTM is the most advanced deep learning technique for stock prediction compared to random forest and logistic regression [4]. Additionally, statistical models such as ARIMA [3] [5-8], linear regression [7], and a deep learning model, LSTM [5-7], have been widely used in financial forecasting. Based on these surveys and comparative analysis with the other two models, LSTM outperforms and makes accurate predictions in portfolio optimization for the financial market. In contrast, ARIMA models often fail to capture the intricate patterns within stock prices, leading to inaccurate predictions.

Recent advancements in deep learning, especially the application of recurrent neural networks (RNNs) like LSTM, have demonstrated superior performance in time-series forecasting. LSTM models are known for their ability to retain long-term dependencies, making them ideal for financial data, which exhibits sequential behaviour over time. Researchers have explored numerous optimization techniques to fine-tune LSTMs, focusing on both structural and hyperparameter optimization [9-12].

III. METHODOLOGY

The methodology in this work is divided into six major steps. The steps are as follows. a) Data acquisition b) Pre-processing and feature extraction c) Design model with optimized parameters d) Processing the output model e) Testing and error correction f) Plotting the output for visualization.

A. Data Acquisition

Ten sectors are selected from the NSE, India. The data used in this study spans the past five years for ten sectors. In each sector, the historical prices are extracted from the Yahoo Finance website using the yfinance package in Python. Python libraries are used to extract this data. All sectors' daily data were extracted from January 1, 2019, to December 31, 2023. It contains 5000 records of the monthly sector. Each record contains information on high, low, opening, closing, and volume value of the sector for that period. Predict the next 3 months' price, convert it into monthly data. The selected sectors are the following: Nifty Auto, Nifty IT, Nifty Bank, Nifty PSU Bank, Nifty Reality, Nifty Metal, Nifty Media, Nifty Consumer and Durable, Nifty FMCG, Nifty Financial Services, based on the contribution of the sectoral index. The selected sector is identified in the report published by NSE on Apr 30, 2024[3].

B. Feature Selection and Data Preprocessing

In the beginning stage of data pre-processing, noisy and redundant records are removed to enhance the quality of the data. Specifically, records with zero trading volume are excluded. The sectoral dataset includes six features: Open, High, Low, Close, Volume, and Adjusted Close. Since the current study focuses on univariate analysis, the "Close" price is selected as the primary variable for modelling, while the remaining features are omitted.

To normalize the input data, the MinMaxScaler function in Python is used to regularize the selected feature. This technique is utilized to rescale the data into a common range between 0 and 1, ensuring uniformity and minimizing the effect of scale differences.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 8, pp 44-49, August-2025

https://doi.org/10.47001/IRJIET/2025.908006

The dataset is divided into training and testing subsets. The training set covers the period from January 1, 2019, to January 1, 2022, while the testing set covers from January 1, 2023, to January 1, 2024. After training the model on the training data, predictions are generated for the testing data. The model is used to forecast the stock price for the next three months and evaluate the performance using various error metrics, such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE).

C. Model Design with Optimize Parameter

Once sectors are selected, employ a machine learning model to predict the performance of sectors. Time-series data for the individual sector is processed and fed into an LSTM model for future price prediction.

LSTM is an extended and advanced version of a recurrent neural network (RNN) [10] with high accuracy and predicts the future value of sequential data. The LSTM model maintains the information in the memory cell and gates. There are three gates (input, forget, and output) in LSTM. [11-13]

- Input gate: To decide how much data is needed for the cell.
- Forget gate: To decide to keep or remove the incoming information from the memory.
- Output gate: To determine which output is to be kept in memory. The predicted value is available at the output gate.

For predicting the sector price, the LSTM model was designed, and the fine-tuned parameters are listed below.

i. Time step is the number of previous days' data that will be used to predict the next day's predictions. The model uses the monthly closing price for the previous 6 months of data.
ii. Feature is the number of input variables. The model uses a single feature that is a closed value. It is a univariate model.
iii. Neuron in each layer, the model uses two layers. The first layer has 64 units. The second layer has 32 units.
iv. Dense layer, the model uses two dense layers, one has 64 units and the second has 1 unit.
v. Dropout rate is used to avoid the overfitting of the model. The model uses a twenty percent dropout rate.

The forecast horizon is adjusted with different parameters by tuning the model. To train the model, a batch size of 64 and 100 epochs is used for the model. The hyperbolic tangent function (tanh) activation function is used for all layers except for the output layer. The sigmoid function is used for activation at the final output layer. The model is compiled using the Adam optimizer and MAE loss function. The Adam optimizer is a popular optimizer for training machine learning models. It is known for its ability to converge quickly to good solutions. Mean absolute error (MAE) is a common loss function for regression tasks. It measures the average absolute difference between the predicted values and the actual values. The hyperparameter values are selected with the help of the grid search method.

D. Processing the Output Model

It involves the execution of the Python function that works on two arguments. 1) Calculate the actual price and the actual return based on past data. 2) Find the future actual value for the given time step and the calculated predicted return. Table 1 illustrate the sector wise actual and predicted data, and calculate forecast return value for next 3 month using LSTM.

Table 1: Sector-wise LSTM prediction and Return

SECTOR NAME	31-DEC-2023		31-MAR-2024		RETURNS	
	Actual Price	Predicted Price	Actual Price	Predicted Price	Actual	Predicted
AUTO	18618	18114	21419	21936	15.04	17.81
IT	35515	32363	34898	34166	-1.73	-3.79
BANK	48292	45520	47125	48212	-2.41	-0.16
FMCG	56987	52514	53949	54594	-5.33	-4.19
METAL	7978	7413	8257	8463	3.49	6.08
MEDIA	2388	2469	1796	2790	-24.80	16.83
PSUBANK	5713	5509	7007	6666	22.64	16.67
REALITY	783	687	901	926	15.01	18.28
CONSUMABLE	9577	9156	10128	10147	5.74	5.93
FIN SERVICES	21487	19564	20989	19746	-2.31	-8.10

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 8, pp 44-49, August-2025

https://doi.org/10.47001/IRJIET/2025.908006

E. Testing and Error Calculation

Once the model is compiled, train it on the training data. After training the model, it is used to predict the new data. Table 2 shows the performance metrics evaluated for accuracy.

i. Mean Squared Error (MSE): Measures the size of the prediction error. Lower MSE value signifying better model accuracy.
ii. Root Mean Squared Error (RMSE): Measures the square root of MSE. A lower RMSE value suggests better model accuracy.
iii. Mean Absolute Percentage Error (MAPE): Indicate the average percentage difference between predicted and actual values.
iv. Coefficient of Determination (R2): Evaluates the extent to which the model explains the variability in the actual data, with values closer to 1 denoting a higher quality fit.

Table 2: Sector wise Evaluation matrix

SECTOR NAME	MAPE	RMSE	R2 SCORE	ACCURACY
AUTO	0.05	0.05	0.58	94.35
IT	0.05	0.05	0.34	94.56
BANK	0.06	0.06	-0.27	93.31
FMCG	0.10	0.11	-1.76	89.86
METAL	0.05	0.05	0.46	94.28
MEDIA	0.07	0.06	0.14	92.48
PSUBANK	0.05	0.05	0.64	94.32
REALITY	0.08	0.09	0.57	91.04
CONSUMABLE	0.04	0.05	0.61	95.22
FIN SERVICES	0.08	0.09	-1.69	91.12

F. Plotting the output for visualization

Figures 1 to 4 illustrate the actual, predicted, and forecasted prices across various sectors. It also is highlighting the model's performance in capturing market trends and future movements.

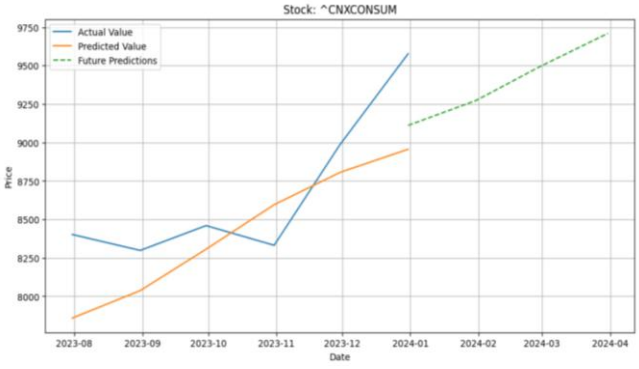

Figure 1: Consumable sector

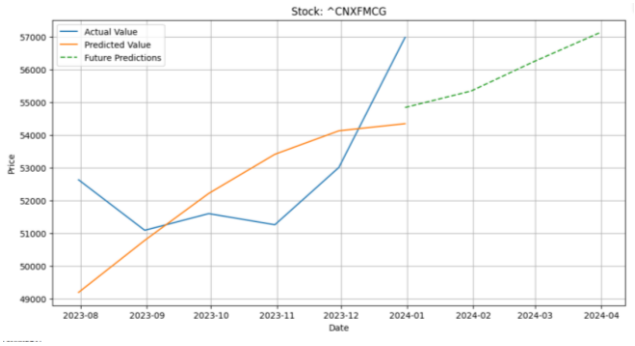

Figure 2: FMCG sector

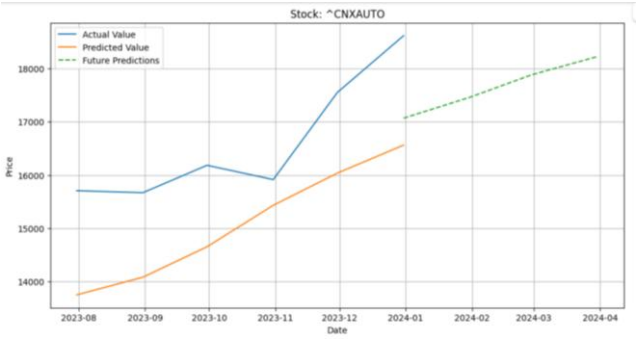

Figure 3: AUTO sector

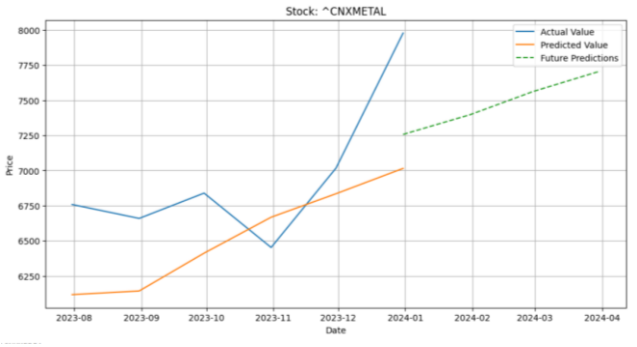

Figure 4: METAL Sector

For the ten sectors, the LSTM model is applied to individual sectors. The model was trained and tested on the historical data and forecast the predicted value. This study shows that REALITY, AUTO, MEDIA, and PSUBANK has most demanding sector, with highest predicted return, making it for future investment.

The results showed that CONSUMABLE had the most consistent upward price trend, with the lowest RMSE score and highest accuracy, making it the top candidate for future investment.

IV. RESULTS AND DISCUSSIONS

For the ten sectors, the LSTM model is applied to individual sectors. The model was trained and tested on the historical data and forecast the predicted value. This study

ISSN (online): 2581-3048
Volume 9, Issue 8, pp 44-49, August-2025
https://doi.org/10.47001/IRJIET/2025.908006

shows that REALITY, AUTO, MEDIA, and PSUBANK has most demanding sector, with highest predicted return, making it for future investment.

The results showed that CONSUMABLE had the most consistent upward price trend, with the lowest RMSE score and highest accuracy, making it the top candidate for future investment.

V. CONCLUSION

This paper presents a technique for selecting progressive sectors using machine learning techniques. Combining machine learning with deep learning provides an effective framework for investors to make informed decisions. The use of LSTM for stock price prediction has shown promising results, particularly in forecasting future price movements based on historical data.

The proposed technique can be extended to different sectors and more complex prediction models, allowing for further refinement and application across various markets. Also, extend the selection of stocks based on the most progressive sector and choose the weightage of each stock to generate the progressive portfolio.

REFERENCES

[1] H. Markowitz, "Portfolio Selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952, doi: 10.1111/j.1540-6261.1952.tb01525.x.

[2] Chaweewanchon, A., & Chaysiri, R. (2022). Markowitz Mean-Variance Portfolio Optimization with Predictive Stock Selection Using Machine Learning. *International Journal of Financial Studies*, 10(3), 64. https://doi.org/10.3390/ijfs10030064.

[3] Dariusz Kobielaa, Dawid Kreftaa, Weronika Krol´, Paweł Weichbroth "ARIMA vs LSTM on NASDAQ stock exchange data", *26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems* (KES 2022).

[4] www.NSE.india

[5] M. Kumbure, "Machine learning techniques and data Stock market forecasting: A literature Review of the portfolio optimization problem," doi: 10.1016/j.eswa2022.116659.

[6] Tanya Garg, Daljeetsingh Bawa, "Evaluation of Statistics model for timeseries data forecasting", *Science and technology journal* vol. II issue: I January 2023, ISSN:2312-3388, https/doi.org/10.2223/stj.2022.11.01.06

[7] K. Vikranth, P.S. Nethravathi and K. Krishna Prasad, "PRICE PREDICTION SYSTEM – A PREDICTIVE DATA ANALYTICS USING ARIMA MODEL", *ICTACT JOURNAL ON SOFT COMPUTING*, JANUARY 2024, VOLUME: 14, ISSUE: 03, ISSN: 2229-6956 (ONLINE), DOI: 10.21917/ijsc 2024.0463.

[8] Kelvin J.L. Koa, "Diffusion Variational Autoencoder for Tackling Stochasticity in Multi-Step Regression Stock Price Prediction", *CIKM '23: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, October 2023, Pages 1087 – 1096, https://doi.org/10.1145/3583780.3614844

[9] J, K., E, H., Jacob, M. S., & R, D. (2021). Stock price prediction based on an LSTM Deep learning model. *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 1–4. https://doi.org/10.1109/icscan53069.2021.9526491

[10] Ahire, P., Lad, H., Parekh, S., Kabrawala, S., Student, & Engineering, C. (2021). LSTM BASED STOCK PRICE PREDICTION.

[11] J. Sen, S. Mehtab, A. Dutta, and S. Mondal, "Precise Stock Price Prediction for Optimized Portfolio Design Using an LSTM Model," *2021 19th OITS International Conference on Information Technology (OCIT)*, Bhubaneswar, India, 2021, pp. 210-215, doi: 10.1109/OCIT53463.2021.00050.

[12] Tripathi, S. S., & Tripathi, S. (2022). Can stock prices be predicted? A Comparative study of LSTM and SVR for financial market forecast. *TechRxiv*. https://doi.org/10.36227/techrxiv.21158932.v2

[13] Pramod, & Pm, Mallikarjuna. (2021). Stock Price Prediction Using LSTM. *Test Engineering and Management*. 83. 5246-5251.

[14] Hiransha Ma, Gopalakrishnan E.Ab , Vijay Krishna Menonab, Soman K.P.*.," NSE stock market prediction using deep learning model", *International Conference on Computational Intelligence and Data Science (ICCIDS 2018*).

[15] Geeta Kolte, Varadraj Kini, Harikrishnan Nair, Prof. Suresh Babu K. S.," Stock Market Prediction using Deep Learning" ISIN NO: 2321-9653, *IJRASET*, https://doi.org/10.22214/ijraset.2022.41159

[16] Sen, J. (2021). Robust Portfolio Design and Stock Price Prediction Using an Optimized LSTM Model. *2021 IEEE 18th India Council International Conference (INDICON)*. OR Sen, Jaydip. "Robust Portfolio Design and Stock Price Prediction Using an Optimized LSTM Model." *2021 IEEE 18th India Council International Conference (INDICON)*, 2021.

[17] Sen, Jaydip. "Analysis of Sectoral Profitability of the Indian Stock Market Using an LSTM Regression Model." *ArXiv (Cornell University)*, 2021.

ISSN (online): 2581-3048
Volume 9, Issue 8, pp 44-49, August-2025
https://doi.org/10.47001/IRJIET/2025.908006

[18] A.Chatterjee, H. Bhowmick, and J. Sen, "Stock Volatility Prediction using Time Series and Deep Learning Approach," *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, Mysuru, India, 2022, pp. 1-6, doi: 10.1109/MysuruCon55714.2022.9972559.

Citation of this Article:

Heena J. Patel, & Dr. Prashant P. Pittalia. (2025). Forecasting Future Trends: An LSTM Approach to Identifying High Growth Sectors. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(8), 44-49. Article DOI https://doi.org/10.47001/IRJIET/2025.908006

****

