2022 International Research Conference on Smart Computing and Systems Engineering (SCSE) | 978-1-6654-7375-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/SCSE56529.2022.9905191

Stock Market Prediction using Artificial Intelligence

M. W. Sachin Dilhan*
Department of Statistics and Computer Science
Faculty of Science, University of Kelaniya, Sri Lanka
sachindilhan2021@gmail.com

N. Mihirini Wagarachchi
Department of Interdisciplinary Studies
University of Ruhuna, Sri Lanka
mihirini@is.ruh.ac.lk

***Abstract* - This research focuses on predicting stock closing prices for one day or the future in specific economic conditions. Today, Sri Lanka faces a financial crisis due to the COVID-19 pandemic. Therefore, lots of investors are bankrupt due to unpredictable stock prices. This work mainly focuses on predicting stock prices in banking sector shares such as Commercial Bank (COMB.N), Hatton National Bank (HNB.N), Seylan Bank (SEYB.N), and Sampath Bank (SAMP.N) on Colombo Stock Exchange (CSE) in Sri Lanka. According to the hypothesis, All Share Price Index (ASPI) and Banking Sector indices have been taken as a numerical sentiment parameter other than the historical prices from each bank. Since ASPI shows overall market performance and Banking sector indices show banking sector capitalization changed over time. There can be a positive and negative sentiment when the ASPI and Sector Indices increase and decrease, respectively. Finally, a dataset is divided into 70% for training and 30% for testing. This study has used Recurrent Neural Networks (RNNs) such as Long short-term memory (LSTM) and Gated Recurrent Unit (GRU) using 25, 50, 100, 150, and 200 epochs. LSTM model has given the lowest Mean Squared Error (MSE) and Root Mean Square Error (RMSE). According to the LSTM model, COMB.N, HNB.N, and SAMP.N were given the lowest MSE, and RMSE for 100 epochs, and SEYB.N was given the lowest MSE and RMSE value for the 150 epochs.**

***Keywords* - Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), Machine Learning, stock market prediction, Streamlit API**

I. Introduction

Today's world is facing a financial crisis due to the COVID-19 pandemic, and many investors are bankrupt or lost a large part of their investment, due to the wide variation in stock prices; This has led to investors losing their confidence in investing in the stock market. The financial market has seen a significant impact on economies. Often Business organizations in the current economy are highly dependent on funds from these stock markets. Therefore, examining the way of acting and identifying the performance of stock markets is an essential area of research. These types of analyses include Predicting the price of shares, for example, stocks and bonds, etc. Exchange rate, securities benefit, market indicators, trade volumes, inventory classification, etc.

Rising share markets indicate favorable economic conditions for companies, which successively translate into higher profits. On the other hand, the declining share prices show an economic downturn. These trends will likely point out the economy and share tandem within the coming days. In the case of stock markets, the task is vital due to massive dynamic changes in stock market behavior and many immeasurable economic benefits. By predicting stock prices, risk managers and professionals can see if their portfolio will be declined in the future and decide whether they want to sell before it loses market value. Therefore, research to predict future trends in financial indices is critical and necessary for investors interested in the stock markets.

This study presents a method to increase investor confidence by predicting the banking stock prices using machine learning techniques. It focuses on predicting stock closing prices for one day or in the future in specific economic conditions. This predictability, could lead to an increase in investor confidence and this would have a favorable impact on the Sri Lankan economy.

II. Ease Of Use

The Graphical User Interface (GUI) is developed with streamlit Application Programming Interface (API). Then the user can interact with the system. If an investor wants to predict a one-day closing price, then the investor needs to enter current market information to the API. If an investor needs to predict the future closing price of a particular share, then the investor needs to enter how many years need to predict. Then an investor can see how the closing price goes graphically throughout the years.

III. Literature Review

The authors in [1] developed a predictive model with Recurrent Neural networks (RNN) and Long Short-term Memory (LSTM) to predict future stock prices. Different numbers of epochs have been used for training the datasets. GOOGL and NKE companies were selected for prediction from the New York Stock Exchange (NYSE). The historical data set is collected from 19/08/2004 to 19/12/2019 from GOOGL and NKE. Future work was proposed to find out the simplest collection for related data, collection length, and by increasing the number of epochs for training to maximize the accuracy of that predictions.

The authors in [2] discussed an algorithm to predict the share prices of the Indian stock market. Stock prices are affected by many factors such as political reasons, news, public safety events, etc. The major limitation of this study is to predict the share prices with that kind of fluctuations. The proposed model predicts stock prices based on news sentiments analysis and historical stock market time series data. Sentiments have been extracted from Twitter and incorporate the polarity to enhance the prediction accuracy. By increasing accuracy, investors can make wiser and better investment decisions. Natural Language Processing (NLP) has been used to identify the psychology of investors. The study introduced some polarity regarding the score level for the social media data and news data. The second step was to input the historical time series data. Finally, the proposed RNN model predicts the historical pattern.

The authors in [3] developed a model to predict stock prices in several companies on the Colombo Stock Exchange (CSE) in Sri Lanka. The study proposed an RNN and identified the limitations of the algorithms to measure the accuracy of this model. Feedforward, LSTM, gated recurrent unit (GRU), and Simple Recurrent Neural Network (SRNN) were used as the architectures. This research mainly focused on three companies under different sectors which are banking, finance and insurance, the manufacturing sector, and the diversified holding sector. Commercial Bank (COMB.N), Royal Ceramics Limited (RCL.N), and John Keels Holding (JKH.N) have been selected respectively. Dataset was collected from 01/01/2002 to 30/06/2016. The LSTM model was the best among all RNN models. When compared with the LSTM model with other feedforward networks, LSTM and SRNN were given the lowest error. GRU network provides the most elevated errors relative to the other neural network methods.

The authors in [4] conducted research based on neural networks to forecast the share prices in the future and trends in the Chinese stock market. The proposed solution was comprehensive. This model has been given high accuracy according to the stock market prediction. Stock market prediction needs both fundamental and technical analysis. This research has three parts: data extraction, preprocessing, and stock market prediction. This research was mainly focused on stock price prediction, which was based on LSTM. Feature expansion and recursive feature elimination approaches were used to clean up the dataset. This model was developed by combining the feature engineering part with the LSTM model and by comparing the most frequently used models.

The authors in [10] conducted research based on stock market perdition using a Generative Adversarial Network (GAN), LSTM, and RNN. The GAN was used as a discriminator for predicting the future closing price of some stocks. This produced a promising result in the closing price of some stocks with the help of the LSTM, which worked as the generator. Multi-layer perceptron was used to work as a discriminator. The model was trained with past several data as a final way to forecast the daily closing price. This research was mainly focused on International Business Machine (IBM) from the New York Stock Exchange (NYSE) and Ping an Insurance Company in China (PAICC). Root Mean Square Error (RMSE) was computed for five datasets to evaluate the criteria. For comparison purposes, this research also used some Support Vector Regression (SVR), Artificial Neural Networks (ANN), and LSTM models. According to the results compared RMSE, this GAN model has given the best result and minimum error.

Several limitations can be seen in previous studies. The researchers didn't consider other relevant factors such as sentiment news, dividends, etc. Not only the historical share prices can predict future prices, but also there are some other factors that affect share prices. For example, COVID-19 pandemic sentiment excessively affected whole world share markets. There are some sectors can be seen in stock market such as banking and finance, appeal sector, manufacturing etc. Therefore, sentiments are affected differently to the sector. It is needed to consider sector-wise analysis for sentiments as well as the historical share prices to obtain a better accuracy.

IV. MATERIALS AND METHODS

A. GRU and LSTM Algorithems

There are three layers in Artificial Neural Networks (ANN) as the Input layer, Hidden layer, and Output layer. LSTM is a special or different kind of RNN that is usually used for stock price prediction. There can be one or more hidden layers in Neural Networks. The number of nodes is relative to the input layer which depends on the dimension of the data. Input nodes related to the hidden layer is called "synapses". Weight is the coefficient and the relation between the input and hidden layer node. Neural networks learn by adjusting the weights of the different nodes. After the learning process is done, ANN has the optimal weights for each synapse.

Sigmoid and tangent hyperbolic(tanh) features are applied for the hidden layer nodes at the sum of the different weights from the other inputs. It is referred to as activation functions. The output of the developed model sometimes will not give the effectual output. For correcting the output, backpropagation needs to be implemented to take the maximum output by reducing the error. Implement the backpropagation procedure. The output and hidden layer need to be connected by sending signals to match the suitable weight of each neuron and the appropriate error for deciding the number of epochs. This procedure is repeated until it meets the maximum prediction and reduces the prediction cost. RNN is used for future stock price forecasting purposes and is a kind of neural network used in earlier stages. Future values are predicted based on the historically collected data. Historical data should be remembered for predicting future share prices. Therefore, in RNNs, the hidden layer acts as a store for storing historical and sequential data. Recurring is used for demonstrating the process of utilizing the historical data for future prediction [2]. The previous stages' data must be remembered for predicting and forecasting future share prices.

There are some limitations in RNNs. It cannot store long-term data. But using the LSTM approach, historical data can be remembered as it is essential for future predictions. Using different gates with inbuilt memory, a memorization process can be done. GRU is an advanced type of recurrent neural network which was introduced in 2014. GRU architecture is more like the LSTM architecture because both LSTM and GRU use gates toflow information. But when comparing LSTM with GRU, GRU is relatively new compared to LSTM.

In GRU, each timestamp is called t. Then input has been taken as Xt. The hidden state is called Ht-1 from the historical timestamp t-1. After several times, it gives an output to the newly identified hidden state Ht. Then it will be passed to the new timestamp. GRU has two gates, one is called the reset gate and the other is called the update gate. The reset gate is responsible for the short-term memory of some neural networks. The update gate is responsible for long-term memory.

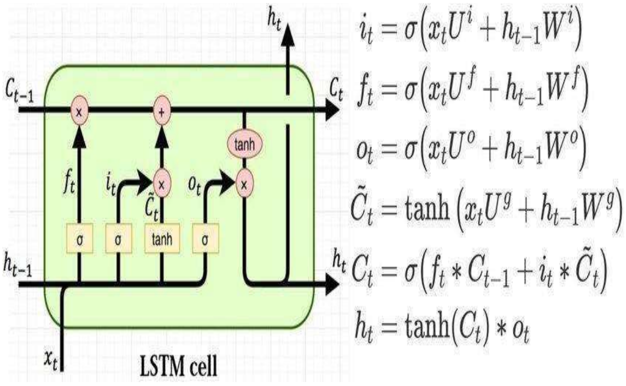

Fig. 1. LSTM Architecture

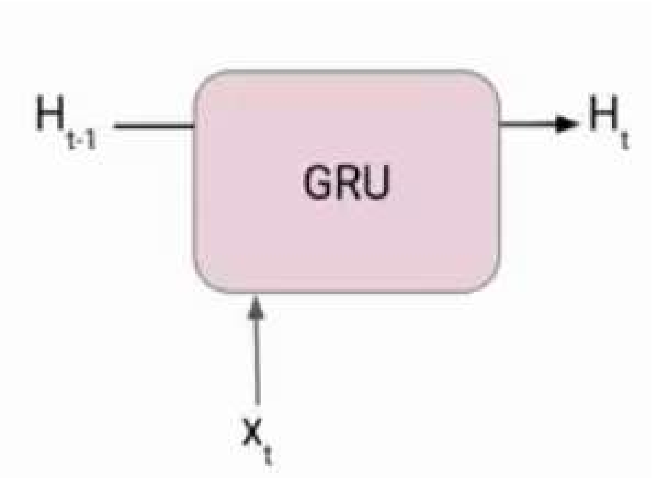

Fig. 2.Gated Recurrent Unit (GRU) Architecture

B. Methods and Materials Used

According to the existing literature the highest accuracy is given by the Recurrent Neural Network (RNN). This study decided to use LSTM and GRU with a different number of epochs to predict the closing price of the shares on the CSE banking sector in Sri Lanka.

1) Python language & libraries

In this research, the latest python version 3.10.2 was used. In the modern software industry, python is used as an interpreted and general-purpose programming language. Python has several advantages, such as readability with its notable use of relevant whitespaces and the object-oriented approach. Python has libraries that can be used for dataset creation, pre-processing, plot graphs, etc.

2) Google collaboratory notebooks

Colab notebook is running in the cloud and is highly integrated with google drive. Therefore, it does not interact with the local machine and performance can also be increased by changing runtime type from TPU to GPU. Google Drive can be used to upload and load data into a colab and results also can be stored in a colab.

3) Streamlit framework

Streamlit is an open-source app framework for the python language, which is used to develop apps or web applications in data science and machine learning projects within a minimal period. Streamlit is very familiar and also compatible with most of the python libraries.

4) Correlation test

The correlation Test is the technique to identify the relationship between two continuous and quantitative variables. If the correlation is between 0 to +1, which represents the positive relationship between two variables, and if the correlation varies from 0 to -1, that shows a negative correlation. However, if it is near ±1, that means it implies there is a strong correlation.

a) Pearson Correlation Test

Pearson correlation test can be used to statistically identify the strength between given two variables.

b) Spearman`s Ranking Correlation Test

Spearman ranking correlation test measures the degree of association between two variables.

C. Methodology

According to existing literature, RNN has given the highest accuracy. Therefore, the following steps have been taken to develop the LSTM and GRU models.

1) Data Collection

Data from the Colombo stock exchange (CSE) was collected for this research. Different data was used as a historical dataset for each banking sector share which includes closing, opening, high, low, trades, volume, and turnover for each bank. Other than these data, there are many data like All Share Price Index (ASPI), Sector indices, foreign holdings, foreign buying, dividends payments, etc. ASPI and Banking Sector indices have been taken as sentiment parameters according to the hypothesis. Daily stock data was collected from CSE for Commercial bank (COMB.N), Hatton National Bank (HNB.N), Seylan Bank (SEYB.N) and Sampath Banking (SAMP.N) from 01/01/2002 to 31/12/2020. ASPI and Sector indices were also used as a numerical sentiment parameter from 01/01/2002 to 31/12/2020.

2) Hypothesis

According to the hypothesis, ASPI and Banking Sector indices have been taken as a numerical sentiment parameter other than the historical prices from each bank. Since ASPI shows overall market performance and Banking Sector Indices show banking sector capitalization changed over time. There can be a positive and negative sentiment when the ASPI and Sector Indices increases and decrease respectively. None of the researchers have used ASPI and Banking Sector Indices to predict stock prices in CSE.

TRADING DATE	PRICE HIGH (Rs.)	PRICE LOW (Rs.)	CLOSE PRICE (Rs.)	OPEN PRICE (Rs.)	TRADE VO\RE VOLUME (No.)		TURNOVER (Rs.)
02-Jan-02	26.00	25.00	25.00	25.00	7	107,100	2,678,100.00
03-Jan-02	25.00	24.50	25.00	25.00	8	4,700	117,000.00
04-Jan-02	25.00	24.50	25.00	25.00	8	7,500	184,800.00
07-Jan-02	24.50	23.50	23.50	25.00	13	13,500	321,650.00
08-Jan-02	25.00	24.75	25.00	23.50	3	8,100	200,550.00
09-Jan-02	24.00	24.00	24.00	25.00	2	600	14,400.00
11-Jan-02	23.50	23.00	23.00	24.00	6	3,500	80,650.00
15-Jan-02	22.50	21.50	22.00	23.00	13	10,600	233,250.00
16-Jan-02	21.75	21.00	21.00	22.00	14	4,800	101,325.00
17-Jan-02	20.25	20.00	20.00	21.00	11	10,600	212,375.00
18-Jan-02	21.75	20.00	21.75	20.00	8	20,200	414,075.00
21-Jan-02	21.50	21.50	21.50	21.75	2	1,300	27,950.00
22-Jan-02	21.50	20.00	20.00	21.50	10	8,900	187,075.00
23-Jan-02	21.00	21.00	21.00	20.00	2	300	6,300.00
24-Jan-02	21.00	21.00	21.00	21.00	3	4,900	102,900.00
25-Jan-02	21.50	21.25	21.50	21.00	3	27,300	580,700.00
29-Jan-02	21.25	20.50	20.50	21.50	3	2,900	60,950.00
31-Jan-02	21.00	21.00	21.00	20.50	2	5,000	105,000.00
01-Feb-02	20.50	20.50	20.50	21.00	3	1,800	36,900.00
05-Feb-02	20.50	20.00	20.25	20.50	14	11,300	229,725.00
06-Feb-02	20.50	20.25	20.50	20.25	6	11,600	237,550.00
07-Feb-02	20.50	20.50	20.50	20.50	1	100	2,050.00
08-Feb-02	20.50	20.50	20.50	20.50	3	2,500	51,250.00
11-Feb-02	20.50	20.50	20.50	20.50	3	2,200	45,100.00
12-Feb-02	20.75	20.50	20.50	20.50	2	1,500	30,875.00

Fig. 3. Historical Data set structure for a bank (Commercial Bank)

Date	ASPI	Sector_Index
2002-01-02 00:00:00	610.70	1,330.50
2002-01-03 00:00:00	611.90	1,345.40
2002-01-04 00:00:00	609.60	1,335.00
2002-01-07 00:00:00	603.50	1,318.30
2002-01-08 00:00:00	596.60	1,304.30
2002-01-09 00:00:00	592.20	1,295.10
2002-01-10 00:00:00	585.10	1,271.40
2002-01-11 00:00:00	577.70	1,251.70
2002-01-15 00:00:00	566.50	1,221.30
2002-01-16 00:00:00	563.50	1,197.70
2002-01-17 00:00:00	549.50	1,166.60
2002-01-18 00:00:00	566.70	1,218.10
2002-01-21 00:00:00	570.10	1,225.70
2002-01-22 00:00:00	561.90	1,193.70
2002-01-23 00:00:00	558.20	1,189.20
2002-01-24 00:00:00	567.50	1,212.20
2002-01-25 00:00:00	571.10	1,218.40
2002-01-30 00:00:00	564.60	1,206.40
2002-01-31 00:00:00	562.20	1,203.80
2002-02-01 00:00:00	564.30	1,200.30
2002-02-05 00:00:00	565.10	1,205.10

Fig. 4. ASPI and Banking Sector Indices

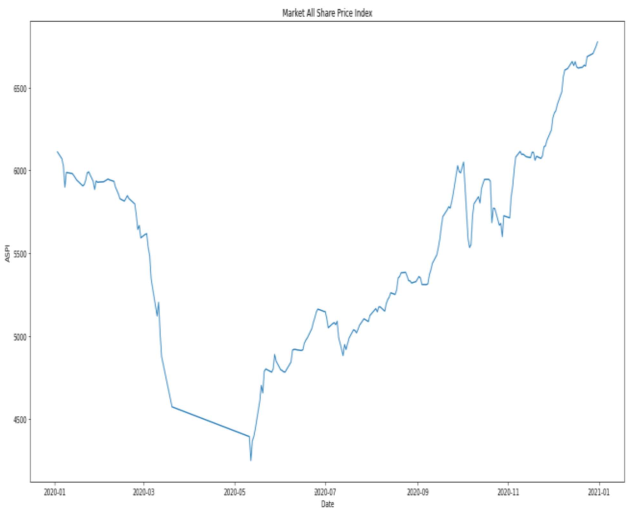

Fig. 5.All Share Price Index (ASPI) in COVID-19 Lockdown Situation

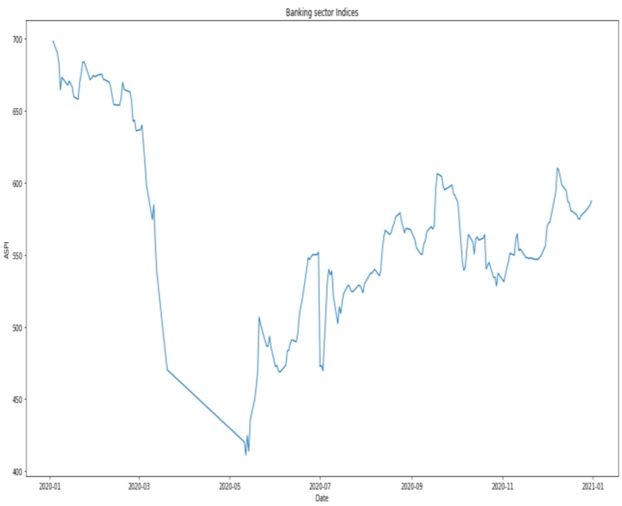

Fig. 6. Banking Sector Indices in COVID-19 Lockdown Situation

The ASPI and Banking sector indices in the COVID-19 period is shown in Fig.5 and Fig.6 respectively. As mentioned in the above hypothesis, Fig.5 and Fig.6 show an idea of how news and other sentiments related to the ASPI and Banking Sector Indices. Downslope can be seen in Fig.5 and Fig.6 during the COVID-19 lockdown period starting from March 2020. According to the hypothesis, ASPI and Banking Sector Indices have been taken as a numerical sentiment parameter other than the historical prices from each bank. Since ASPI shows overall market performance and Banking sector indices show banking sector capitalization changed over time. There can be a positive and negative sentiment when the ASPI and Sector Indices increase and decreases respectively.

3) Detailed architecture

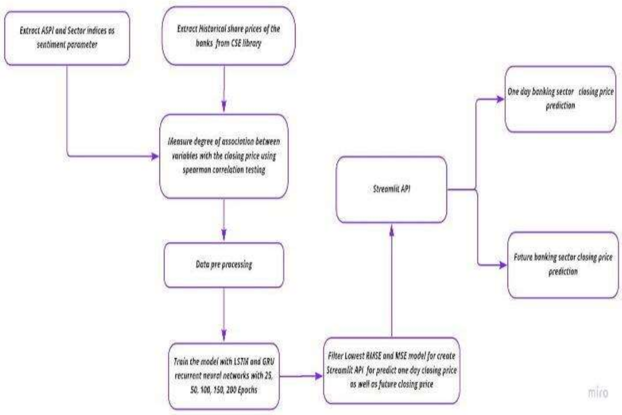

Fig. 7. The detailed architecture of the predictor model

4) Data preprocessing

In this research, data pre-processing is done to detect whether the dataset has null values. Pandas data frame was used to remove null value rows from the dataset. In some cases, closing prices of the previous day for the relevant stock have been considered for opening prices of that stock in the next day. Because there was not any column for opening prices in some years. Data pre-processing part is the most important process unless model validation cannot happen with null or empty values. Then spearman correlation test was performed for each column according to the historical dataset. Then extracted those columns to train the model according to the p value and r value. Then a decision can be taken whether the need to add that column into the trained model or not. Because Spearman correlation ranking is used to measure the degree of association between two variables. The total dataset has been divided into 70 % for training and 30 % for testing. Due to large values in the dataset, Machine Learning models need to use scalers. These scalers make it easy to learn and understand the problem. Then MinMaxScaler is used as a scaler for data pre-processing.

5) Building the algorithm

This research used RNNs such as LSTM and GRU. The LSTM algorithm is used for remembering the sequence for a long duration of time. GRU is also used for sequence analysis and time series analysis like LSTM. But there are some differences in the gates between LSTM and GRU models. When building LSTM and GRU models opening, high, low, trades, volume, turnover, ASPI, and Banking Sector Indices were taken as independent input data and closing price as the dependent data. ASPI and Banking Sector Indices were used as a numerical sentiment parameter. Finally, the model was trained using LSTM and GRU models with different numbers of epochs such as 25, 50, 100, 150, and 200 epochs for each bank with a batch size of 16. Total training times were 5 per bank and 40 times for all four selected banks using those two algorithms.

6) Evaluation of the model

Mean squared error (MSE) and Root Mean Square Error (RMSE) are used as evaluation criteria. A model validation graph is also used to evaluate the model and decide how the model is fitted.

7) Building the graphical user interface with streamlit

Streamlit is an open-source app framework for the python language which is used to develop apps or web applications for data science and machine learning projects in a minimal period.

Streamlit is very familiar and compatible with most of the python libraries. Therefore, GUI can be created quickly and effortlessly.

V. Results and Evaluation

	MSE VALUE FOR LSTM AND GRU WITH DIFERENT NO OF EPOCHS									
	25 Epochs		50 Epochs		100 Epochs		150 Epochs		200 Epochs	
	LSTM	GRU	LSTM	GRU	LSTM	GRU	LSTM	GRU	LSTM	GRU
Commercial bank	8.73648	9.56949	7.24548	8.32860	5.80531	6.03224	5.92386	6.42373	6.12443	6.87463
Sampath Bank	7.42348	7.83247	7.02372	6.93234	6.14385	6.24275	6.73127	7.02643	7.24759	7.12195
HNB Bank	8.34984	8.42746	8.12327	7.63328	7.01324	7.17324	7.42637	7.83633	7.57359	7.96452
Seylan Bank	7.42937	7.73224	6.23377	7.16583	5.81875	6.42738	5.81344	6.32134	6.24566	6.95484

Fig. 8. MSE for LSTM and GRU Architectures

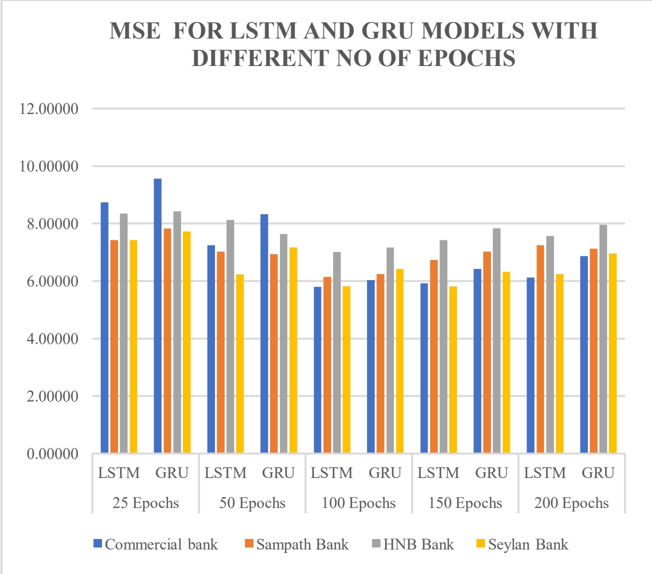

Fig. 9. Chart for Mean Squared Error (MSE) for Different No of Epochs

	RMSE VALUE FOR LSTM AND GRU WITH DIFERENT NO OF EPOCHS									
	25 Epochs		50 Epochs		100 Epochs		150 Epochs		200 Epochs	
	LSTM	GRU	LSTM	GRU	LSTM	GRU	LSTM	GRU	LSTM	GRU
Commercial bank	2.95575	3.09346	2.69174	2.88593	2.40942	2.45606	2.43390	2.53451	2.47476	2.62195
Sampath Bank	2.72461	2.79866	2.65023	2.63293	2.47868	2.49855	2.59447	2.65074	2.69213	2.66870
HNB Bank	2.88961	2.90301	2.85014	2.76284	2.64825	2.67829	2.72514	2.79934	2.75202	2.82215
Seylan Bank	2.72569	2.78069	2.49675	2.67691	2.41221	2.53523	2.41111	2.51423	2.49913	2.63720

Fig. 10. RMSE for LSTM and GRU Architectures

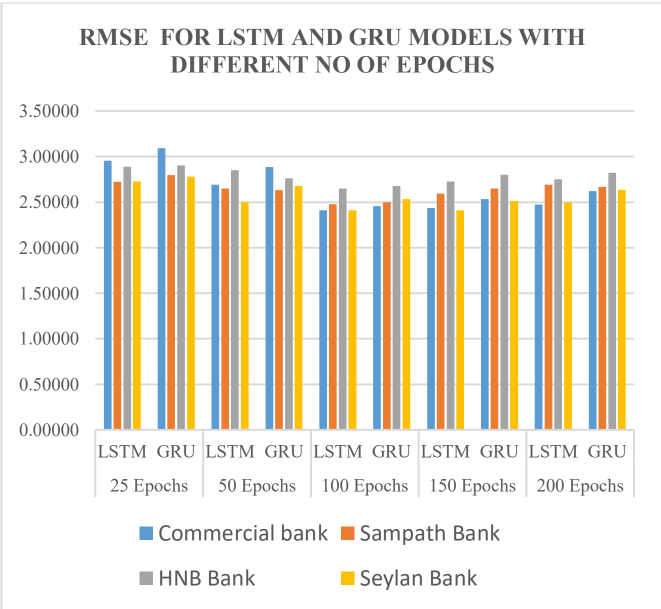

Fig. 11. Chart for Root Mean Squared Error (RMSE) for Different No of Epochs.

According to the results of this research, the evaluation criteria are the MSR and RMSE. The stock market prediction model was trained 40 times with LSTM and GRU models using different no of epochs with historical datasets with numerical sentiments (ASPI and Sector indices). According to the above bar charts, the LSTM model has given the lowest MSE and RMSE values for every training. Finally, commercial bank, HNB bank, and Sampath bank give the lowest MSE and RMSE values for the 100 epochs. Seylan Bank gives the lowest MSE and RMSE value for the 150 epochs. When the epoch is increased over and above the lowest MSE and RMSE values, the model has been overfitted. The lowest MSE and RMSE value model has been selected from each bank for future prediction.

The commercial bank 100 epochs model has the lowest MSE and RMSE value. Then testing dataset vs predicted closing price could be plotted as follows.

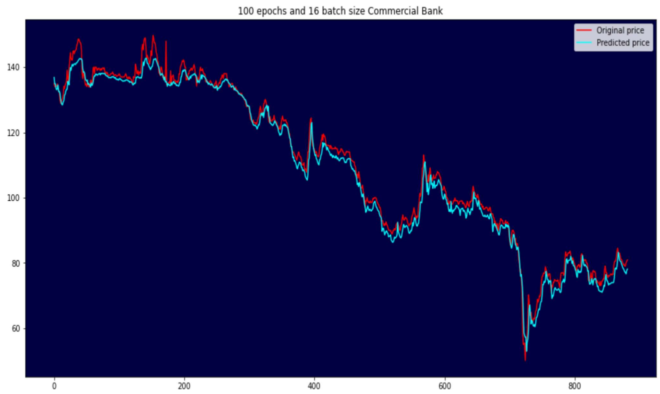

Fig. 12. Original testing closing price vs predicted closing prices with 100 epoch model in COMB.N

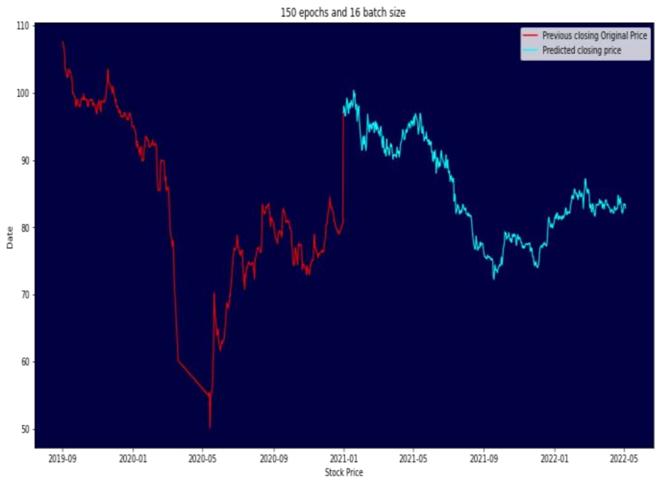

Fig. 13. Previous original closing prices vs future predicted closing prices

Using the above 100 epoch model for the commercial bank, the future closing prices can be plotted as follows.

In this research, the lowest error model has been selected for the prediction of each bank closing price for one day as well as future days. There are 2 types of prediction of each bank; the first one is future value prediction for that as an input need to give how many years need to predict starting from the 2020 January then the result will be come, and which gives closing price for the future days. The drop-down box should be closing price in future then future closing price can be predicted according to the investor opinion.

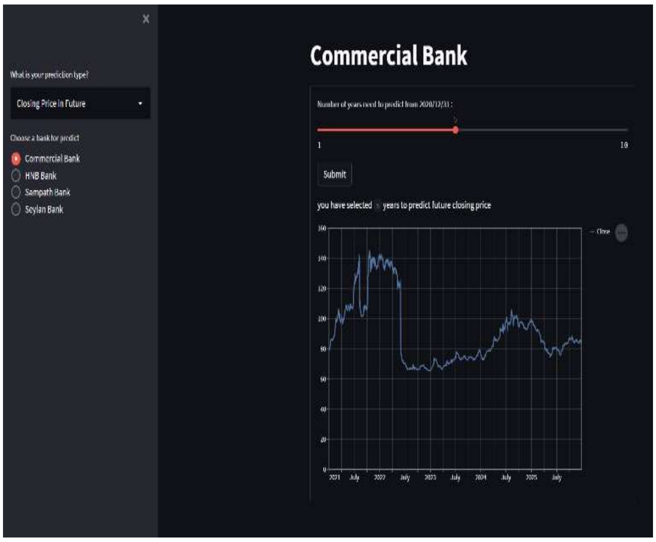

Fig. 14. Future closing price prediction Commercial Bank

Another type is predicting the share price for the one day for that investor needs to fill the current market values according to the form given then the investor can get the closing price for one day, above preferred bank can be selected, and the preferred type should be one-day Prediction in the dropdown box then one-day closing price can be predicted using the existing market values. The following Prediction has been made for the commercial bank; likewise, each bank prediction can be made with this API.

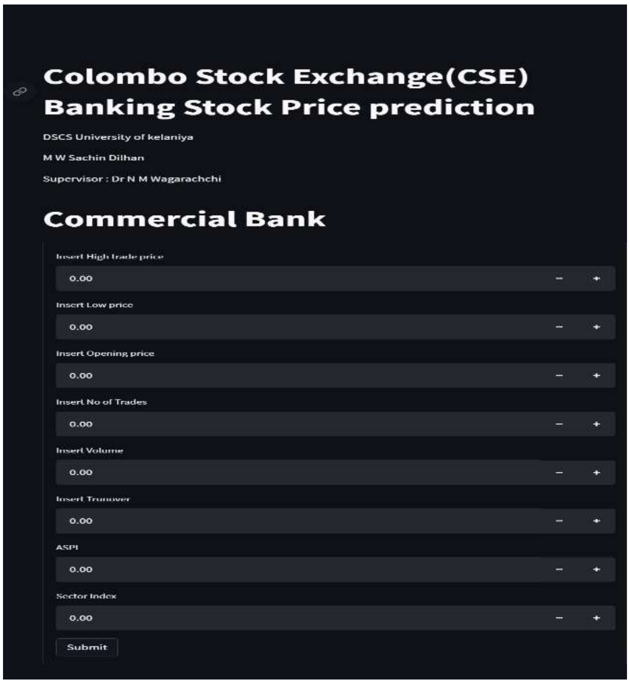

Fig. 15. API form for predict one-day closing price

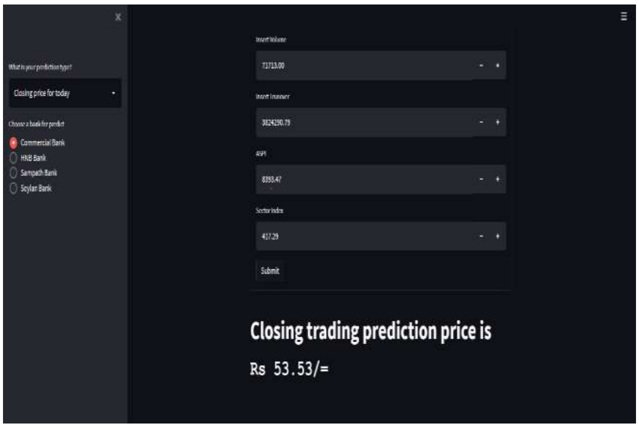

Fig. 16. Output of the one-day closing price prediction model

VI. CONCLUSION

This research aimed to develop a stock market prediction model, particularly for the banking sector, for the historical dataset for the different banks collected from the Colombo Stock Exchange (CSE) data library that mainly focused on HNB COMMERCIAL, CEYLAN, and SAMPATH banks. In this prediction, investors need to know the future closing price of a particular banking share or what will the one-day closing price of a specific share. Prediction of stock prices can also have many advantages for the economy. Here prediction can be made in 2 types one is predicting the future closing value of the bank, and the other on in predict one-day closing price according to the current market performance. Spearman correlation has been used to check the association between two variables with the closing price of the banking sector share. According to the hypothesis, the All-share price index (ASPI) and Banking sector indices are also used because All Share Price Index (ASPI) and Banking Sector indices have been taken as a numerical sentiment parameter other than the historical prices from each bank. Since ASPI shows overall market performance and Banking sector indices show banking sector capitalization changed over time. There can be a positive and negative sentiment when the ASPI and Sector Indices increase and decrease, respectively. ASPI and the Sector index perform a major role because most of the investor's ASPI and Sector indices are taken as a sentiment value, including the total sentiment impact for the overall market and the sector. Finally, models were trained using the LSTM and GRU architectures with 25,50,100,150 and 200 epochs. In a COMB.N, HNB.N and SAMP.N were given the lowest MSE value for the 100 epochs, and SEYB.N gave the lowest MSE value for the 150 epochs. Then the lowest MSE and RMSE value model was used to create the streamlet API for predicting the one-day closing price and future closing price value.

VII. FUTURE WORK

This research can be extended to other sectors, such as manufacturing, consumer durables, and apparel etc. This can be done for these sectors using ASPI and particular sector indices. Twitter or other social media news and text sentiment analysis can be done as future work using some Natural Language Processing techniques.

REFERENCES

[1] Adil Moghar, Mhamed Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network", Procedia Computer Science, Volume 170, 2020, Pages 1168-1173, 2020.

[2] M. Kesavan, J. Karthiraman, R. T. Ebenezer and S. Adhithyan, "Stock Market Prediction with Historical Time Series Data and Sentimental Analysis of Social Media Data," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020.

[3] T. F. A. Samarawickrama, A Recurrent Neural Network Approach in Predicting Daily Stock Prices, University of Sri Jayewardenepura, 2018. .

[4] Shen, J., Shafiq, M.O. Short-term stock market price trend prediction using a comprehensive deep learning system. J Big Data 7, 66, 2020.

[5] Zhao J, Zeng D, Liang S, Kang H, Liu Q. "Prediction model for stock price trend based on recurrent neural network", Journal of Ambient Intelligence and Humanized Computing. 2021 Jan;12(1):745-53.

[6] Sirimevan N, Mamalgaha IG, Jayasekara C, Mayuran YS, Jayawardena C. Stock market prediction using machine learning techniques. In 2019 International Conference on Advancements in Computing (ICAC) 2019 Dec 5 (pp. 192-197). IEEE.

[7] Yan Y, Yang D., "A stock trend forecast algorithm based on deep neural networks", Scientific Programming, Jul 5;2021.

[8] L. L. H. P. S. Yao, "High-Frequency Stock Trend Forecast Using LSTM,", 13th International Conference on Computer Science & Education (ICCSE), 2018.

[9] Kesavan. et. al, "A Machine Learning Model for Stock Market Prediction,", 2020.

[10] R. A. M. Rahman, Forecasting Stock Market Price Using Multiple Machine Learning Technique, Seen at https://www.researchgate.net/publication/348976652_Forecasting_Stock_Market_Price_Using_Multiple_Machine_Learning_Technique, 2021.