****

**Unraveling Stock Market Insights with Advanced Learning Models**

**A Capstone Project report submitted in partial fulfillment of the requirements for the Post Graduate Program in Data Science at Praxis Tech School, Kolkata, India**

#### By

**SUYASH MAURYA (Roll No: A23046)**

**SK MONIRUL ISLAM (Roll No: A23040)**

**NIKITA DAS (Roll No: A23027)**

**SK RITURAJ ALAM (Roll No: A23042)**

**SAYANTAN SAHA (Roll No: A23037)**

**SHREYAN BHATTACHARJEE (Roll No: A23038)**

**BAISAKHI BISWAS (Roll No: A23011)**

**SK MOSIUR RAHMAN (Roll No: A23041)**

**Of fall 2024 Batch of Data Science**

**Under the supervision of**

**Prof. Jaydip Sen**

### Praxis Tech School, Kolkata, India

**Acknowledgment**

We express our heartfelt gratitude to all those who contributed to the successful completion of this project. First and foremost, we extend our sincere appreciation to our supervisor Dr. Jaydip Sen, whose guidance, support, and invaluable insights were instrumental throughout the research process. Their expertise and encouragement propelled us forward, enabling us to navigate challenges and achieve our objectives effectively. We are also deeply thankful to Praxis Tech School, for providing the necessary resources and facilities essential for conducting this study. The conducive environment fostered a culture of innovation and collaboration, facilitating our research endeavors. Furthermore, we acknowledge the participants and stakeholders whose cooperation and contributions were integral to the data collection process. Their willingness to share insights and provide access to relevant information significantly enriched our analysis and findings. We are grateful to our colleagues and peers for their encouragement, constructive feedback, and stimulating discussions, which enriched our understanding and enhanced the quality of our work. Last but not least, we extend our heartfelt thanks to our family and friends for their unwavering support, understanding, and patience throughout this journey. Each of these individuals and entities played a crucial role in the realization of this project, and for that, we extend our sincerest appreciation.

Thank you.

**Abstract**

This study endeavors to construct predictive models for forecasting stock prices within the National Stock Exchange (NSE) of India, focusing on five diverse sectors: IT, steel, auto, banking, and FMCG. Incorporating data from two major indices, NIFTY and S&P 500. The analysis encompasses a comprehensive dataset spanning five years, from 2018 to 2022. The first three years' data (2018-2020) are dedicated to training the models, while the subsequent two years (2021-2022) are utilized for rigorous testing. Data points are collected at daily intervals to ensure accuracy and reliability in the analysis process. Advanced deep learning architectures, including Long-and-Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), are employed to extract meaningful insights and patterns from the data. Both univariate and multivariate analyses are conducted to discern relevant factors influencing stock price movements across the selected sectors and indices. This research aims to challenge the efficient market hypothesis by demonstrating the feasibility of accurate stock price prediction despite market stochasticity. By meticulously designing models and selecting pertinent variables, the study endeavors to provide stakeholders with actionable insights for informed decision-making in stock trading and investment activities. The performance of each predictive model is thoroughly evaluated, considering factors such as prediction accuracy and computational efficiency. Comparative analyses are conducted to assess the effectiveness of different model architectures and methodologies.

Ultimately, this research contributes to advancing the field of stock price forecasting within the NSE of India, offering valuable insights that can enhance risk management strategies and decision-making processes for market participants.

**Keywords**: Stock Price Forecasting, Deep Learning, Univariate Analysis, Multivariate Analysis, Time Series Regression, Root Mean Square Error (RMSE), Long-and-Short-Term Memory (LSTM) Network, Convolutional Neural Network (CNN)

### TABLE OF CONTENT

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No** |
|  | **List of Figure** | **5** |
| **1** | **Introduction** | **6** |
| **2** | **Related Work** | **7** |
| **3** | **Methodology** | **8** |
| **4** | Deep Learning ModelsConvolutional Neural NetworkLong- and Short-Term Memory Network | **9 10 13** |
| **5** | Performance Results and Analysis1. Results of Deep Learning Models2. Summary of Performance Results of the Models | **16** |
| **6** | **Conclusion and Future Work** | **30** |
| **7** | **References** | **31** |
|  | | |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| Fig No. | Description of Figure | Page No |
| 1(a) | Variation of RMSE with different days in week for Nifty50 (CNN\_UNI-5) | 24 |
| 1(b) | Actual vs Predicted graph for Nifty50 (CNN\_UNI-5) | 24 |
| 2(a) | Variation of RMSE with different days in week for Nifty50 (CNN\_UNI-10) | 25 |
| 2(b) | Actual vs Predicted graph for Nifty50 (CNN\_UNI-10) | 25 |
| 3(a) | Variation of RMSE with different days in week for Nifty50 (CNN\_MULTI-10) | 26 |
| 3(b) | Actual vs Predicted graph for Nifty50 (CNN\_MULTI-10) | 26 |
| 4(a) | Variation of RMSE with different days in week for S&P500 (LSTM\_UNI-5) | 27 |
| 4(b) | Actual vs Predicted graph for S&P500 (LSTM\_UNI-5) | 27 |
| 5(a) | Variation of RMSE with different days in week for S&P500 (LSTM\_UNI-10) | 28 |
| 5(b) | Actual vs Predicted graph for S&P500 (LSTM\_UNI-10) | 28 |
| 6(a) | Variation of RMSE with different days in week for Nifty50 (LSTM\_MULTI-10) | 29 |
| 6(b) | Actual vs Predicted graph for Nifty50 (LSTM\_MULTI-10) | 29 |

**Chapter 1**

**Introduction**

The stock market, often referred to as the share market, serves as a primary platform where business entities and individuals closely monitor their ownership stakes in various companies and organizations. Given its pivotal role in determining a nation's economic prosperity, the stock market holds significant sway over economic activities and investor sentiments. Predicting the trajectory of the stock market proves to be an arduous task owing to the dynamic and ever-evolving nature of market data. Despite numerous attempts using conventional methods in the past, accurate stock market predictions have remained elusive, casting doubt on the reliability of traditional forecasting approaches.

In recent years, the emergence of machine learning (ML) algorithms has brought renewed hope for improving the accuracy of stock market forecasts. By leveraging sophisticated algorithms and vast datasets, ML methodologies aim to provide investors with insights into potential market trends and opportunities for financial growth. However, it's important to note that while ML algorithms have shown promise in forecasting company stock performance, they still face challenges in achieving precision.

This study delves into the realm of stock market prediction, shedding light on various research endeavours aimed at enhancing predictive capabilities. One notable technique highlighted in the study is the CNN-LSTM Neural Network model, which combines the strengths of convolutional neural networks (CNNs) and long short-term memory networks (LSTMs). This innovative approach holds the potential to improve data prediction accuracy and empower investors with valuable insights for informed decision-making in the realm of financial investments and strategic planning.

**Chapter 2**

**Related Work**

The literature on stock price forecasting is broadly categorized into regression, econometrics, and machine learning models. Each category offers unique approaches and methodologies to tackle the complexities of predicting stock market movements.

Regression-based investigations face challenges with high stock market volatility. These investigations often rely on techniques such as multivariate adaptive regression splines (MARS) and ordinary least square (OLS) regression. However, the inherent volatility of stock market data can pose limitations to the accuracy of these models.

Econometric approaches, which utilize time series models like autoregressive integrated moving average (ARIMA) and Granger Causality, are commonly employed in stock price forecasting. Yet, when dealing with extremely volatile data, these methods may not always provide sufficient predictive power.

In recent years, machine learning and deep learning models have emerged as promising alternatives for stock price prediction. Algorithms such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks have demonstrated the ability to effectively capture complex patterns in stock market data, offering more robust forecasting capabilities.

Hybrid models, which combine inputs from social web sources with deep learning and machine learning frameworks, have gained popularity for their potential to improve prediction accuracy. By blending the strengths of different approaches, hybrid models aim to overcome the limitations of individual methods and enhance overall forecasting performance.

Deep learning algorithms, in particular, excel in handling the volatility and randomness inherent in time series data. Their ability to learn intricate patterns and dependencies makes them well-suited for capturing the dynamic nature of stock market movements, thus offering valuable insights for investors and financial analysts alike.

**Chapter 3**

**Methodology**

The goal of this research is to create accurate prediction models for predicting stock prices on the Indian National Stock Exchange (NSE). By employing advanced deep learning methods, we will build strong frameworks that can produce precise stock price forecasts. Furthermore, our technique incorporates data from a variety of industries and indexes, such as IT, steel, automotive, banking, FMCG, NIFTY, and S&P 500. With the help of this all-inclusive method, we can perform in-depth market analysis and improve our comprehension of the numerous variables affecting stock price movement.

Apart from developing the model, we will meticulously assess its performance by employing metrics like accuracy and Root mean squared error (MSE). Additionally, we will contrast different approaches, examining variations in model topologies and training strategies to maximise computing efficiency and predictive accuracy. Our main objective is to continuously improve our models in order to increase their efficacy in practical applications.   
Yahoo Finance provided daily data for NSE trading days from January 1, 2018, to December 31, 2022. The data covered two indices (NIFTY and S&P 500), five industries (IT, steel, auto, banking, and FMCG). Variables like date, open, high, low, close, and volume are all included in the raw data.

The models underwent training from January 01, 2018, to December 31, 2020, historical data, and testing from January 01, 2021, to December 31, 2022, historical data. The models produced weekly predictions for the 'Close' value for the test dataset using a multi-step forecasting technique with walk-forward validation. To enable ongoing improvement and forecasting for the next week, real 'Close' values were added to the training dataset at the end of each week.

Top of Form

**Chapter 4**

**Deep Learning Models**

Our analysis employs two sophisticated deep learning architectures: Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network. These models are tailored to predict future stock market trends with precision and accuracy.

**Convolutional Neural Network Models:**

CNN\_UNI-5: This model adopts a univariate approach, exclusively leveraging past 'Close' values as input to forecast future 'Close' values, thus capturing trends based solely on historical closing prices.

CNN\_UNI-10: Extending the temporal scope, CNN\_UNI-10 utilizes a longer historical context (10 days) of past 'Close' values, enabling deeper insight into evolving market trends while maintaining the univariate focus.

CNN\_MULTI-10: In contrast to univariate models, CNN\_MULTI-10 embraces a multivariate perspective by incorporating all five variables (Open, High, Low, Close, Volume) as input. By considering a broader range of market indicators, this model provides a comprehensive outlook on future 'Close' values.

**Long and Short-Term Memory Network Models:**

LSTM\_UNI-5: Employing a univariate strategy akin to CNN\_UNI-5, this model capitalizes on past 'Close' values to predict future 'Close' values, thereby capturing nuanced temporal patterns in isolation.

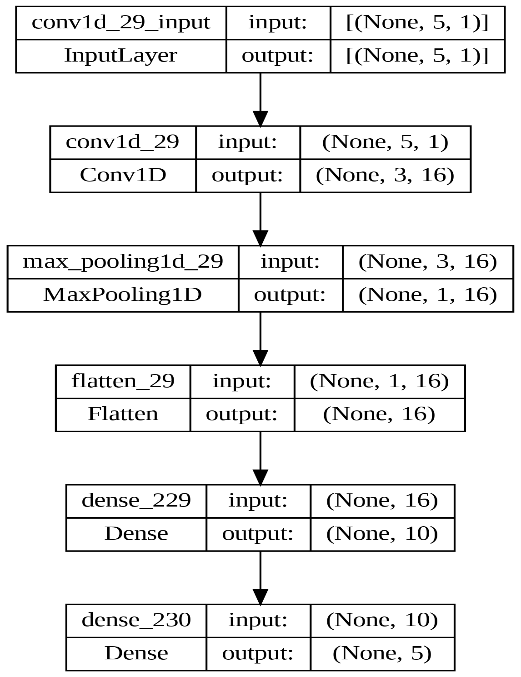
LSTM\_UNI-10: Analogous to LSTM\_UNI-5, LSTM\_UNI-10 extends its predictive horizon by incorporating a longer historical context (10 days) of past 'Close' values, enriching its forecasting capabilities.

LSTM\_MULTI-10: Departing from the univariate paradigm, LSTM\_MULTI-10 embraces a multivariate approach by integrating all five variables (Open, High, Low, Close, Volume) as input. This comprehensive model leverages a holistic view of market dynamics, enabling nuanced predictions of future 'Close' values.

**Convolutional Neural Network Models**

## CNN\_UNI-5

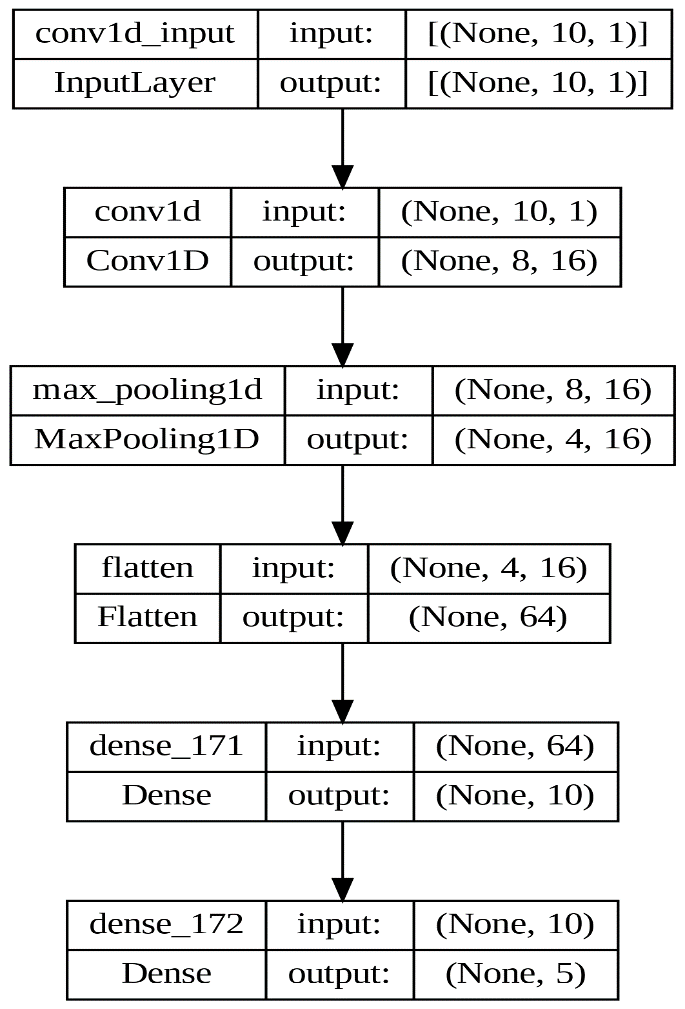
## The univariate model processes one-week (N = 5) preceding data, reading five days in three time-steps with a single convolution layer featuring 16 maps and a kernel size of three, extracting 16 features per reading. It utilizes max pooling of size two to reduce feature map size before flattening into a single vector. A fully-connected layer precedes the output layer, utilizing ReLU activation and ADAM optimizer, except for the final layer which uses sigmoid activation. Training involves a batch size of 4, 20 epochs, and MSE loss for RMSE computation.



Architecture for CNN\_UNI-5 model

## CNN\_UNI-10

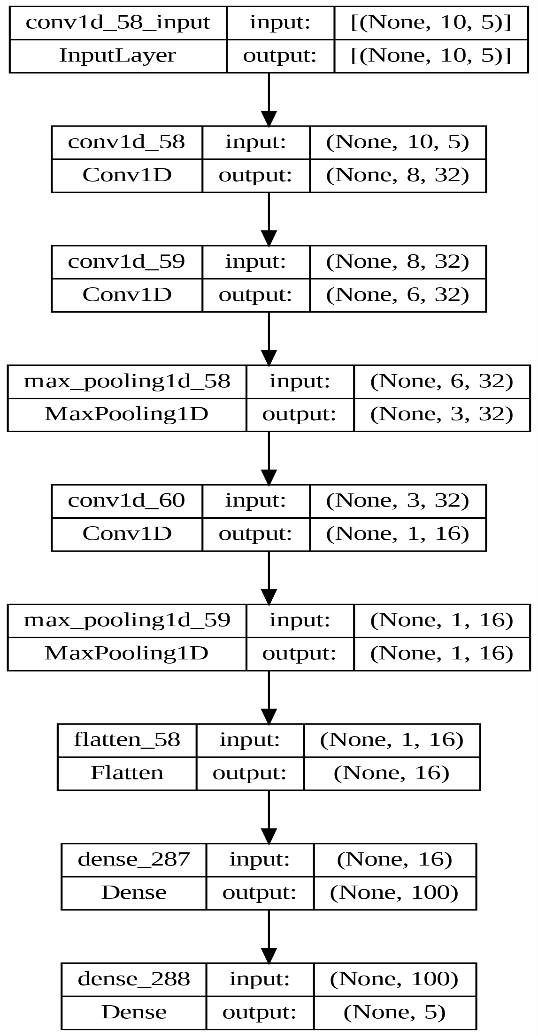
## The univariate model processes data from the last two weeks (N = 10) as input. The sequence of 10 days is read in three time-steps, repeated 16 times with a single convolution layer featuring 16 maps and a kernel size of 3. Max pooling of size 2 reduces the feature map size before flattening into a single vector of size 64. A fully-connected layer interprets the flattened vector before the output layer predicts the open values. ReLU activation and ADAM optimizer are utilized, except for the final layer which employs sigmoid activation.



Architecture for UNI\_CNN-10 model

## CNN\_MULTI-10

## For multivariate analysis, the model utilizes data from the previous two weeks (N = 10) with all five variables. The sequence of 10 days is read in three time-steps, repeated 16 times using a single convolution layer featuring 16 feature maps and a kernel size of 3. Max pooling of size 2 reduces feature map size before flattening into a single vector of size 64. A fully connected layer interprets the flattened vector before the output layer predicts open values. ReLU activation and ADAM optimizer are used, except for the final layer which employs sigmoid activation. Training involves a batch size of 16, 70 epochs, MSE loss, and computation of RMSE values.



Architecture for CNN\_MULTI-10 model

**Long and Short-Term Memory Models**

## LSTM\_UNI-5

## The univariate model utilizes data from the past week (N = 5) as input. The LSTM layer comprises 200 nodes, processing the input sequence of shape (5,1) to generate a 200-element vector, representing features extracted from the input values. Subsequently, the first dense layer consists of 100 nodes, receiving the 200-element vector from the preceding LSTM layer. The output layer, with nodes equal to the number of outputs, employs ReLU activation. MSE serves as the loss function and ADAM as the optimizer in the output layer. Training encompasses 20 epochs with a batch size of 4.

A diagram of a function

Description automatically generated

Architecture for UNI\_LSTM-5 model

## LSTM\_UNI-10

## The univariate model utilizes past week's data (N = 10) as input. The LSTM layer consists of 200 nodes, processing the input sequence of shape (10,1) to generate a 200-element vector representing extracted features. The first dense layer comprises 100 nodes, receiving the 200-element vector from the preceding LSTM layer. The output layer, with nodes equal to the number of outputs, employs ReLU activation. MSE serves as the loss function and ADAM as the optimizer in the output layer. Training encompasses 20 epochs with a batch size of 4.

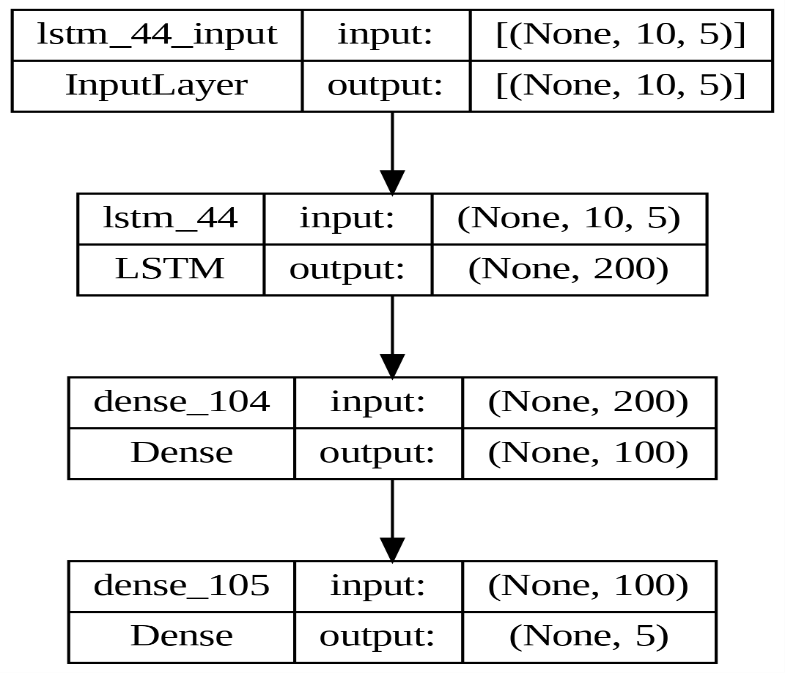
A diagram of a function

Description automatically generated

Architecture for LSTM\_UNI-10 model

## LSTM\_MULTI-10

## The multivariate model utilizes past week's data (N = 10) as input. The LSTM layer consists of 200 nodes, processing the input sequence of shape (10,5) to generate a 200-element vector representing extracted features. The first dense layer comprises 100 nodes, receiving the 200-element vector from the preceding LSTM layer. The output layer, with nodes equal to the number of outputs, employs ReLU activation. MSE serves as the loss function and ADAM as the optimizer in the output layer. Training encompasses 70 epochs with a batch size of 16.



Architecture for LSTM\_MULTI-10 model

**Chapter 5**

**Performance Results and Analysis**

**Result of Deep Learning Models**

**CNN\_UNI-5**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Time** | **RMSE/Mean** | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** | **RMSE** |
| **Tata Steel** | 16.5 | **8.9** | 5.3 | 5.2 | 5.2 | 6.5 | 6.6 | 5.8 |
| **Tata Consultancy Services** | 17.4 | 3.9 | 67.4 | 82.3 | 86.3 | 107.9 | 115.1 | 93.4 |
| **State Bank of India** | 21.6 | 6.6 | 16.0 | 19.7 | 23.5 | 25.1 | 25.5 | 22.3 |
| **NIFTY** | 21.3 | **2.8** | 313.5 | 333.5 | 359.6 | 403.7 | 452.0 | 375.8 |
| **S&P** | **16.4** | 3.2 | 91.0 | 100.4 | 102.8 | 116.3 | 134.0 | 109.9 |
| **Maruti Suzuki India Limited** | 21.9 | 4.0 | 250.6 | 267.6 | 285.8 | 300.2 | 339.8 | 290.4 |
| **Mahindra & Mahindra Limited** | **22.1** | 5.2 | 33.0 | 36.5 | 37.9 | 41.1 | 45.7 | 39.1 |
| **JSW Steel Limited** | 21.4 | 7.1 | 22.5 | 25.1 | 27.1 | 32.7 | 34.6 | 28.8 |
| **Infosys Limited** | 21.3 | 5.2 | 43.2 | 49.0 | 50.1 | 55.6 | 62.2 | 52.4 |
| **ITC Limited** | 16.6 | 3.4 | 5.9 | 6.2 | 7.2 | 8.2 | 9.2 | 7.5 |
| **Hindustan Unilever Ltd** | 20.3 | 3.8 | 62.3 | 67.6 | 71.1 | 85.7 | 88.0 | 75.6 |
| **HDFC Bank Limited** | 19.9 | 4.0 | 43.2 | 45.8 | 47.5 | 50.6 | 53.7 | 48.3 |
| **MEAN** | **19.7** | **4.917** | **79.5** | **86.6** | **92.0** | **102.8** | **113.9** | **95.8** |
| **MIN** | **16.4** | **2.8** | **5.3** | **5.2** | **5.2** | **6.5** | **6.6** | **5.8** |
| **MAX** | **22.1** | **8.9** | **313.5** | **333.5** | **359.6** | **403.7** | **452.0** | **375.8** |
| **SD** | **2.3** | **1.8** | **98.9** | **105.0** | **112.8** | **123.6** | **139.5** | **116.8** |

**CNN\_UNI-10**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Time** | **RMSE/Mean** | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** | **RMSE** |
| **Tata Steel** | 18.5 | **8.6** | 5.1 | 4.8 | 5.1 | 6.4 | 6.6 | 5.6 |
| **Tata Consultancy Services** | 18.1 | 4.7 | 98.7 | 92.4 | 101.4 | 131.0 | 129.9 | 111.9 |
| **State Bank of India** | 21.4 | 6.3 | 16.5 | 18.6 | 18.9 | 23.1 | 26.6 | 21.0 |
| **NIFTY** | 17.6 | **3.4** | 489.2 | 355.8 | 397.6 | 535.4 | 482.4 | 456.8 |
| **S&P** | 17.4 | 3.8 | 109.5 | 117.4 | 134.9 | 129.5 | 154.2 | 130.0 |
| **Maruti Suzuki India Limited** | 19.0 | 4.0 | 273.1 | 287.7 | 278.5 | 289.4 | 331.2 | 292.7 |
| **Mahindra & Mahindra Limited** | 21.6 | 5.2 | 34.6 | 33.8 | 36.9 | 39.3 | 46.6 | 38.5 |
| **JSW Steel Limited** | 18.0 | 6.8 | 21.9 | 27.5 | 24.3 | 30.1 | 32.4 | 27.5 |
| **Infosys Limited** | 21.1 | 5.5 | 47.4 | 51.4 | 56.8 | 57.4 | 62.3 | 55.3 |
| **ITC Limited** | **16.8** | 4.6 | 5.7 | 8.6 | 9.5 | 11.8 | 13.0 | 10.1 |
| **Hindustan Unilever Ltd** | 20.5 | 4.5 | 58.8 | 79.7 | 89.2 | 100.1 | 107.1 | 88.6 |
| **HDFC Bank Limited** | **21.7** | 4.0 | 43.7 | 48.5 | 45.2 | 53.9 | 52.1 | 48.8 |
| **MEAN** | **19.3** | **5.1** | **100.3** | **93.9** | **99.9** | **117.3** | **120.4** | **107.2** |
| **MIN** | **16.8** | **3.4** | **5.1** | **4.8** | **5.1** | **6.4** | **6.6** | **5.6** |
| **MAX** | **21.7** | **8.6** | **489.2** | **355.8** | **397.6** | **535.4** | **482.4** | **456.8** |
| **SD** | **1.8** | **1.5** | **142.8** | **112.7** | **120.6** | **153.3** | **145.0** | **135.5** |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Time** | **RMSE/Mean** | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** | **RMSE** |
| **Tata Steel** | **18.0** | 142.3 | 89.0 | 100.8 | 88.4 | 97.5 | 89.6 | 93.2 |
| **Tata Consultancy Services** | 20.9 | 98.9 | 2356.5 | 2998.7 | 2306.5 | 2130.1 | 1961.3 | 2377.0 |
| **State Bank of India** | 19.2 | 2317.1 | 4712.4 | 11074.2 | 8480.5 | 5144.4 | 7713.6 | 7781.1 |
| **NIFTY** | 20.8 | 3.7 | 409.8 | 430.0 | 499.0 | 491.2 | 606.8 | 492.2 |
| **S&P** | 21.3 | 14986.3 | 176982.9 | 397882.8 | 468818.3 | 335330.5 | 902040.7 | 516826.5 |
| **Maruti Suzuki India Limited** | 19.2 | 8.6 | 681.3 | 599.0 | 635.1 | 569.0 | 627.6 | 623.5 |
| **Mahindra & Mahindra Limited** | 23.1 | 115.4 | 972.4 | 649.6 | 656.2 | 990.4 | 970.8 | 862.7 |
| **JSW Steel Limited** | 18.7 | 838.0 | 2626.8 | 5036.4 | 2127.1 | 3765.6 | 2556.3 | 3391.3 |
| **Infosys Limited** | 21.7 | 126.8 | 1065.4 | 1330.5 | 1466.3 | 1208.7 | 1231.4 | 1267.5 |
| **ITC Limited** | **33.5** | 3221.4 | 8441.3 | 4249.4 | 12144.6 | 1338.4 | 2183.0 | 6976.6 |
| **Hindustan Unilever Ltd** | 22.6 | 95.0 | 1685.9 | 2305.4 | 1942.3 | 1732.9 | 1626.6 | 1875.0 |
| **HDFC Bank Limited** | 19.3 | 92.0 | 1108.9 | 1100.2 | 1058.7 | 1196.5 | 1138.1 | 1121.4 |
| **MEAN** | **21.5** | **1837.1** | **16761.1** | **35646.4** | **41685.3** | **29499.6** | **76895.5** | **45307.3** |
| **MIN** | **18.0** | **3.7** | **89.0** | **100.8** | **88.4** | **97.5** | **89.6** | **93.2** |
| **MAX** | **33.5** | **14986.3** | **176982.9** | **397882.8** | **468818.3** | **335330.5** | **902040.7** | **516826.5** |
| **SD** | **4.1** | **4270.5** | **50510.5** | **114116.0** | **134561.9** | **96322.5** | **259860.9** | **148511.0** |

**CNN\_MULTI-10**

**LSTM\_UNI-5**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Time** | **RMSE/Mean** | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** | **RMSE** |
| **Tata Steel** | **32** | **8.8** | 4.6 | 5.2 | 4.8 | 6.8 | 7 | 5.8 |
| **Tata Consultancy Services** | 32.5 | 4.2 | 62.9 | 111.7 | 101.4 | 104.7 | 120.3 | 102.1 |
| **State Bank of India** | 33.2 | 5.9 | 15.8 | 16 | 19.7 | 20.7 | 25.6 | 19.9 |
| **NIFTY** | 32.8 | 3.7 | 309.6 | 491.3 | 354.3 | 524.8 | 693.9 | 493.9 |
| **S&P** | 32.5 | **3** | 74.7 | 85.2 | 119 | 107.3 | 119.9 | 102.9 |
| **Maruti Suzuki India Limited** | **34.8** | 3.8 | 271.2 | 237.7 | 242.1 | 268.6 | 332 | 272.4 |
| **Mahindra & Mahindra Limited** | 32.5 | 6 | 33.6 | 45.6 | 50.3 | 46.8 | 47.7 | 45.2 |
| **JSW Steel Limited** | 32.7 | 7.6 | 24.8 | 21.9 | 37.5 | 33.7 | 33 | 30.7 |
| **Infosys Limited** | 34.4 | 4.7 | 37.9 | 42.4 | 44.1 | 52 | 57.1 | 47.2 |
| **ITC Limited** | 34.4 | 3.9 | 7.9 | 7.2 | 8.2 | 8.7 | 9.6 | 8.4 |
| **Hindustan Unilever Ltd** | 32.8 | 3.4 | 49.7 | 54.1 | 68.2 | 75.1 | 80.5 | 66.6 |
| **HDFC Bank Limited** | 33.8 | 3.5 | 36 | 40.2 | 42.8 | 45.3 | 48.9 | 42.8 |
| **MEAN** | **33.2** | **4.97** | **77.4** | **96.5** | **91** | **107.9** | **131.3** | **103.2** |
| **MIN** | **32** | **3** | **4.6** | **5.2** | **4.8** | **6.8** | **7** | **5.8** |
| **MAX** | **34.8** | **8.8** | **309.6** | **491.3** | **354.3** | **524.8** | **693.9** | **493.9** |
| **SD** | **0.9** | **1.8** | **102** | **139.8** | **105.3** | **149.1** | **197.7** | **142.5** |

**LSTM\_UNI-10**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Time** | **RMSE/Mean** | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** | **RMSE** |
| **Tata Steel** | 94.3 | **10.8** | 7.3 | 5.4 | 8.2 | 6.6 | 7.6 | 7.1 |
| **Tata Consultancy Services** | 92.9 | 6.4 | 147.1 | 167.9 | 129.5 | 149.8 | 168.1 | 153.2 |
| **State Bank of India** | **95.8** | 6.5 | 17.3 | 19.2 | 25.1 | 21.6 | 25.3 | 21.9 |
| **NIFTY** | 94.4 | 3.8 | 510.2 | 461.7 | 479.7 | 515.3 | 539.4 | 502.0 |
| **S&P** | 68.0 | **3.2** | 89.1 | 94.2 | 100.0 | 130.8 | 128.9 | 110.0 |
| **Maruti Suzuki India Limited** | 93.1 | 4.6 | 315.6 | 266.9 | 301.6 | 357.7 | 393.5 | 330.0 |
| **Mahindra & Mahindra Limited** | 94.0 | 7.0 | 46.0 | 48.1 | 46.9 | 57.1 | 60.4 | 52.0 |
| **JSW Steel Limited** | 95.8 | 10.7 | 37.6 | 39.6 | 42.4 | 46.4 | 49.1 | 43.2 |
| **Infosys Limited** | 65.8 | 5.7 | 49.4 | 48.2 | 60.3 | 59.8 | 66.1 | 57.2 |
| **ITC Limited** | **64.4** | 5.3 | 9.6 | 11.9 | 11.8 | 13.1 | 11.3 | 11.6 |
| **Hindustan Unilever Ltd** | 93.9 | 4.1 | 60.3 | 76.1 | 76.9 | 82.5 | 98.9 | 79.9 |
| **HDFC Bank Limited** | 69.9 | 5.0 | 49.8 | 61.4 | 49.9 | 59.4 | 78.2 | 60.6 |
| **MEAN** | **85.2** | **6.1** | **111.6** | **108.4** | **111.0** | **125.0** | **135.6** | **119.1** |
| **MIN** | **64.4** | **3.2** | **7.3** | **5.4** | **8.2** | **6.6** | **7.6** | **7.1** |
| **MAX** | **95.8** | **10.8** | **510.2** | **461.7** | **479.7** | **515.3** | **539.4** | **502.0** |
| **SD** | **13.5** | **2.5** | **151.5** | **133.8** | **140.4** | **155.4** | **164.4** | **149.2** |

**LSTM\_MULTI-10**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Time** | **RMSE/Mean** | **Mon** | **Tue** | **Wed** | **Thu** | **Fri** | **RMSE** |
| **Tata Steel** | **116.8** | **58807.7** | 52945.8 | 30668.1 | 41098.4 | 34895.8 | 27774.6 | 38528.0 |
| **Tata Consultancy Services** | 72.6 | 210.0 | 3566.4 | 6526.8 | 2608.9 | 5731.5 | 5690.7 | 5047.0 |
| **State Bank of India** | 74.0 | 5060.7 | 18689.4 | 10883.4 | 16358.4 | 23567.1 | 12378.4 | 16994.0 |
| **NIFTY** | **67.5** | **6.8** | 806.0 | 484.3 | 1228.4 | 1003.8 | 814.9 | 901.7 |
| **S&P** | 92.5 | 40555.1 | 1059661.0 | 1359717.1 | 1076852.0 | 2071776.1 | 1164894.7 | 1398610.6 |
| **Maruti Suzuki India Limited** | 92.3 | 13.5 | 931.0 | 859.1 | 976.5 | 1212.5 | 850.8 | 974.9 |
| **Mahindra & Mahindra Limited** | 92.9 | 1018.0 | 10218.6 | 9274.6 | 7051.9 | 2704.2 | 6471.7 | 7607.4 |
| **JSW Steel Limited** | 71.5 | 4799.4 | 13329.1 | 27478.9 | 14430.0 | 22425.1 | 15565.8 | 19422.5 |
| **Infosys Limited** | 68.8 | 1120.3 | 9036.5 | 3496.0 | 12383.7 | 17070.5 | 9395.9 | 11197.5 |
| **ITC Limited** | 71.0 | 3644.3 | 12253.2 | 9595.8 | 4265.3 | 4531.6 | 5522.0 | 7892.3 |
| **Hindustan Unilever Ltd** | 75.3 | 61.4 | 1060.7 | 944.5 | 1023.1 | 1205.1 | 1679.1 | 1211.3 |
| **HDFC Bank Limited** | 93.0 | 189.3 | 2804.0 | 2748.9 | 2270.8 | 1447.9 | 1992.5 | 2308.3 |
| **MEAN** | **82.3** | **9623.9** | **98775.1** | **121889.8** | **98379.0** | **182297.6** | **104419.3** | **125891.3** |
| **MIN** | **67.5** | **6.8** | **806.0** | **484.3** | **976.5** | **1003.8** | **814.9** | **901.7** |
| **MAX** | **116.8** | **58807.7** | **1059661.0** | **1359717.1** | **1076852.0** | **2071776.1** | **1164894.7** | **1398610.6** |
| **SD** | **15.0** | **19202.1** | **302939.0** | **389941.9** | **308349.1** | **595139.5** | **334053.2** | **400948.7** |

**Summary of Performance Results of the Models**

We observe that CNN -2 model – a univariate CNN model with past two week’s data as its input – turns out to be the fastest model in execution.

The model CNN -1 – a univariate CNN model with past one week’s data as its input – turned out to be the most accurate model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Execution Time** | | | **RMSE/MEAN** | | |
| **Rank** | **Model** | **Value** | **Rank** | **Model** | **Value** |
| 1 | CNN\_UNI-10 | 19.3 | 1 | CNN\_UNI-5 | 4.91 |
| 2 | CNN\_UNI-5 | 19.7 | 2 | LSTM\_UNI-5 | 4.97 |
| 3 | CNN\_MULTI-10 | 21.5 | 3 | CNN\_UNI-10 | 5.1 |
| 4 | LSTM\_UNI-5 | 33.2 | 4 | LSTM\_UNI-10 | 6.1 |
| 5 | LSTM\_MULTI-10 | 82.3 | 5 | CNN\_MULTI-10 | 1837.1 |
| 6 | LSTM\_UNI-10 | 85.2 | 6 | LSTM\_MULTI-10 | 9623.9 |

**List of Figures**

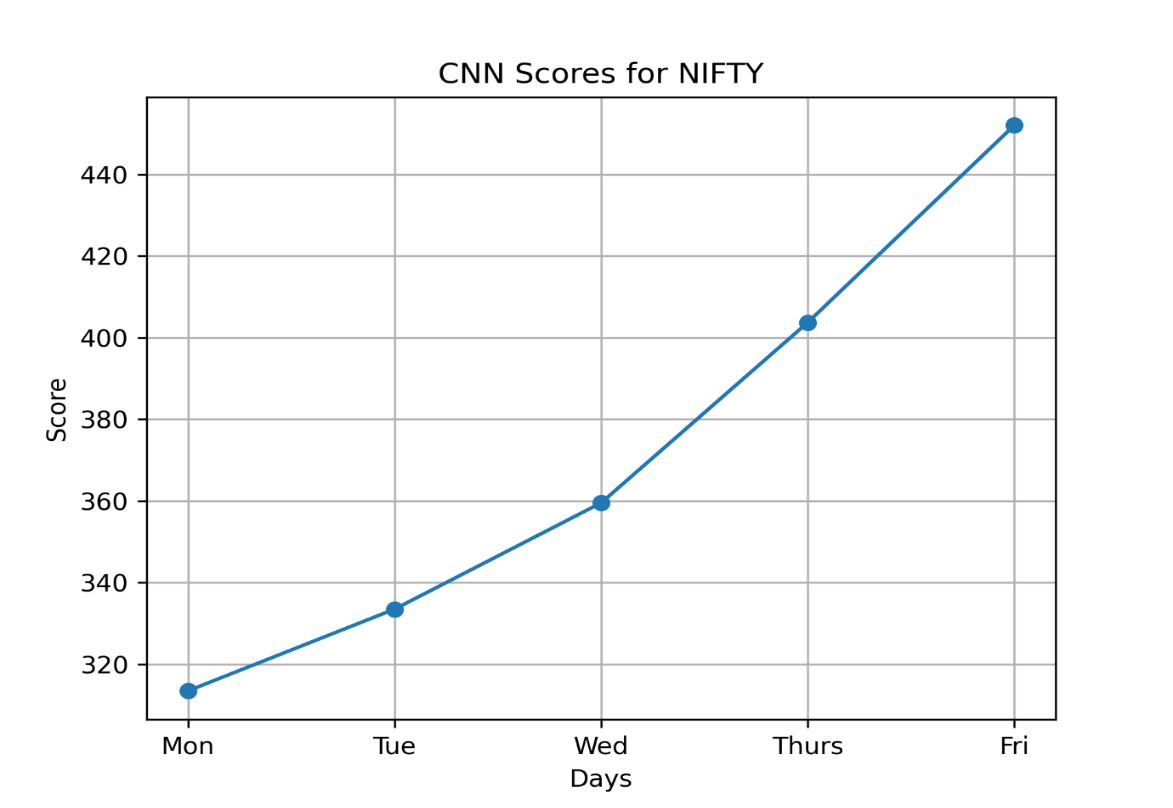


Fig 1(a) Variation of RMSE with different days in week for Nifty50 (CNN\_UNI-5)

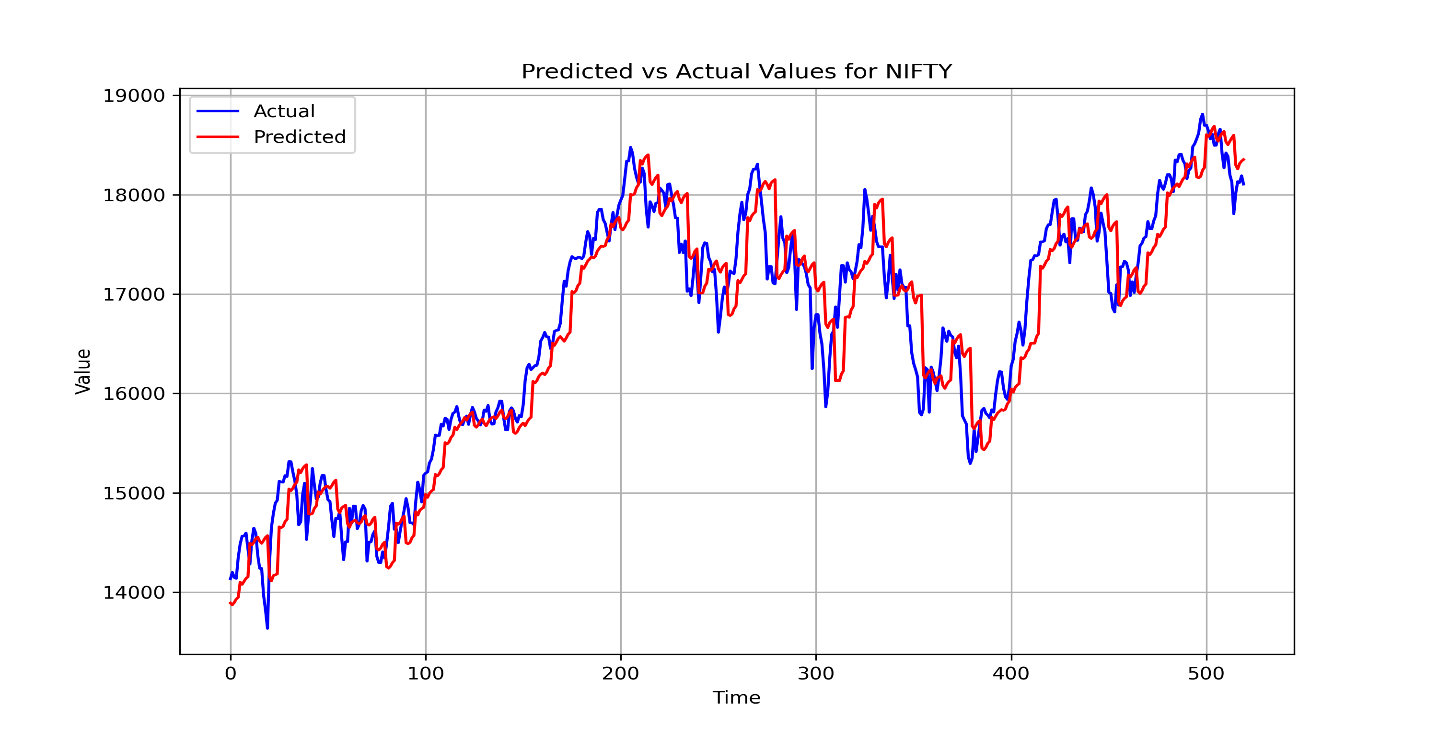


Fig 1(b) Actual vs Predicted graph for Nifty50 (CNN\_UNI-5)

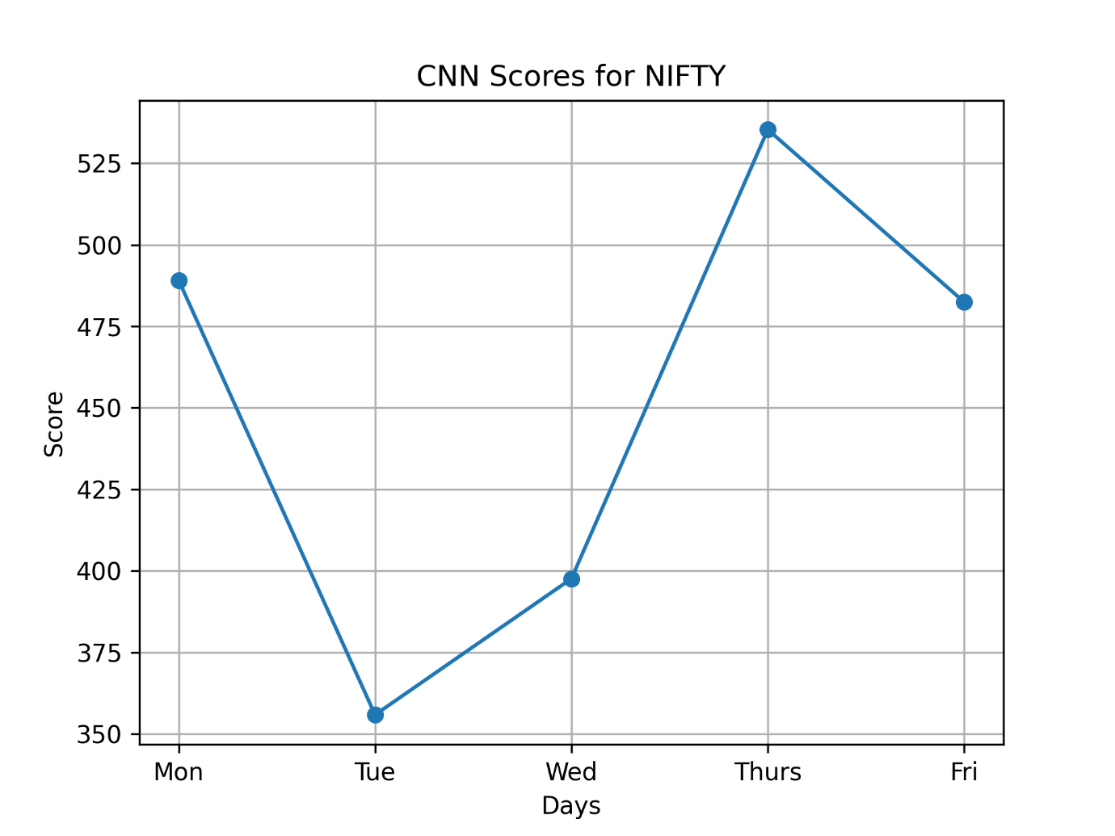
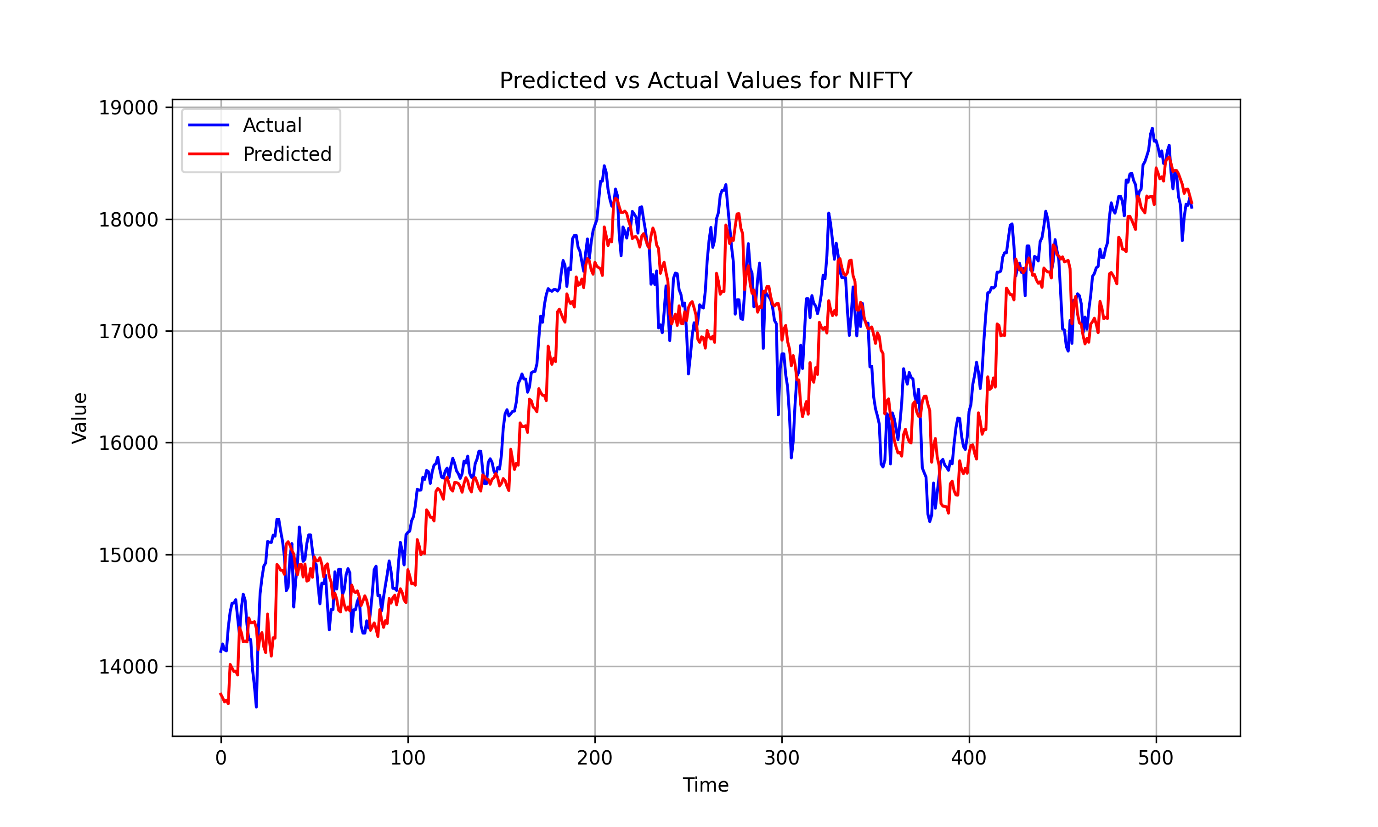


Fig 2(a) Variation of RMSE with different days in week for Nifty50 (CNN\_UNI-10)

Fig 2(b) Actual vs Predicted graph for Nifty50 (CNN\_UNI-10)

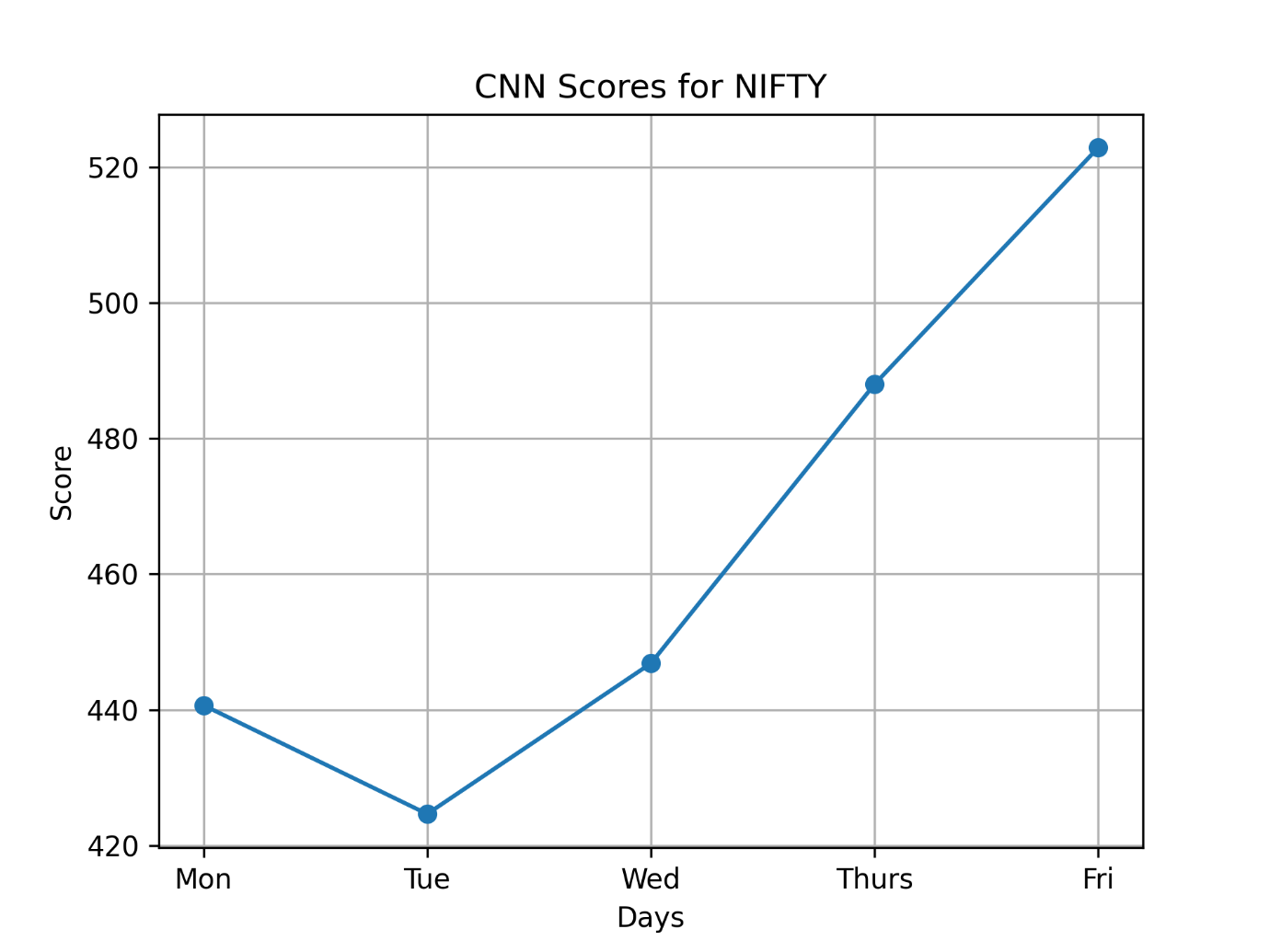
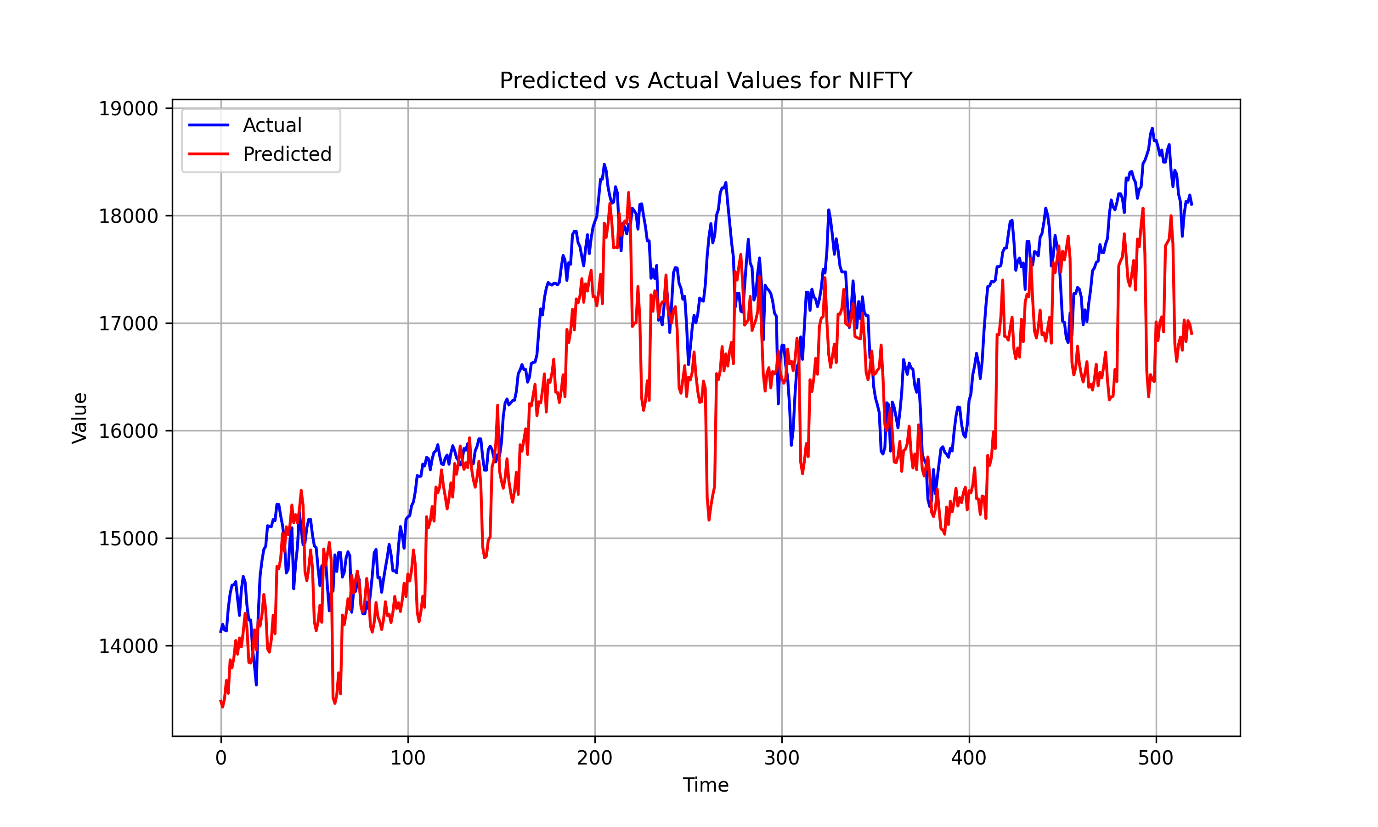


Fig 3(a) Variation of RMSE with different days in week for Nifty50 (CNN\_MULTI-10)

Fig 3(a) Actual vs Predicted graph for Nifty50 (CNN\_MULTI-10)

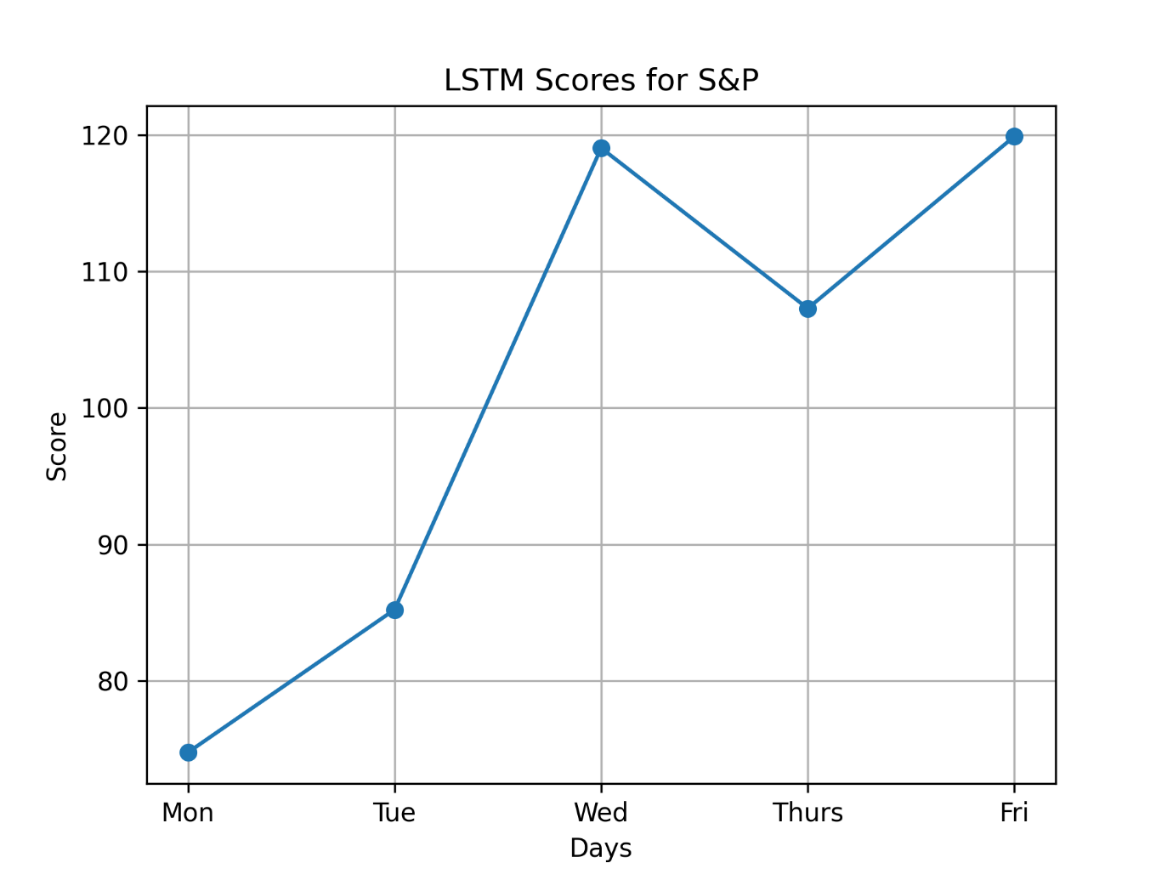


Fig 4(a) Variation of RMSE with different days in week for S&P500 (LSTM\_UNI-5)

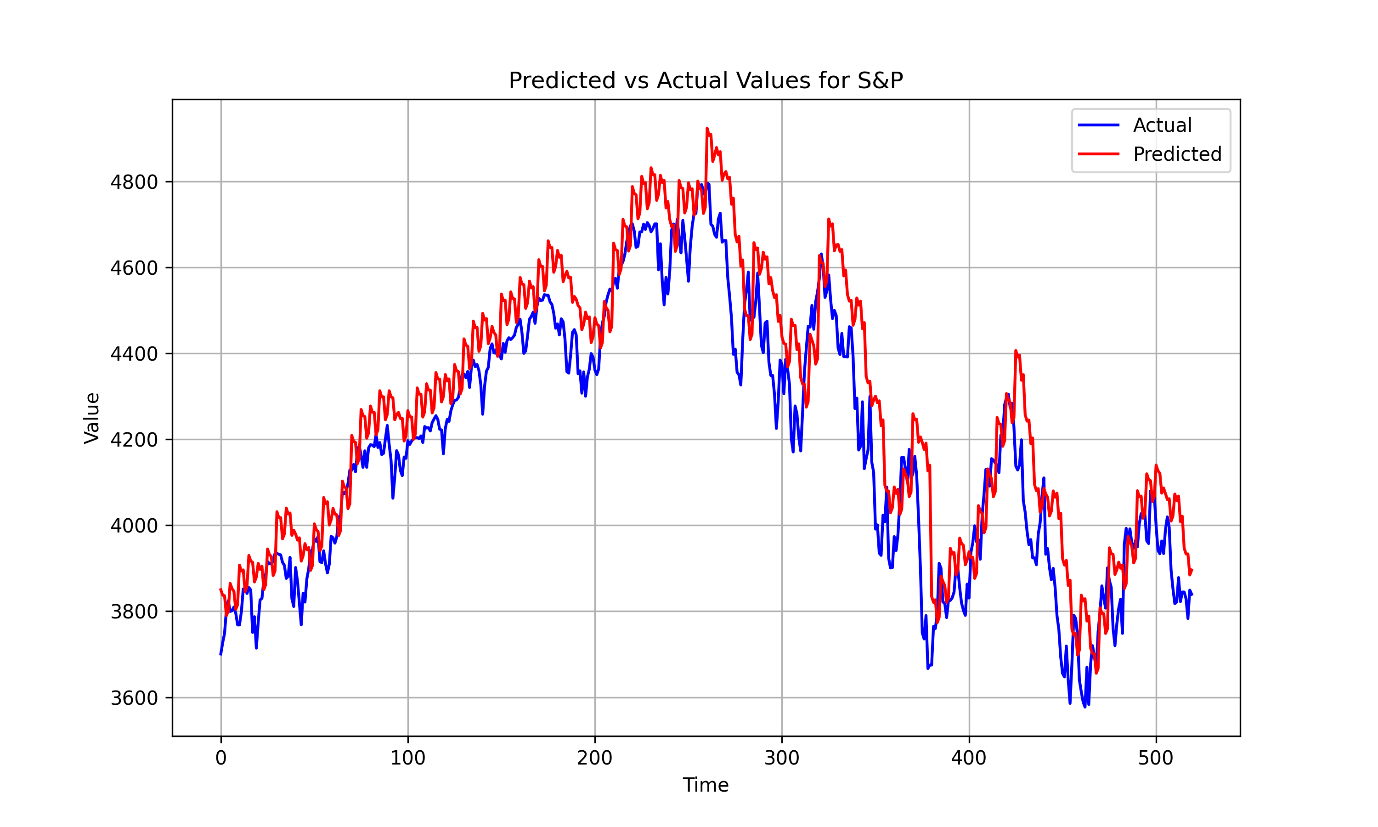


Fig 4(b) Actual vs Predicted graph for S&P500 (LSTM\_UNI-5)

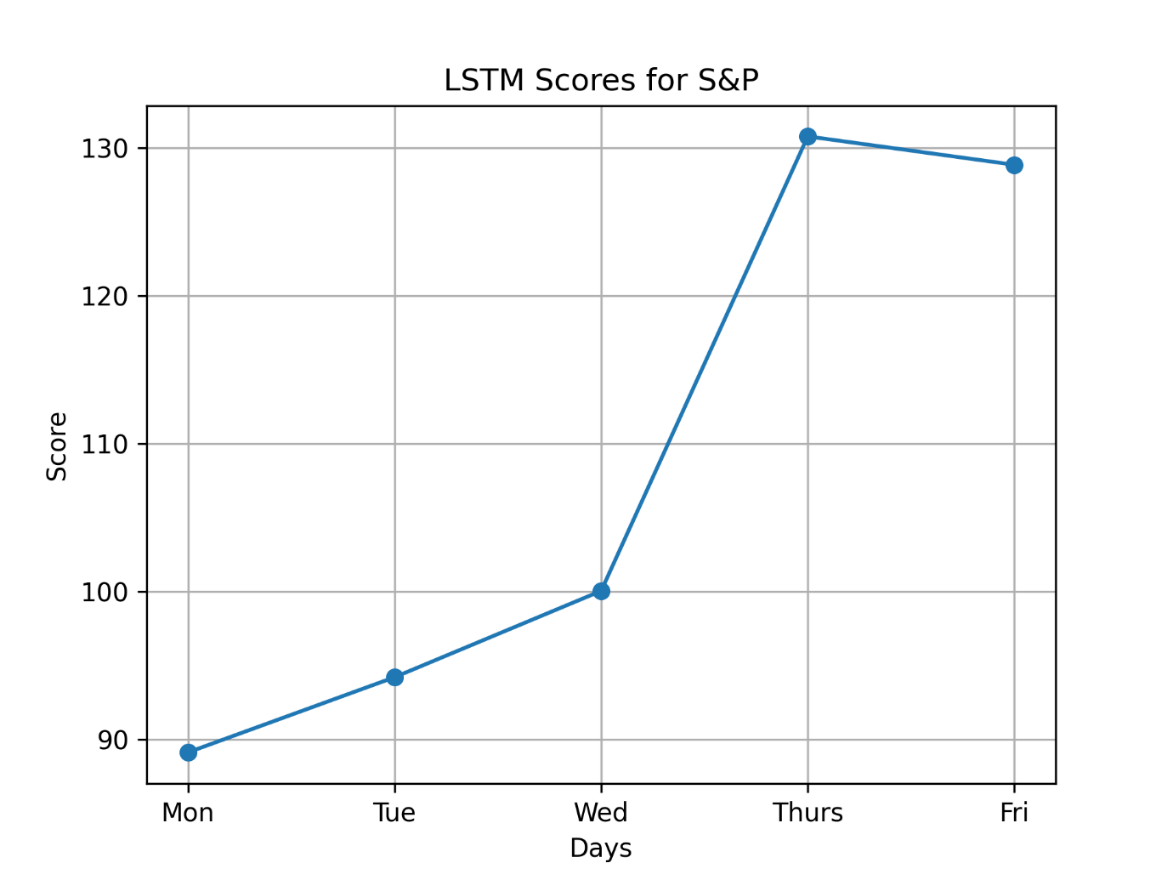
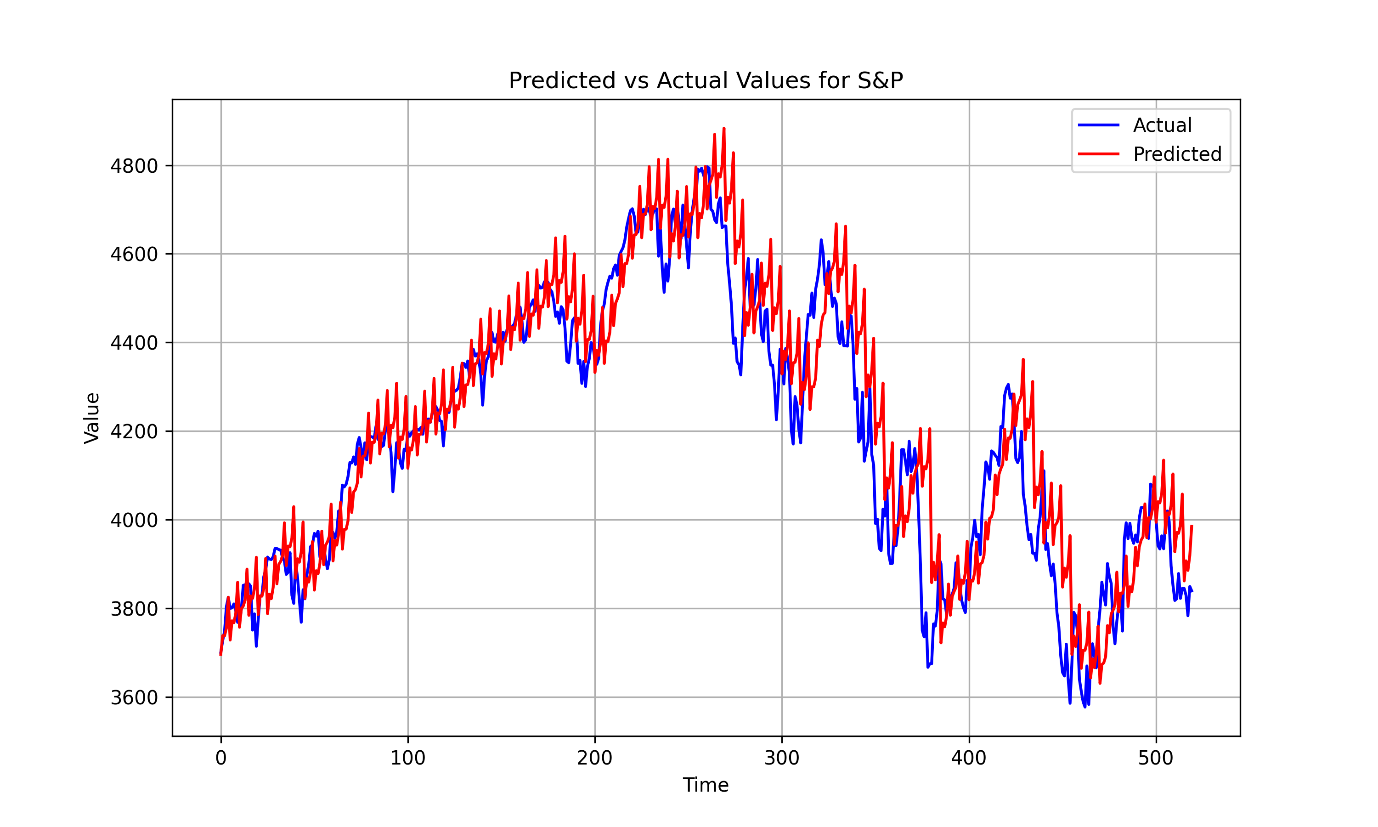


Fig 5(a) Variation of RMSE with different days in week for S&P500 (LSTM\_UNI-10)



### Fig 5(b)Actual vs Predicted graph for S&P500 (LSTM\_UNI-10)

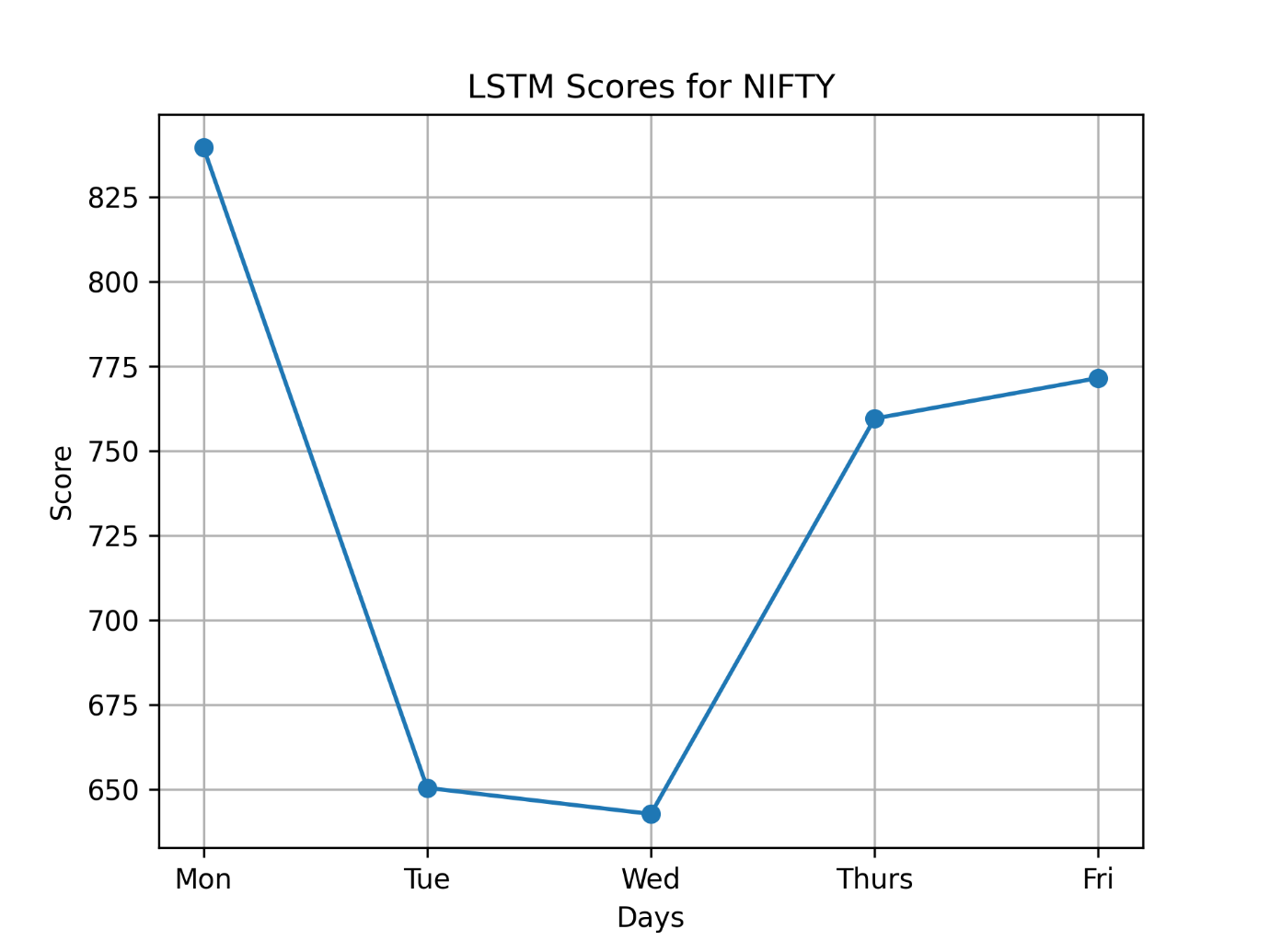
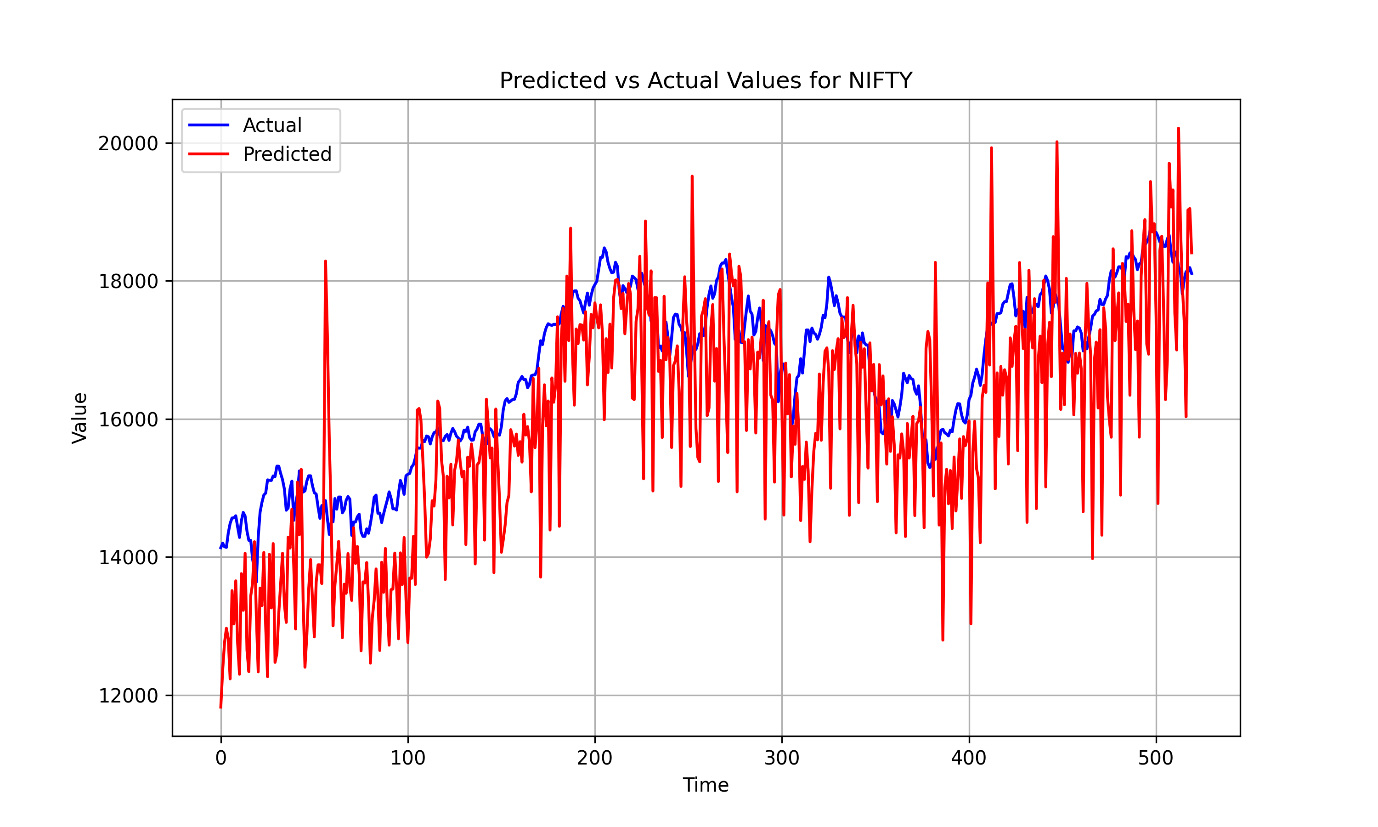


Fig 6(a) Variation of RMSE with different days in week for Nifty50 (LSTM\_MULTI-10)



### Fig 6(b) Actual vs Predicted graph for Nifty50 (LSTM\_MULTI-10)

**CHAPTER 6**

**Conclusion and Future Work**

Our study focused on developing deep learning-driven regression models for forecasting daily stock prices on the National Stock Exchange (NSE) of India. Six models were constructed and tested, including two univariate CNN models, two univariate LSTM models, and one multivariate model combining CNN and LSTM. Using data from January 1, 2021, to December 31, 2022, we achieved exceptional accuracy levels across all models. Particularly, the CNN model utilizing data from the previous one and two weeks emerged as the top performer in terms of accuracy and speed.

Looking ahead, we plan to explore the integration of generative adversarial networks (GANs) to potentially enhance prediction accuracy. Additionally, we aim to investigate the incorporation of alternative data sources, such as sentiment analysis from news articles or social media data, to improve prediction robustness. Furthermore, we will consider the inclusion of exogenous variables, such as macroeconomic indicators, to refine forecasting models. Additionally, we aim to explore various hybrid LSTM models and integrate LSTM, CNN, and GAN models to further advance stock price prediction research.

Through these future directions, we seek to contribute to the development of more accurate and reliable stock price forecasting models, aiding decision-making in financial markets.

**Chapter 7**

**References**

Damrongsakmethee,T. and Neagoe,V.-E. "Stock Market Prediction Using a Deep Learning Approach," 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, 2020, pp. 1-6, DOI: 10.1109/ECAI50035.2020.9223142.

Mehtab, S. and Sen, J. "Stock Price Prediction Using CNN and LSTM-Based Deep Learning Models," 2020 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 2020, pp. 447-453, doi: 10.1109/DASA51403.2020.9317207.

Maiti, A. and Shetty D, P. "Indian Stock Market Prediction using Deep Learning," 2020 IEEE Region 10 Conference (TENCON), Osaka, Japan, 2020, pp. 1215-1220, doi: 10.1109/TENCON50793.2020.9293712.

P. S and V. P. R, "Stock Price Prediction using Machine Learning and Deep Learning," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Hassan, India, 2021, pp. 660-664, DOI:10.1109/MysuruCon52639.2021.9641664.

P. Praveen, S. Shravani, R. Srija and M. Tajuddin, "A Model to Stock Price Prediction using Deep Learning," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 242-252, doi: 10.1109/ICSCSS57650.2023.10169558.

Biswas, M., Shome, A., Islam, M. A., Nova, A. J. and Ahmed, S. "Predicting Stock Market Price: A Logical Strategy using Deep Learning," 2021 IEEE 11th IEEE Symposium on Computer Applications &Industrial Electronics (ISCAIE), Penang,Malaysia, 2021, pp.218-223, doi:10.1109/ISCAIE51753.2021.9431817

Hossain, M. A., Karim, R., Thulasiram, R., Bruce, N. D. B. and Y. Wang, "Hybrid Deep Learning Model for Stock Price Prediction," 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 2018, pp. 1837-1844, doi: 10.1109/SSCI.2018.8628641.

Khan, S., Rabbani, M. R., Bashar, A. and Kamal, M., "Stock Price Forecasting Using Deep Learning Model," 2021 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 2021, pp. 215-219, doi: 10.1109/DASA53625.2021.9682319.

Parida, N., Balabantaray, B. K., Nayak, R. and Rout, J. K., "A Deep Learning based Approach to Stock Market Price Prediction using Technical indicators," 2023 International Conference on Advances in Intelligent Computing and Applications (AICAPS), Kochi, India, 2023, pp. 1-7, doi: 10.1109/AICAPS57044.2023.10074445.

Akita, R., Yoshihara, A., Matsubara, T., and Uehara, K., "Deep learning for stock prediction using numerical and textual information," 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, Japan, 2016, pp. 1-6, doi: 10.1109/ICIS.2016.7550882

Singh, T., Kalra, R., Mishra, S. et al. "An efficient real-time stock prediction exploiting incremental learning and deep learning". Evolving Systems 14, 919–937 (2023). https://doi.org/10.1007/s12530-022-09481-x.

Li, Y., Pan, Y. "A novel ensemble deep learning model for stock prediction based on stock prices and news". Int J Data Sci Anal 13, 139–149 (2022). <https://doi.org/10.1007/s41060-021-00279-9>.

Yu, P., Yan, X., "Stock price prediction based on deep neural networks". Neural Comput & Applic 32, 1609– 1628 (2020). <https://doi.org/10.1007/s00521-019-04212-x>.

Billah, M.M., Sultana, A., Bhuiyan, F. et al., "Stock price prediction: comparison of different moving average techniques using deep learning model". Neural Comput & Applic (2024). https://doi.org/10.1007/s00521- 023-09369-0.

Shah, A., Gor, M., Sagar, M. et al., "A stock market trading framework based on deep learning architectures". Multimed Tools Appl 81, 14153–14171 (2022). <https://doi.org/10.1007/s11042-022-12328-x>.

Ding, G., Qin, L., "Study on the prediction of stock price based on the associated network model of LSTM". Int. J. Mach. Learn. & Cyber. 11, 1307–1317 (2020). <https://doi.org/10.1007/s13042-019-01041-1>.

Bandhu, K.C., Litoriya, R., Jain, A. et al., "An improved technique for stock price prediction on real-time exploiting stream processing and deep learning". Multimed Tools Appl (2023). <https://doi.org/10.1007/s11042-023-17130-x>.

Swathi, T., Kasiviswanath, N. & Rao, A.A. "An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis". Appl Intell 52, 13675–13688 (2022). https://doi.org/10.1007/s10489-022- 03175-2.

Olorunnimbe, K., Viktor, H. "Deep learning in the stock market—a systematic survey of practice, backtesting, and applications". Artif Intell Rev 56, 2057–2109 (2023). https://doi.org/10.1007/s10462-022- 10226-0.

Mehtab S. and Sen, J. (2022) “Analysis and forecasting of financial time series using CNN and LSTM-based deep learning models,” in: Sahoo, J. P., Tripathy, A. K., Mohanty, M., Li, K. C., Nayak, A, K. (eds), Advances in Distributed Computing and Machine Learning, Lecture Notes in Networks and Systems, Vol 302, pp. 405-423, Springer, Singapore, doi: doi.org/10.1007/978-981-16-4807-6\_39.