Stock Market Prediction During Covid Using LSTM

ANANYA SINGH
21MCED01

INSTITUTE OF TECHNOLOGY
NAAC ACCREDITED 'A' GRADE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING INSTITUTE OF TECHNOLOGY NIRMA UNIVERSITY AHMEDABAD-382481

May 2023

Stock Market Prediction During Covid Using LSTM

Major Project

Submitted in partial fulfillment of the

requirements for the degree of

Master of Technology in Computer Science and Engineering (Data Science)

Submitted By

ANANYA SINGH

21MCED01

Guided By

Dr. Swati Jain

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING INSTITUTE OF TECHNOLOGY NIRMA UNIVERSITY AHMEDABAD-382481

May 2023

Certificate

This is to certify that the major project entitled **"Stock Market Prediction During COVID Using LSTM"** submitted by **Ananya Singh (Roll No: 21MCED01)**, towards the partial fulfillment of the requirements for the award of the degree of Master of Technology in Computer Science and Engineering (Data Science) of Nirma University, Ahmedabad, is the record of work carried out by her under my supervision and guidance. In my opinion, the submitted work has reached the level required for being accepted for examination. The results embodied in this major project, to the best of my knowledge, haven't been submitted to any other university or institution for the award of any degree or diploma.

Dr. Swati Jain
Guide & Associate Professor,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. Swati Jain
Associate Professor,
Coordinator M.Tech - CSE (Specialization)
Institute of Technology,
Nirma University, Ahmedabad

Dr. Madhuri Bhavsar
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad

Dr R.N. Patel
Director,
Institute of Technology,
Nirma University, Ahmedabad

Statement of Originality

I, **Ananya Singh**, Roll. No. **21MCED01**, give an undertaking that the Major Project entitled "**Stock Market Prediction During COVID Using LSTM**" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering (Data Science)** of the Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

____

Signature of Student

Date: 23rd May 2023

Place: Ahmedabad

Endorsed by

Dr. Swati Jain

(Signature of Guide)

Acknowledgments

It gives me immense pleasure in expressing thanks and profound gratitude to **Dr. Swati Jain**, Associate Professor, Computer Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his valuable guidance and continual encouragement throughout this work. The appreciation and continual support she has imparted have been a great motivation to me in reaching a higher goal. Her guidance has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.

A special thank you is expressed wholeheartedly to **Dr. R.N. Patel**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout the course of this work.

I would also thank the Institution, and all faculty members of the Computer Engineering Department, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

- Ananya Singh
21MCED01

ABSTRACT

Researchers have always faced significant challenges when attempting to predict stock market movements in the field of computation. This difficulty arises due to the influential nature of stock prices, which are impacted by a multitude of factors, including both tangible and intangible elements such as physical and physiological factors, rational and irrational behavior, geopolitical stability, and investor sentiment. Successful investors strive to anticipate future market conditions for profitable investments. In light of these considerations, we propose the utilization of a stacked long-short-term-memory (LSTM) model to forecast the closing index of stock prices during the uncertain period of the pandemic. The model's performance is evaluated using the root mean square error (RMSE) as a performance metric. Our objective is to optimize the model to enhance prediction accuracy and achieve superior stock market forecasting. The dataset employed in this study encompasses stock market data from NIFTY 50 (India), DAX 40 Index (Germany), FTSE Indices (UK), and S&P 500 Indices (USA), spanning across four sectors: Banking, Information Technology, Healthcare, and Retail. The duration of the dataset ranges from January 30, 2020, to March 31, 2022. The primary goal of this research paper is to analyze historical data and extract future patterns and insights.

CONTENTS

List of Figures

List of Tables

CHAPTER 1 : Introduction

1.1 General Introduction

The outbreak of COVID-19, caused by the SARS-CoV-2 virus, had a devastating impact on the global economy, leading to its declaration as a pandemic by the World Health Organization on March 11, 2020 [1]. India, with the second-largest population in the world, reported its first case on January 27, 2020 [2]. As the number of cases increased, measures like lockdowns, mobility restrictions, and healthcare challenges adversely affected the economy. Additionally, the influence of culture played a significant role in shaping the volatility and magnitude of returns [5], which can be observed through fluctuations in stock market indices.

The task of predicting future outcomes, which is highly sought after, is considered one of the most challenging aspects in the field of science and technology. Many researchers and experts have focused on studying financial market forecasting due to the crucial role financial markets play in the overall economy of a country. Time series data analysis is key to forecasting prices in financial markets, as it examines the characteristics of historical data to make predictions. Time series forecasting involves evaluating the past to predict the future and has implications in various sectors such as engineering, economics, science, and finance. Therefore, developing models that appropriately fit the underlying time series is essential.

Traditionally, two fundamental methodologies have been used for estimating the stock prices of organizations. To forecast future prices, technical analysts analyze historical stock prices, including opening and closing prices, volume, and adjacent close values. On the other hand, economic analysts rely on external factors such as corporate profiles, market conditions, political changes, economic considerations, and information from financial news articles and social media to conduct qualitative analysis [12]. Forecasting methodologies can be categorized into linear models (AR, MA, ARIMA, ARMA) and non-linear models (ARCH, GARCH, and Neural Network).

Feedforward neural networks adjust connection weights to minimize the gap between predicted and actual outcomes. Recurrent neural networks, which have loops that maintain information persistence, have been developed to process sequential data [13]. Among the recurrent neural network architectures, Long Short-Term Memory (LSTM) was introduced by Sepp Hochreiter et al. LSTM is known for its effectiveness in handling sequential data, providing the network with the ability to retain or forget information and assign different levels of importance to alter the outcome. In this study, we employ a stacked LSTM model to forecast stock prices for specific sectors based on historical price data prices.

1.2 Objective of Study

In order to forecast the closing index of stock prices during the challenging period of the pandemic, we propose the utilization of a stacked long-short-term memory (LSTM) model. The model's performance is evaluated using the root mean square error (RSME) as a performance metric. Our objective is to optimize the model to enhance prediction accuracy and achieve superior stock forecasting capabilities. To conduct our analysis, we consider a dataset comprising stock market data from NIFTY 50 (India), DAX 40 Index (Germany), FTSE Indices (UK), and S&P 500 Indices (USA) across four sectors: Bank, Information Technology, Healthcare, and Retail. The dataset covers the period from January 30, 2020, to March 31, 2022. The primary aim of this research is to leverage historical data to identify future patterns and gain valuable insights into stock market behavior.

1.3 Scope of Work

This paper presents several key contributions:

•Utilizing closing stock prices as input, we aim to forecast future prices of the target stock.

•We introduce a stacked LSTM model, a type of recurrent neural network that effectively captures the order dependence in sequence prediction tasks. This model leverages the cell state as a transportation highway, facilitating the transmission of relevant data throughout the sequence and thereby enhancing the accuracy of predictions.

•The performance of our proposed model is evaluated using the root mean square error (RMSE) metric. By comparing the predicted values with the observed values, we assess the effectiveness of the stacked LSTM model in accurately forecasting the closing price of stock data.

CHAPTER 2 : Literature survey

PAPER REFERENCE	TIME PERIOD	SOURCE	INDEX PARAMETER(S)	LEARNI NG RATE METHOD	MODEL
[6]	Jan 1, 2011 to Dec 31, 2016	NIFTY 50	Open/Close/High/ Low	RMSprop	LSTM
[7]	April 25, 2013 to May 15, 2017	CSI-300	Open	RMSProp learning rate 0.001	Multi Input LSTM
[8]	-	BSE (Bombay Stock Exchange	Close	-	LSTM
[9]	0:30 pm on April 3, 2017 to 2:15 pm on May 2, 2017 (41,250 minutes)	S&P 500	Closing Price	SGD	CNN
[10]	Jan 1, 2015 to Feb 14, 2017	Shanghai - Shenzhen 300 Stock Index (HS300)	Opening Price, Closing Price, Volume, Amplitude, Difference, Lowest Price, Highest Price, Volume change, Volume limit, Price change, Price limit	-	RNN Boost
[11]	-	Taiwan Stock Exchange Corporation (TWSE)	Opening Price, Closing Price, D Value, Upper BBands, MA20 and Lower BBands, Highest Price, Lowest Price, RSI (5 Day), RSI (10day), K Value, Transaction, Trade Volume	Adaptive Learning Rate Method	LSTM

[12]	Trade dates between 1998-2018	S&P 500 Index, Microsoft Corporation from NASDAQ, Shanghai Composite Index in China, Ping An Insurance Company of China (PAICC), IBM from NYSE	Opening, Low, Closing, High, Volume.	-	LSTM with GAN
[13]	Jan 2, 2008 to Nov 27, 2018	S&P 500 index	Closing Price	ReLU activation function	Multiple pipeline CNN and Bidirectional LSTM
[14]	May 4, 2009 to May 4, 2019	Five companies-Goldman Sachs, Pfizer, JP Morgan Chase and Co., Johnson and Johnson, Nike from Yahoo Finance	Closing Price	-	ANN and Random Forest

Table 1: Different stock market prediction models have been explored.

Machine learning has significantly expanded the realm of predictions, particularly in the field of stock market forecasting. Researchers from various disciplines, such as computer science, finance, and other communities, have been actively exploring different methodologies and models to predict future stock prices by analyzing patterns and trends from historical data. The advancement of machine learning has enabled the analysis of large datasets for more accurate forecasting.

Although convolutional neural networks (CNNs) were originally developed for image processing, their applications have expanded to include time series data. For instance, Hyun Sik Sim et al. [9] proposed a CNN-based prediction technique that transformed technical parameters into images and employed time series functions for analysis. Jithin Eapen et al. [13] introduced a bi-directional CNN model that demonstrated improved predictions compared to support vector machines (SVM) and single pipeline models, reducing the potential effects of overfitting. In a novel approach, Min Wen et al. [16] reconstructed time series data using motifs and utilized CNNs, highlighting their efficiency over sequential models in terms of computational complexity. Lina Ni et al. [17] proposed a unique method that considered the spatio-temporal characteristics of Forex time series data, showing superior performance of their model compared to LSTM and CNN models.These studies illustrate the versatility and effectiveness of CNNs in analyzing time series data for stock market prediction, showcasing their potential for improving computational efficiency and enhancing forecasting accuracy.

Deep learning models, which encompass a multitude of learning algorithms, have contributed to advancements in analyzing complex prediction models. Xiao Ding et al. [15] introduced a deep-learning model specifically designed for forecasting event-driven stock markets. Their approach utilized financial news as indices, and their findings indicated that more accurate predictions could be achieved by focusing on news titles rather than the news content itself.

In a comparative study, Mehar Vijh et al. [14] examined stock market price prediction using artificial neural networks (ANN) and Random Forest machine learning models. Based on metrics such as MAPE, MBE, and RMSE, the research concluded that ANN outperformed Random Forest in predicting stocks for the following day.

Recurrent neural networks (RNNs) have garnered significant attention for market prediction tasks. Weiling Chen et al. [10] proposed a hybrid model called RNN Boost, which combines RNN with Adaboost. Their research incorporated the use of news content from an online social media platform, Sina Weibo (China's largest social media platform), instead of traditional news media. By leveraging these online platforms, which excel in analyzing public sentiments, the RNN Boost model demonstrated promising results for stock market prediction. These studies highlight the effectiveness of deep learning models, including RNNs and hybrid approaches, in analyzing complex prediction tasks for stock market forecasting. Furthermore, they showcase the advantages of leveraging non-traditional data sources, such as news titles and online social media platforms, to enhance the accuracy of predictions.

LSTM, a type of recurrent neural network (RNN) well-suited for sequential data prediction, is widely used either in its basic form or in combination with other models. Murtaza Roondiwala et al. [6] proposed using LSTM to forecast stock prices based on opening, closing, high, and low prices as parameters. Through various simulations with different parameter combinations, the study concluded that utilizing all four parameters together yielded the best results.

In a comparison study on the iShares MSCI United Kingdom dataset, Mahla Nikou et al. [18] demonstrated that LSTM outperformed support vector regression, neural network, and random forest methods in minimizing prediction errors. Bo-Sheng Lin et al. [11] employed LSTM to predict the closing prices of the top 10 companies listed in the TWSE. The majority of the results, except for Midland Holdings Limited, showed acceptable performance.

Lakshminarayanan et al. [19] proposed an LSTM model considering two variants of datasets: one comprising the standard Dow Jones Index (DJI) stock prices, and the other combining the previous dataset with crude oil and gold prices. The findings indicated that the latter model performed better, shedding light on the impact of crude oil and gold prices on stock price predictions.

To predict the market's closing price, Kang Zhang et al. [12] combined LSTM with Generative Adversarial Networks (GAN) using a Multi-Layer Perceptron (MLP) as the discriminator. This model exhibited significantly improved performance compared to other feed-forward neural network models.

Hao Li et al. [7] introduced a multi-input LSTM (MI-LSTM) architecture that outperformed the basic LSTM model by 9.96% according to the mean square error metric. The MI-LSTM architecture effectively filters noise from the data and extracts useful information.

Achyut Ghosh et al. [8] focused on determining the optimal time to study a company's stock by analyzing its growth. LSTM was utilized to calculate the error rate for companies in each sector.

These studies demonstrate the continuous development and modification of different LSTM-based models to achieve improved results in making future stock market predictions.

CHAPTER 3: Methodology

3.1 Description of Dataset

DATA	SITE	SCALING INDEX
India NIFTY Indices	https://www1.nseindia.com/products/content/equities/ indices/historical_index_data.htm	10000
Germany DAX 40 Indices	https://www.marketwatch.com/investing/index/dax/download-data?countrycode=dx&mod=mw_quote_tab	10000
UK FTSE Indices	https://www.marketwatch.com/investing/index/ukx/download-data?countrycode=uk&mod=mw_quote_tab	1000
USA S&P 500 Indices	https://finance.yahoo.com/quote/%5EGSPC/	1000
COVID-19 Data	https://ourworldindata.org/coronavirus	-

Table 2: Dataset Description

These indices consist of historical daily opening, closing, high and low stock prices along with Adjacent close and Volume. The dataset includes data from 30th January 2020 to 31st March 2022 of sectors Bank, Information Technology, Healthcare, and Retail. The date 30th January 2020 has been selected as the start date of our analysis as 27th January 2020 was the date of detection of India's first Covid-19 case. The frequency of the dataset is "daily except holidays and weekends". Thus, the dataset has 540 sequences or instances of data.

3.2 Proposed Model

The LSTM (Long Short-Term Memory) architecture is well-known for its ability to address the vanishing gradient problem that can occur in traditional feed-forward neural networks. It achieves this by incorporating a cell state that carries information from previous time steps to later ones, allowing for better memory retention. The LSTM model utilizes three gate mechanisms: the Forget gate, Input gate, and Output gate, which selectively control the flow of information in and out of the cell state. Activation functions like sigmoid and tanh are used to regulate the values within the network and compress the data.

In our proposed model, we adopt a stacked LSTM architecture consisting of three LSTM layers stacked on top of each other. Each LSTM layer has a different number of hidden layers: 128, 64, and 16, respectively. Sequential data, such as closing price values, is fed into the LSTM model.

The model then predicts future values based on this input. The final step in the model is the output dense layer. During training, the model uses a mean squared error loss function to measure the discrepancy between predicted and actual values. The Adam optimizer is employed for optimization, with a learning rate of 0.01.

CHAPTER 4 : Experimental Settings

4.1 Technology/ Tools

The Python programming language was utilized to develop the model, leveraging various libraries for different functionalities:

- Keras: We employed Keras, an open-source deep learning framework, which offers a user-friendly interface for defining and training various deep learning models. It integrates well with TensorFlow and provides efficient data parallelism, enabling faster processing of large datasets.

- TensorFlow: TensorFlow, developed by Google, is a widely used open-source platform for machine learning and deep learning. It provides a comprehensive ecosystem of tools and resources for training and deploying ML/DL models.

- Scikit-learn: As a powerful machine-learning library in Python, it offers a wide range of functionalities, including dimensionality reduction, classification, clustering, and regression. We utilized scikit-learn for various machine learning and statistical modeling tasks.

- Matplotlib: Matplotlib is a plotting library used for visualizing data. In our paper, we used Matplotlib to plot time series data, showcasing the actual and predicted values.

- Pandas: The Pandas library was employed to handle data manipulation tasks. It allowed us to read comma-separated values (CSV) files and convert them into dataframe format for further analysis.

- NumPy: NumPy is a fundamental library for numerical computations in Python. We utilized NumPy for matrix operations, such as reshaping data and generating random matrices.

- Math: The Math library was used for calculating performance metrics. In our paper, we specifically employed the root mean square error (RMSE) metric.

- Datetime: The Datetime library was used to convert the date column from an object format to a datetime format, enabling easier handling and manipulation of date-related information.

These libraries collectively provided a comprehensive set of tools and functionalities to support the development and analysis of our model.

4.2 Implementation Steps:

- *Stage 1*: Data Collection: During the data collection stage, we obtained historical stock data from reputable sources, including the official websites of

NIFTY 50 (India), DAX 40 Index (Germany), FTSE Indices (UK), and S&P 500 Indices (USA). The data collected encompassed four sectors: Bank, Information Technology, Healthcare, and Retail. This extensive dataset was then utilized to analyze market trends and make predictions for the closing prices of stocks over a one-month period. The predicted data covered the duration from February 4, 2022, to March 31, 2022. By leveraging this historical data, we aimed to gain insights into future patterns and forecast stock performance during this specific timeframe.

- *Stage 2*: Data Preprocessing: This stage involves:

    a) Data discretization: As part of data reduction, we have performed data discretization by selecting daily data for analysis, excluding holidays and weekends. This ensures that we have a consistent and manageable dataset for our analysis.

    b) Data transformation: To ensure that the data is suitable for LSTM analysis, we have applied data normalization. In this process, we have used the MinMax scalar technique to scale the values between 0 and 1. This normalization step helps to standardize the data and make it more compatible with the LSTM model.

    c) Data cleaning: The data cleaning step involves obtaining uniformity across all datasets thereby ensuring that the dataset of stock market and covid cases cover daily data except that of holidays and weekends.

    d) Data integration: In this step, we have collected data from various sources and integrated them into a unified dataset. After the data has been transformed and cleaned, we divided the dataset into training and testing sets. For our analysis, we have allocated 65% of the data for training purposes, allowing the model to learn patterns and relationships from the training data and reserved the rest 35% for testing purpose.

- *Stage 3:* Feature Extraction: In this stage, all the relevant features for stock price prediction are fed into the neural network model. The primary analytical parameter used in our study is the closing price index of stocks. This parameter serves as the target variable for the model to learn and make predictions. By using the closing price index as the analytical parameter, we aim to capture the patterns and trends in the stock prices that can help in predicting future price movements.

- *Stage 4:* Training Neural Network: At this stage, the neural network is trained using the dataset reserved for training. The training process involves feeding the sequential input layer, three LSTM layers, and a dense layer with an Adam activation function into the stacked LSTM model. The model is initialized with various weights, biases, and optimizers to optimize the learning process. In our study, we trained the model for 100 epochs, with each epoch consisting of multiple iterations or batches of data. The batch size used for training is 64, indicating that the model processes 64 data points at a time before updating the weights and biases. This helps in optimizing the training process and improving the efficiency of the model. Furthermore, the step size used for evaluation is 150, which means that the model makes predictions for every 150th data point in the dataset. This step size allows us to evaluate the performance of the model at regular intervals and assess its ability to predict stock prices accurately.
  Overall, the training process involves iteratively adjusting the weights and biases of the model based on the training data to minimize the difference between the predicted and actual closing prices, aiming to improve the model's predictive capabilities.

- *Stage 5:* Output Generation: In the final phase, the output layer of the model generates the forecasted value, which is the predicted closing price of the stock. This forecasted value is compared to the actual target value, which is the true closing price of the stock. To evaluate the performance of the model and measure the discrepancy between the predicted and target values, the root mean square error (RMSE) performance metric is used. RMSE is a commonly used metric in regression tasks and represents the square root of the average of the squared differences between the predicted and target values. It provides a measure of the overall accuracy of the model's predictions, with lower RMSE values indicating better performance. By calculating the RMSE, we can assess how well the model is able to capture the patterns and trends in the stock prices and make accurate predictions. A lower RMSE indicates that the model's predictions are closer to the actual values, suggesting that the model has higher prediction accuracy. Overall, the RMSE performance metric is used in this study to quantitatively evaluate the effectiveness of the model in predicting stock prices by assessing the discrepancy between the predicted and target values.

CHAPTER 5: Research Result And Analysis

Different researchers have explored various machine learning algorithms and techniques to develop accurate stock market prediction models. LSTM has gained popularity due to its stability, although it can have higher runtime compared to other algorithms [21]. The key to obtaining the best-fit model lies in tuning the hyperparameters such as activation function, number of epochs, hidden layers, optimizers, batch size, and more. By experimenting with different configurations of these parameters, researchers aim to find the optimal model.

For example, Murtaza Roondiwala et al. [6] achieved the lowest training and testing error after 500 epochs when working with the NIFTY 50 dataset. Kamalov et al. [22] compared optimization techniques like Stochastic Gradient, RMSprop, and Adam while using ReLU as the optimizer for predicting the next day's price of the S&P 500 index. Their study concluded that the RMSprop optimizer produced the best results, with mean and validation Mean Absolute Error of 0.0148 and 0.0150, respectively. Moghar et al. [23] considered the number of epochs as a tuning factor in their study using data from companies like GOOGL and NKE from the New York Stock Exchange.

In this particular paper, the COVID-19 period was chosen to understand the data dynamics during an emergency situation. The prediction was conducted for specific sectors, consisting of companies with similar business models, to analyze sector performance. While direct comparisons between individual companies are not possible, analyzing the performance of the entire sector provides valuable insights. When considering different countries, it becomes possible to compare how these sectors globally fared during that time span.

The paper employs a stacked LSTM model to predict the next-day closing index for a span of 38 days. Figures 1, 2, 3, and 4 in the paper illustrate the closing prices of stocks in the bank, IT, healthcare, and retail sectors (represented by the blue line), along with the training error (orange line) and testing error (green line). These figures also include the daily COVID-19 cases recorded from January 30, 2020, to March 31, 2022, providing context for the analysis.

TABULAR REPRESENTATION OF ERROR VALUES

SECTOR	TRAINING ERROR	TESTING ERROR
Bank	1.71	3.74
IT	547.69	704.42
Healthcare	62.68	106.36
Retail	19.84	27.22

Table 3: Error value for sectors in Germany

SECTOR	TRAINING ERROR	TESTING ERROR
Bank	545.15	543.23
IT	350.79	1818.56
Healthcare	109.48	129.64
Retail	5.37	26.38

Table 4: Error value for sectors in India

SECTOR	TRAINING ERROR	TESTING ERROR
Bank	39.67	135.00
IT	24.29	28.89
Healthcare	6.99	22.51
Retail	6.39	12.39

Table 5: Error value for sectors in United Kingdom

SECTOR	TRAINING ERROR	TESTING ERROR
Bank	6.55	9.34
IT	35.75	105.81
Healthcare	11.82	27.31
Retail	3.32	3.75

Table 6: Error value for sectors in the United States of America

i. GRAPHICAL REPRESENTATION OF GERMANY

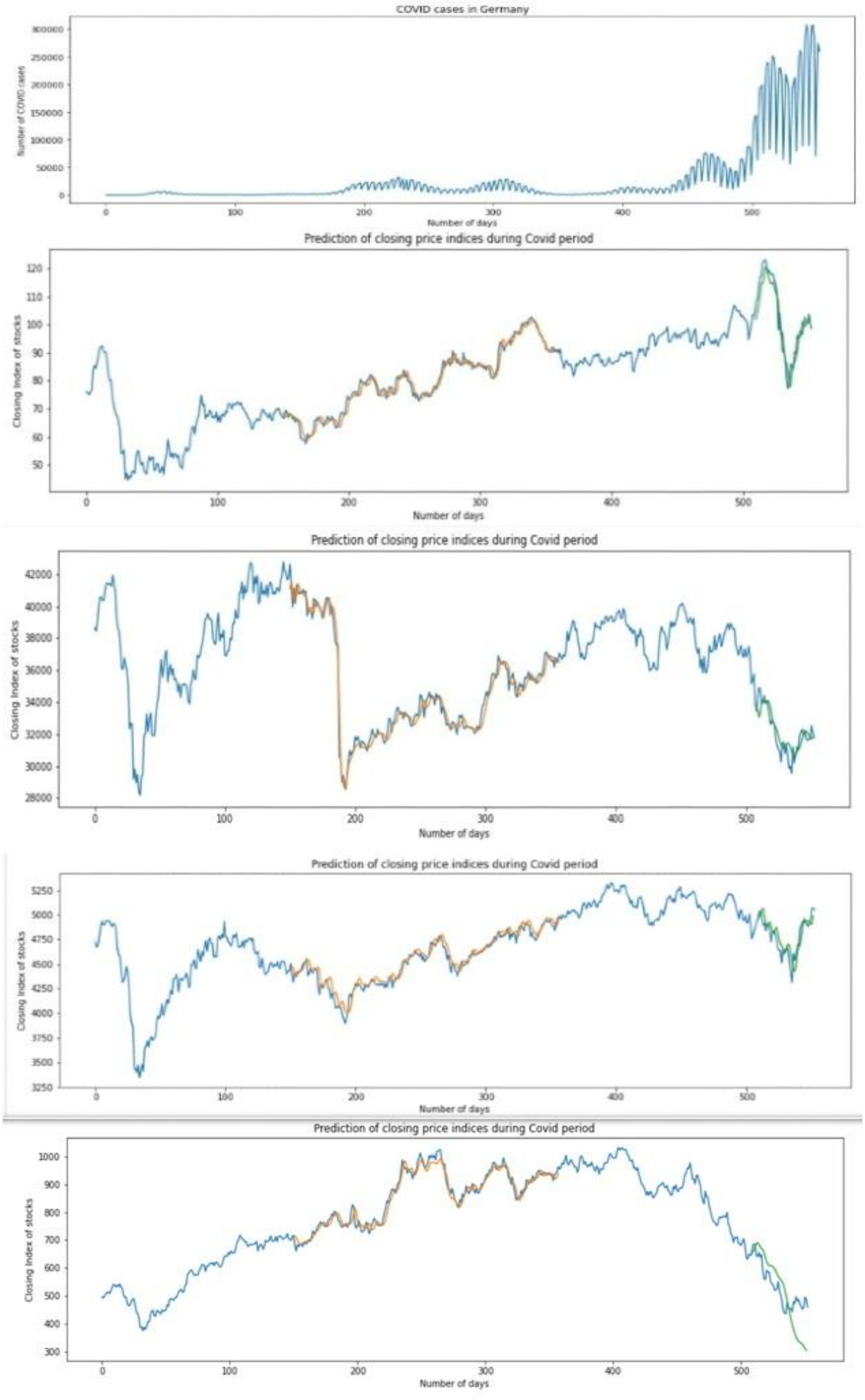

Figure 1: Graph (i) Daily new COVID cases in Germany (ii) Prediction graph for Germany’s Bank sector (iii) Prediction graph for Germany’s IT sector (iii) Prediction graph for Germany’s Healthcare sector (iv) Prediction graph for Germany’s Retail sector

ii. GRAPHICAL REPRESENTATION OF INDIA

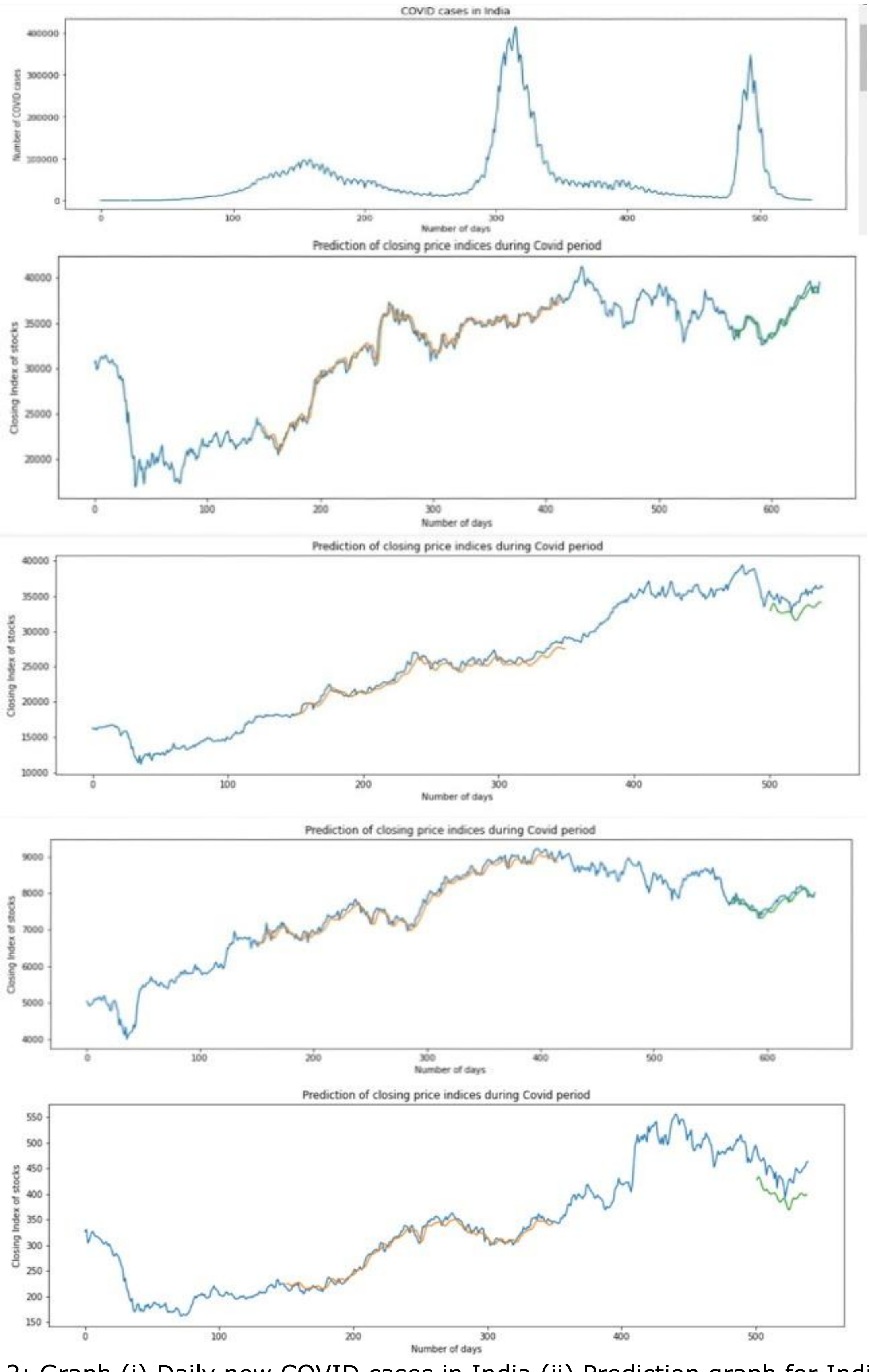

Figure 2: Graph (i) Daily new COVID cases in India (ii) Prediction graph for India Bank sector (iii) Prediction graph for India IT sector (iii) Prediction graph for India Healthcare sector (iv) Prediction graph for India Retail sector

iii. GRAPHICAL REPRESENTATION OF THE UNITED KINGDOM

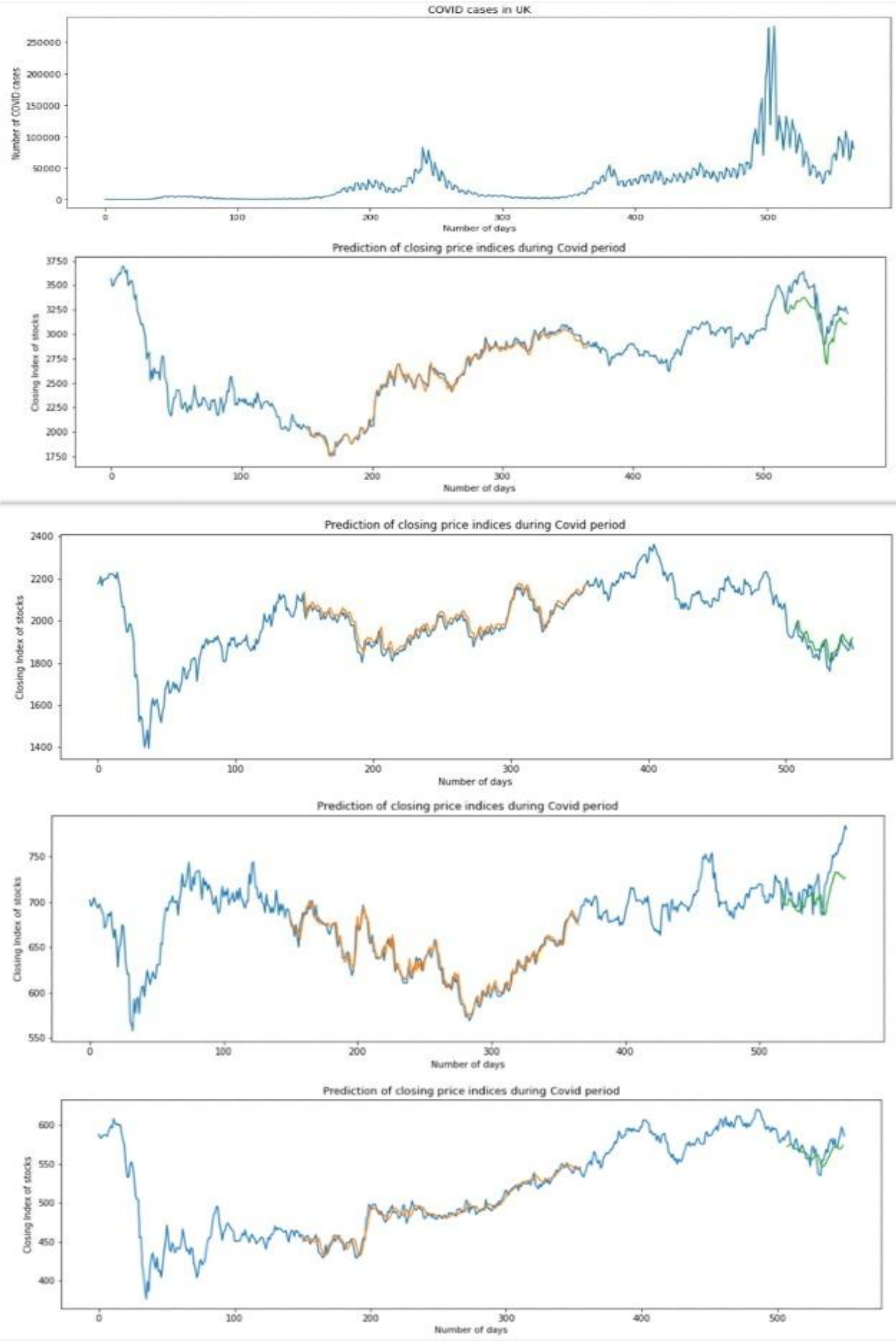

Figure 3: Graph (i) Daily new COVID cases in UK (ii) Prediction graph for UK Bank sector (iii) Prediction graph for UK IT sector (iii) Prediction graph for UK Healthcare sector (iv) Prediction graph for UK Retail sector

iv. GRAPHICAL REPRESENTATION OF THE UNITED STATES OF AMERICA

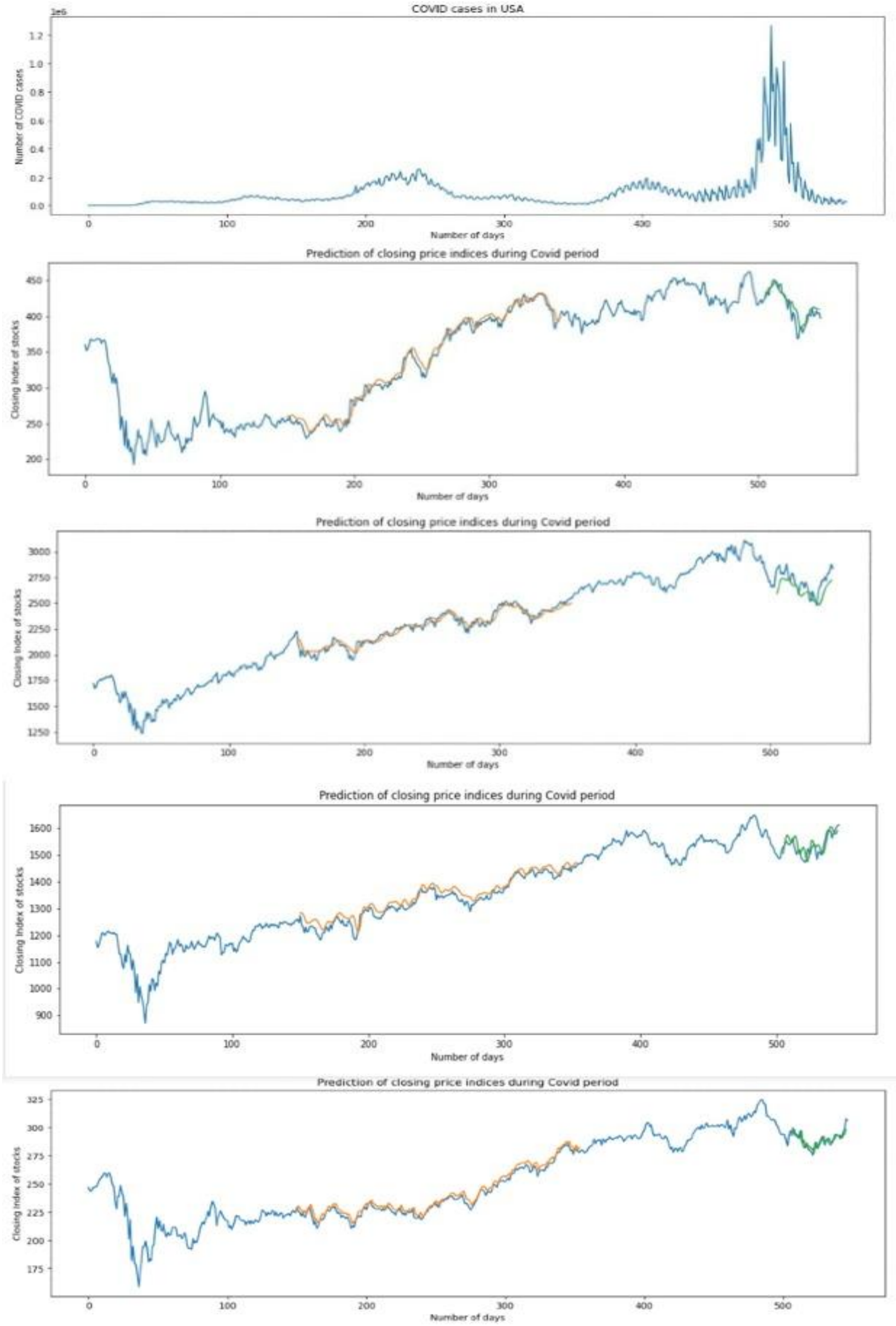

Figure 4: Graph (i) Daily new COVID cases in USA (ii) Prediction graph for USA Bank sector (iii) Prediction graph for USA IT sector (iii) Prediction graph for USA Healthcare sector (iv) Prediction graph for USA Retail sector

The implementation of the suggested stacked LSTM model is carried out using the Python programming language. The model is trained and tested on data from all four sectors (Bank, Healthcare, IT, and Retail), and the predictions are plotted to assess the effectiveness of the model. The errors in the predicted prices for each sector are presented in Table 3, 4, 5, and 6.

To evaluate the accuracy of the model's predicted prices, the root mean square error (RMSE) is utilized as the performance metric. RMSE measures the difference between the predicted closing prices and the actual values, providing an indication of the model's prediction error. By calculating the RMSE, the model's ability to accurately forecast stock prices can be assessed.

$$\text{RSME}= \sqrt{\frac{1}{N}\Sigma(\hat{y}_i - y_i)^2}$$

Where y_i refers to the predicted closing price, y_i refers to original closing price.

OBSERVATIONS:

The COVID-19 pandemic has had significant effects on various sectors globally, including the bank, healthcare, IT, and retail sectors. Here is an overview of the effects in India, USA, Germany, and the UK:

1. Bank Sector:

- India: The banking sector in India faced challenges due to the economic slowdown, reduced business activity, and disruptions in loan repayments. Non-performing assets (NPAs) increased, and banks had to implement measures such as loan restructuring and moratoriums [26].
- USA: Banks in the USA experienced volatility in financial markets, decreased demand for loans, and increased loan defaults. The Federal Reserve implemented monetary policy measures to stabilize the financial system and support banks [27].
- Germany: German banks faced similar challenges, including increased credit risk and lower interest rate margins. The government provided support through financial stimulus packages and loan guarantee programs [28].
- UK: UK banks faced the impact of reduced economic activity, loan defaults, and increased provisions for bad debts. The government implemented measures to provide liquidity support and regulatory flexibility [29].

2. Healthcare Sector:

- India: The healthcare sector in India faced immense pressure due to the surge in COVID-19 cases. Hospitals experienced high patient loads, shortages of medical supplies, and staffing challenges. The sector also saw increased demand for telemedicine services [30].

- USA: The healthcare sector in the USA faced a strain on healthcare infrastructure, overwhelmed hospitals, and shortages of personal protective equipment (PPE). The pandemic highlighted the need for healthcare system reforms and increased investment in public health [31].

- Germany: German healthcare system coped relatively well, with efficient testing and contact tracing. Hospitals were prepared for the surge in cases, and the government provided financial support to healthcare providers [32].

- UK: The UK healthcare sector faced challenges in managing the high number of COVID-19 cases. Hospitals faced capacity issues, shortages of PPE, and staff burnout. The vaccination drive played a crucial role in managing the pandemic [33].

3. IT Sector:

- India: The Indian IT sector witnessed disruptions in operations due to lockdown measures. Many companies transitioned to remote work, and demand for digital solutions, cybersecurity, and cloud services increased [34].

- USA: The IT sector in the USA experienced a mixed impact. While some segments, such as remote work solutions and e-commerce, thrived, other areas like IT consulting and hardware faced challenges due to reduced business activity [35].

- Germany: German IT companies benefited from the increased demand for digitalization and technology adoption. Remote work solutions, e-commerce platforms, and cybersecurity services saw growth [36].

- UK: The UK IT sector faced disruptions and uncertainties. Remote work solutions, digital communication tools, and e-commerce platforms saw increased demand, while IT projects and investments were affected [37].

4. Retail Sector:

- India: The retail sector in India was significantly impacted by the pandemic and lockdowns. Physical retail stores faced closures, supply chain disruptions, reduced consumer spending, and a shift towards e-commerce platforms [38].

- USA: The US retail sector experienced a mix of challenges and opportunities. Brick-and-mortar retailers faced closures and reduced foot traffic, while e-commerce and online retail saw significant growth [39].

- Germany: German retail businesses faced temporary closures and reduced consumer spending. Essential goods and online retail segments performed relatively well during the pandemic [40].

- UK: The UK retail sector faced store closures, supply chain disruptions, and reduced consumer confidence. E-commerce and online retail gained prominence, and there was an increased focus on safety measures in physical stores [41].

- Market Volatility: In March 2020, there was a sudden and significant dip in stock market values globally. This volatility was observed across countries and sectors, reflecting the high level of uncertainty caused by the pandemic.

- Panic Trading: The rapid spread of the virus and the absence of a known cure resulted in panic trading among investors. This led to a sharp decline in stock market values.

- Global Stock Market Decline: Major global indexes, including the S&P 500, FTSE 100, DAX, and Nikkei, experienced double-digit percentage declines in 2020. The S&P 500, in particular, faced a loss of 37% of its value [24].

- Economic Disruptions: The pandemic caused disruptions in labor markets, leading to significant contractions in GDP. Factories closed, and unemployment rates soared, with millions of jobs lost worldwide.

- Stock Market Reaction to COVID-19: Stock market losses coincided with daily increases in COVID-19 infections. The stock market's performance was closely tied to the progression of the virus and its impact on economic activities.

- Quick Recovery: Unlike the Stock Market Crash of 1929, the stock markets recovered quickly in 2020. By November, many indexes had reached or surpassed their pre-pandemic levels [24].

- Investor Behavior: Despite the initial crash, many investors continued to invest rather than sell, driven by confidence in federal stimulus measures and vaccine development.

- Impact on Industries: The IT industry experienced both challenges and opportunities. Remote work arrangements led to a loss of opportunities for companies with international dealers and manufacturers. However, there was an increased demand for software and social media platforms, contributing to the industry's expected market boom [25].

- Construction and Housing Sector: The lack of laborers and migrant workers returning to their hometowns had a significant impact on housing and construction projects, leading to reduced workforce and material shortages.

- Banking Sector: Banks faced challenges such as low net interest margins, decreased retail spending, and increased non-performing loan ratios, particularly among SMEs.

- Healthcare Industry: Many medical offices not directly involved in virus-related needs were temporarily closed. Deferred elective surgeries and treatments for chronic illnesses resulted in revenue losses for healthcare providers.

Overall, the COVID-19 pandemic had profound and varied effects on the stock market and different industries, highlighting the complex interplay between global health crises and economic dynamics.

CHAPTER 6 : Conclusion

In this research paper, we proposed a stacked LSTM model to predict closing prices for the auto, healthcare, metal, and bank sectors during the challenging times of the Covid-19 pandemic. This period witnessed significant events such as lockdowns, vaccination drives, and the Russia-Ukraine crisis, all of which had substantial impacts on stock markets. Consequently, forecasting stock prices has become a difficult task. While analyzing stock price data, variables such as high, low, open, close, and adjacent close values, as well as trading volume, are commonly considered. However, these variables alone do not capture the entirety of the stock market. To improve accuracy, it is essential to incorporate other factors such as geopolitical stability, investor sentiment, market rumors, and physical factors, among others. These additional parameters contribute to the outcome of stock prices. The stock market is currently highly popular and is expected to continue gaining traction. This has motivated researchers to explore new techniques and propose innovative models for prediction. These forecasting approaches assist investors in understanding the market, analyzing data, and making informed investment decisions.

Looking ahead, the future scope of this paper involves conducting a detailed comparison by analyzing stock market data before and during the Covid-19 period across various sectors. This analysis would provide valuable insights into the effects of the pandemic on different sectors. Uncertain times often lead to new opportunities, and it has been observed that some sectors thrived during the Covid-19 era while others faced significant setbacks. By utilizing these analytical techniques, it becomes possible to anticipate future possibilities in the stock market.

BIBLIOGRAPHY

[1] Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W.C., Wang, C.B. and Bernardini, S., 2020. The COVID-19 pandemic. Critical reviews in clinical laboratory sciences, 57(6), pp.365-388.

[2] Ghosh, A., Nundy, S. and Mallick, T.K., 2020. How India is dealing with COVID-19 pandemic. Sensors International, 1, p.100021.

[3] Hur, J., Raj, M. and Riyanto, Y.E., 2006. Finance and trade: A crosscountry empirical analysis on the impact of financial development and asset tangibility on international trade. World Development, 34(10), pp.1728-1741.

[4] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. Neural computation, 9(8), pp.1735-1780.

[5] Fernandez-Perez, A., Gilbert, A., Indriawan, I. and Nguyen, N.H., 2021. COVID-19 pandemic and stock market response: A culture effect. Journal of Behavioral and Experimental Finance, 29, p.100454.

[6] Roondiwala, M., Patel, H. and Varma, S., 2017. Predicting stock prices using LSTM. International Journal of Science and Research (IJSR), 6(4), pp.1754-1756. [7] Li, H., Shen, Y. and Zhu, Y., 2018, November. Stock price prediction using attention-based multi-input LSTM. In Asian conference on machine learning (pp. 454-469). PMLR.

[8] Ghosh, A., Bose, S., Maji, G., Debnath, N. and Sen, S., 2019, September. Stock price prediction using LSTM on Indian Share Market. In Proceedings of 32nd international conference on (Vol. 63, pp. 101-110).

[9] Sim, H.S., Kim, H.I. and Ahn, J.J., 2019. Is deep learning for image recognition applicable to stock market prediction?. Complexity, 2019.

[10] Chen, W., Yeo, C.K., Lau, C.T. and Lee, B.S., 2018. Leveraging social media news to predict stock index movement using RNN-boost. Data & Knowledge Engineering, 118, pp.14-24.

[11] Lin, B.S., Chu, W.T. and Wang, C.M., 2018, July. Application of stock analysis using deep learning. In 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI) (pp. 612-617). IEEE.

[12] Zhang, K., Zhong, G., Dong, J., Wang, S. and Wang, Y., 2019. Stock market prediction based on generative adversarial network. Procedia computer science, 147, pp.400-406.

[13] Eapen, J., Bein, D. and Verma, A., 2019, January. Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. In 2019 IEEE 9th annual computing and communication workshop and conference (CCWC) (pp. 0264-0270). IEEE.

[14] Vijh, M., Chandola, D., Tikkiwal, V.A. and Kumar, A., 2020. Stock closing price prediction using machine learning techniques. Procedia computer science, 167, pp.599-606.

[15] Ding, X., Zhang, Y., Liu, T. and Duan, J., 2015, June. Deep learning for event-driven stock prediction. In Twenty-fourth international joint conference on artificial intelligence.

[16] Wen, M., Li, P., Zhang, L. and Chen, Y., 2019. Stock market trend prediction using high-order information of time series. Ieee Access, 7, pp.28299-28308. [17] Ni, L., Li, Y., Wang, X., Zhang, J., Yu, J. and Qi, C., 2019. Forecasting of forex time series data based on deep learning. Procedia computer science, 147, pp.647-652.

[18] Nikou, M., Mansourfar, G. and Bagherzadeh, J., 2019. Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. Intelligent Systems in Accounting, Finance and Management, 26(4), pp.164-174.

[19] Lakshminarayanan, S.K. and McCrae, J.P., 2019, December. A Comparative Study of SVM and LSTM Deep Learning Algorithms for Stock Market Prediction. In AICS (pp. 446-457).

[20] Jin, Z., Yang, Y. and Liu, Y., 2020. Stock closing price prediction based on sentiment analysis and LSTM. Neural Computing and Applications, 32(13), pp.9713-9729.

[21] Yadav, A., Jha, C.K. and Sharan, A., 2020. Optimizing LSTM for time series prediction in Indian stock market. Procedia Computer Science, 167, pp.2091-2100.

[22] Kamalov, F., Smail, L. and Gurrib, I., 2020, November. Stock price forecast with deep learning. In 2020 International Conference on Decision Aid Sciences and Application (DASA) (pp. 1098-1102). IEEE.

[23] Moghar, A. and Hamiche, M., 2020. Stock market prediction using LSTM recurrent neural network. Procedia Computer Science, 170, pp.1168-1173.

[24] https://www.thestreet.com/dictionary/c/covid-19-stock-market-crash-of-2020

[25] https://www.marketdataforecast.com/blog/impacts-of-covid19-on-information-technology-industry

[26] "COVID-19 Impact: Banking sector may be staring at a huge NPA shock"- https://economictimes.indiatimes.com/industry/banking/finance/banking/covid-impact-banks-might-be-staring-at-a-spike-in-retail-npas/articleshow/80615103.cms

[27] "The Impact of COVID-19 on the US Banking System" - Board of Governors of the Federal Reserve System: https://www.federalreserve.gov/publications/files/financial-stability-report-20210506.pdf

[28] "Covid-19 and German Banking: Vulnerabilities, Stability Measures and Future Perspectives" - Deutsche Bundesbank: https://www.bundesbank.de/resource/blob/852106/6e8f415fd6fb449c99dd1f41efb376b3/mL/2020-07-covid-19-and-german-banking-data.pdf

[29] "COVID-19: Impact on the UK Banking Sector" - PwC UK: https://www.pwc.co.uk/industries/banking-capital-markets/insights/covid-19-impact-on-the-uk-banking-sector.html

[30] "COVID-19 Pandemic: Impact on the Indian Healthcare System" - Journal of Family Medicine and Primary Care: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7646922/

[31] "Impact of COVID-19 on healthcare systems: A review" - Journal of Clinical Nursing: https://pubmed.ncbi.nlm.nih.gov/32678419/

[32] "Germany's Response to COVID-19: A Primer on Early Actions and Lessons Learned" - Center for Global Development: https://www.cgdev.org/publication/germanys-response-covid-19-primer-early-actions-and-lessons-learned

[33] "The impact of COVID-19 on the NHS" - The Health Foundation: https://www.health.org.uk/publications/long-reads/the-impact-of-covid-19-on-the-nhs

[34] "The Impact of COVID-19 on Indian IT Services Companies" - International Journal of Applied Engineering Research: http://www.ripublication.com/ijaer19/ijaerv14n11_84.pdf

[35] "How COVID-19 is Impacting the IT Industry in the US" - CompTIA: https://www.comptia.org/content/research/how-covid-19-is-impacting-the-it-industry-in-the-us

[36] "Digital Transformation in Germany's Information and Communications Technology (ICT) Market" - International Trade Administration: https://www.trade.gov/germanys-ict-market

[37] "COVID-19 and the UK Tech Sector" - Tech Nation: https://technation.io/insights/reports/covid-19-and-the-uk-tech-sector/

[38] "COVID-19 impact: A look at the Indian retail landscape" - EY India: https://www.ey.com/en_in/consumer-products-retail/covid-19-impact-a-look-at-the-indian-retail-landscape

[39] "The Impact of COVID-19 on the US Retail Industry" - McKinsey & Company: https://www.mckinsey.com/industries/retail/our-insights/the-impact-of-covid-19-on-the-us-retail-industry

[40] "The impact of COVID-19 on retail sales in Germany" - Statista: https://www.statista.com/statistics/1102694/covid-19-impact-on-retail-sales-by-category-germany/

[41] "UK retail sales hit by largest drop on record due to coronavirus" - The Guardian: https://www.theguardian.com/business/2020/apr/24/uk-retail-sales-hit-by-largest-drop-on-record-due-to-coronavirus

Covid-19

ORIGINALITY REPORT

5%	6%	0%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

8%

★ Submitted to Institute of Technology, Nirma University

Student Paper

Exclude quotes	On	Exclude matches	< 2%
Exclude bibliography	On		