*Journal of Statistics* *ISSN-P: 1684–8403*
*Volume 29, 2025, pp. 128–142* *ISSN-E: 2957–8981*

# Long Short-Term Memory Method: A Case Study of Pakistan Stock Market Volatility

*Zahid Iqbal* [1*], *Muhammad Shoaib* [1]

## Abstract

Forecasting volatility of stock market has an important role to minimize risk in financial markets. Investors track stock volatility to avoid market risk. Prediction of stock volatility gets attention of many investors and scholars. This study investigated Pakistan stock market volatility by employing Long Short-Term Memory method. Data contains eight macroeconomic variables named as Inflation, Money-Supply, Exchange-rate, Crude oil, Crop production, Gold prices, GDP growth and Unemployment. The data was obtained from World Development Indicators, the State Bank of Pakistan, the Pakistan Bureau of Statistics, and Investing.com. The study covers the period from January 1, 2000 to June 30, 2023. The Long Short-Term Memory (LSTM) Method is type of recurrent neural network (RNN) and the information in matrices form of the order (5772, 14, 9). It can remember information over a long period of time due to its cell state. Three gates, namely forget gate, input gate, and output gate and one cell state which is used for saving of data. Forget gate is identified by $f_t$ and is responsible to forget information which is not necessary. Input gate is denoted by $i_t$ output gate is denoted by $o_t$ and cell state is denoted by $c_t$. The pivotal element within the LSTM model is the cell state $c_t$, which retains information over extended durations. The LSTM is used for prediction purposes. The results are produced using LSTM method in Spyder in python. The long short-term memory method has mean squared error of 0.01951.

**Keywords:** Artificial neural network, LSTM, Pakistan stock market, Volatility, Forecasting.

## 1. Introduction

Forecasting stock volatility has a crucial part to minimize risk in financial markets and other finance related fields. Prediction of stock volatility takes place in decision making of the participants. Investors track stock volatility to avoid market risk. Thus prediction of stock volatility gets attention of many investors and scholars. Volatility in finance related markets is a matter of significant concern. Volatility forecasting of stocks plays a crucial part in investors' decision-making regarding their investments. Policy makers also see the volatility because the goal is stable stock market as effectiveness of the market.

*Corresponding author
[1]Department of Statistics, Allama Iqbal Open University, Islamabad, Pakistan.
Email: zahid.iqbal@aiou.edu.pk

Volatility is often known as measure of uncertainty. The higher volatility the higher risk involved. A lower volatility means that stock market does not fluctuate dramatically. Numerous factors influence volatility of stock markets. We are examining macroeconomic variables that are affecting stock market volatility using two different methods. The macroeconomic variables are inflation, gold prices, crop production, real-effective-exchange-rate, unemployment rate, interest-rate, gross domestic product, crude oil prices and money supply.

Inflation means that rise in prices that decline the purchasing power over period of time. High inflation causes uncertainty that slowdown the economic growth. On the other hand low inflation rate reflects the growth of an economy. Stock market tends to be more volatile when inflation is high. Volatility in gold prices has inverse relation in stock markets. Low volatility in gold prices means safe conditions for investment while increase in gold prices cause the investors withdraw their money from stock market that tends to decrease in stock index.

The significance of agriculture in economic development is considerably greater. By using modern machinery lands produces high quality crop production. Due to high quality and huge crop production we can export the crops that cause foreign exchange and economic development. Exchange rate is fundamentally defined as the comparative value of one nation's currency against that of another. Exchange rates are used for foreign trade purpose. A balanced stock market attract investor's capital which causes demand of country currency. There is a tandem relation between them. Gross domestic product usually measures personal consumptions, investments, government spending and net exports. Investors are optimistic about the economy based on GDP numbers.

In recent years, emerging economies have experienced significant expansion in their stock markets, presenting profitable opportunities for investment (Bosworth & Collin, 1999; Stiglitz, 2000). Pakistan is a developing country facing numerous challenges on the economic and political front. Pakistan's stock market performance has recorded periods of boom and bust in the recent past.

Predicting the stock market is tricky because of its dynamic, stochastic, and uncertain nature (Khan et al., 2020; Parmar et al., 2018). Two primary analytical methodologies are employed in forecasting stock market trends: technical analysis and fundamental analysis. A researcher identifies the primary elements influencing the stock market in basic analysis. On the other hand, in technical analysis, a researcher can predict stock market trends by using stock prices and their trends. Predicting the stock market trend is a difficult task due to the nonlinear pattern of the stock market (Rather et al., 2015; Vaisla & Bhatt, 2010). The objective of the study is to Predict the stock market volatility using Long Short–Term Memory method.

The paper is structured as follows: Section 2 gives an overview of the recent studies, employing different methodologies to forecast the various economic variables. Section 3 explains the methodology used in this study. Results are discussed in Section 4. Section 5 concludes the study.

## 2. Review of literature

Bhandari et al. (2022) predicted closing price of S&P500 using long short-term memory method. The macroeconomic variables employed include EFFR (Effective Federal Funds Rate), CUR (Civilian Unemployment Rate), CSI (Consumer Sentiment Index), and US Dollar Index. Data was manipulate from 2006 to 2020. The outcome indicates that the LSTM model offers a superior fit and achieves high levels of prediction accuracy.

Kim and Won (2018) employed LSTM technique to forecast KOSPI (KOREAN COMPOSITE STOCK PRICE INDEX). The data was used from January 1 2001 to September 30 2011. The macroeconomic variables used are gold and crude oil prices. The findings suggest LSTM method can effectively forecast time series data. Ku et al.(2023) forecasting Malaysia stock market volatility by applying LSTM method. The historical data were taken from yahoo finance. The dataset comprises open, close, high, low, volume, and adjusted closing prices of Malaysian stock market. The data was taken from 1 Jan 2019 to 30 June 2022. Outcomes demonstrate that proposed method outperforms the prediction approach.

Wang (2024) endeavors to forecast Shanghai Stock Exchange 50 Index using an LSTM technique. The data was used from August 2013 to August 2023. The data was gathered from Yahoo Finance. The main indicators was utilized are opening price, minimum, maximum, closing-price to predict next day closing-price of SSE. LSTM method correctly predict SSE 50 index closing prices.

Zhang et al. (2024) analyzed impacts of seven variability indices on volatility in Chinese energy prices through the utilization of LSTM methodology. The indicators are EPU (Economic Policy Uncertainty), GRI (Geopolitical Risk Index), CPU (Climate Policy Uncertainty). The data was used from Jan 2009 to Aug 2022. The consequence indicates that EPU significantly influences Chinese energy price volatility.

Abuein et al. (2024) forecasted Japan stock exchange (JSE) opening price by apply LSTM model. They utilized five variables: opening price, high, low, closing, and trading volume. The data was used from January 4, 2017 to May 27, 2022. The result shows that LSTM method correctly forecast the opening price of Japan stock exchange.

You (2024) examined the volatility of S&P500 by using long short term memory method. Three variables were used open price, close price and volume to predict open price of S&P500. The daily data was used from January 1, 2000 to December 31, 2015. The LSTM model achieves remarkable 97.5% accuracy. Chen (2024) used LSTM technique to forecast the closing price of S&P500. The data of five years was collected from yahoo finance. The open price, minimum, maximum and close price were used. The data was used from December 1, 2018 to June 30, 2021. The LSTM method gains maximum forecasting accuracy. Gupta et al. (2024) employed Long Short-Term Memory method to forecast opening price of Nifty 50 index. They used historical data of Nifty 50 from July 15, 2003 to April 30, 2021. Data was acquired from Yahoo Finance. Results reveal a notable correlation between daily returns and daily closing prices.

In summary, the empirical investigation into predictive modelling of stock price fluctuations in the Pakistan Stock Market utilizes sophisticated machine learning models and discerns crucial technical features impacting prediction precision. The study suggests potential avenues for future research aimed at improving market analysis. The results enhance comprehension regarding the prediction of stock prices, offering vital perspectives for investors navigating the intricate Pakistan Stock Market. The findings emphasize the potential effectiveness of machine learning algorithms in forecasting changes in stock prices, aligning with the purpose of the research. One crucial factor in improving the precision of these models involves incorporating a wide range of data sources, corresponding closely to the inclusion of different technical indicators and features in the research.

## 3. Material and methods

The LSTM technique is a renowned deep learning method within RNNs utilized for time series prediction. LSTMs are employed for both regression and classification tasks, extending beyond stock market prediction. While RNNs excel in retaining information compared to traditional networks, they encounter challenges in learning long term dependencies, primarily due to the vanishing gradient problem. LSTM method leverages memory cells to address the challenge of vanishing gradients. LSTM comprises four elements: forget gate, input gate, output gate, and cell state. Forget gate is identified by $f_t$ and is responsible to forget information which is not necessary. Input gate is denoted by $i_t$ , output gate is denoted by $o_t$ and cell state is denoted by $c_t$. The pivotal element within the LSTM model is the cell state $c_t$, which retains information over extended durations.

### 3.1. Artificial neural network method

Following are some important advantages of Artificial Neural Network which make them beneficial for prediction.

- Artificial Neural Networks (ANN) can generalize and correctly predict the unseen part of the data.
- The Neural Networks are known as universal functional approximates and produce a continuous function for a required accuracy.
- Artificial Neural Networks are nonlinear. The real world systems are usually nonlinear.

An artificial neural network (ANN) is constructed using a network of interconnected units or nodes known as artificial neurons, which loosely emulate the neurons present in a biological brain. Each connection, similar to synapses in a biological brain, can transmit signals to other neurons. It receives signals, processing them, and then transmits signals to neurons connected to it. Signal transmitted through a real number, and each neuron's outcome is measured by applying a non-linear function to sum of its inputs. The connections are referred to as edges. Edges and Neurons usually possess weight adapts as learning progresses. Weight either amplifies or diminishes power of signal transmitted through connection. Neurons may include a threshold, ensuring that a signal is sent only when the combined signal exceeds that threshold. Usually, neurons organized into layers. Different layers can perform various transformations on their inputs. Signals propagate from the initial layer to final layer, possibly passing through multiple layers multiple times. These networks are trained through the processing of examples, each consisting of a known "input" and "result", establishing connections among the two, which stored within the network's data structure. Training of a neural network with example typically involves calculating the disparity among a prediction and a target output. Variance represents error. Subsequently, the network adjusts its weighted connections according to a learning rule, utilizing this error value. Continuous adjustments will lead to the neural network producing output that increasingly resembles the target output. Once a sufficient number of these adjustments have been made, the training can be stopped based on specific criteria. This constitutes a type of supervised learning.This constitutes a type of supervised learning.

Multiple layers are arranged in neurons particularly in deep learning. Neurons within unit layer exclusively link to neurons in preceding and succeeding layers. Input layer receives external data. Layer responsible for producing final outcome is known as output layer. Among them, may exist zero or more hidden layers. Single layer and non-layered networks are also employed. Among two layers, various connections are feasible. They may be 'fully connected', indicating that every neuron in one layer connects to every neuron in the next layer. They may include pooling, where a group of neurons in one layer connects to a single neuron in the next layer, effectively decreasing the number of neurons in that layer. Neurons linked solely in this manner create a feed forward networks. In contrast, networks that allow connections between neurons in the same or previous layers are referred to as recurrent networks. The structure of human neuron structure of artificial neural networks (ANN) are shown in Figures 1 and 2, respectively.

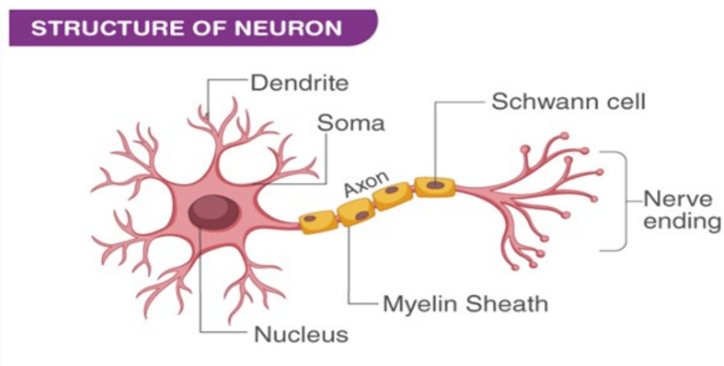



Figure 1: The structure of neuron.

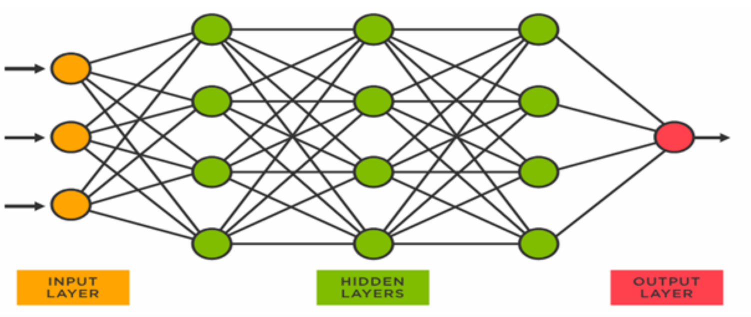



Figure 2: The structure of artificial neural network (ANN).

### 3.2. Long Short-Term Memory (LSTM) method

The LSTM model was introduced by Hochreiter and Schmidhuber in 1997. LSTM is type of RNN which remember information over long period of time. LSTM has memory blocks which contain memory cells. LSTM has three gates, the forget gate is denoted by $f(t)$, the input gate by $i_t$, and the output gate by $o_t$ as shown in Figure 3. $X_t$ represents the vector of input variables, $c_{t-1}$ denotes cell state

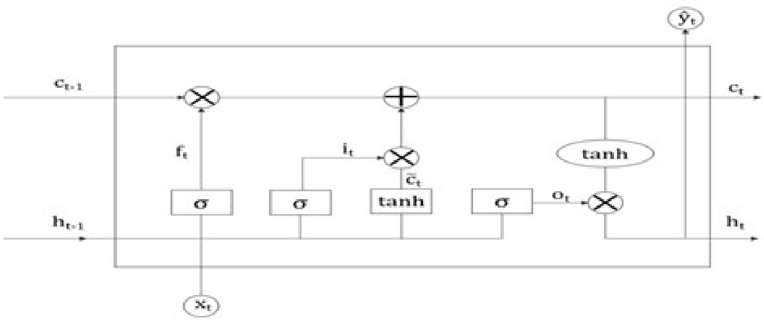



Figure 3: The structure of LSTM.

at $(t-1)$ time, $c_t$ represents cell-state whereas $h_{t-1}$ indicates hidden state at $(t-1)$ time and $h_t$ is hidden state. The unnecessary information has been removed by forget gate. The observations from $h_{t-1}$ and $X_t$ are passed through sigmoid function then unnecessary information removed through forget gate and other observations retain. One implies that the value should be retain and zero response from sigmoid function means the observations should be discorded. Next step is the observations pass through input gate and added in cell-state, where activation function is tanh which update candidate values, $\tilde{c}_,$. In equation form entire process described as

$$f_t = \sigma\left(W_f h_{t-1} + N_f x_t + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_i h_{t-1} + N_i x_t + b_i\right) \tag{2}$$

$$\tilde{c}_t = \tanh\left(W_c h_{t-1} + N_c x_t + b_c\right) \tag{3}$$

$$c_t = f_t c_(t-1) + i_t \tilde{c}_t \tag{4}$$

$$o_t = \sigma\left(W_o h_{t-1} + N_o x_t + b_o\right) \tag{5}$$

$$h_t = o_t \tanh(c_t) \tag{6}$$

$$\hat{y}_t = h_t \tag{7}$$

Where $W_f$ , $W_i$ , $W_o$, $W_c$, $N_f$, $N_i$, $N_o$, $N_c$ are the weights and $b_f$, $b_i$, $b_o$, $b_c$ are the biases corresponding to the forget gate, input gate, and output gate, respectively. Sigmoid activation function is represented by $\sigma$ and tanh is a nonlinear function to create vector of candidates values $\tilde{c}$ and update cell state $c_t$. $\hat{y}_t$ represents the final output of the LSTM model. Hence, forecasting stock market volatility gains much attention around the globe.

### 3.3. Activation function

Activation function dictates whether a neuron should activate by computing the weighted sum and subsequently incorporating bias. Activation function serves non-linearity into output of neuron. The query arises as to why a neural network necessitates a non-linear activation function. Without an activation function, a neural network essentially functions as a linear regression model. It conducts a non-linear transformation on input, enabling it to learn and execute more intricate tasks. The structure of activation function is displayed in Figure 4.

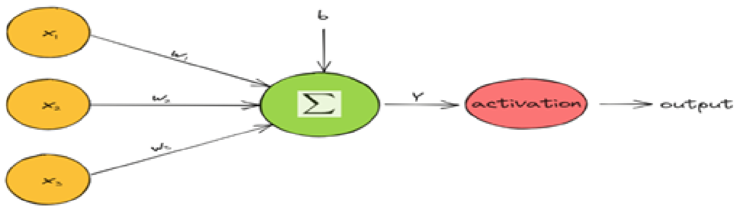



Figure 4: Structure of activation function.

$$\text{Where}, Y = \sum_{i=1}^{m}(x_i * w_i) + b. \tag{8}$$

#### 3.3.1. Sigmoid activation function

The primary reason for using the sigmoid activation function is its output range, which lies between 0 and 1. It is particularly utilized in models where predicting probabilities as output is required. Since probabilities always fall within the range of 0 to 1. Thus, sigmoid function is the appropriate choice. Sigmoid activation function is also called Logistic activation function. The sigmoid activation function curve has an S-shape (Figure 5). The formula of sigmoid activation function is given by (9).

$$\sigma(x) = 1/(1 + e^{-x}). \tag{9}$$

#### 3.3.2. tanh activation function

Tanh activation function is an activation function used for neural networks. It is also known as hyperbolic tangent activation function. Range of the tanh activation function spans from -1 to 1. It also has an S-shaped curve. It transforms the input values to produce output values between -1 and 1. It gives better performance than sigmoid activation function. The formula of tanh activation function is given by (10) and its diagram is given by 6. A comparison of both activation function is shown in Figure 7.

$$\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}) \tag{10}$$

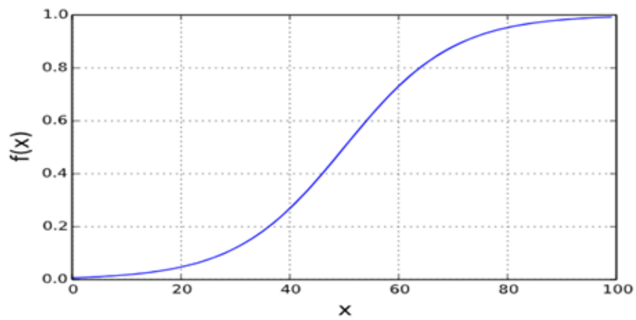



Figure 5: Curve of sigmoid activation function.

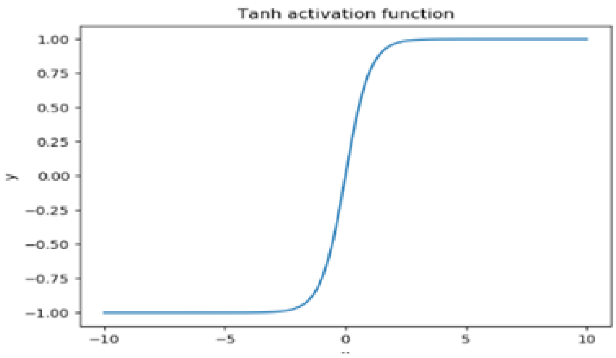



Figure 6: Curve of tanh activation function.

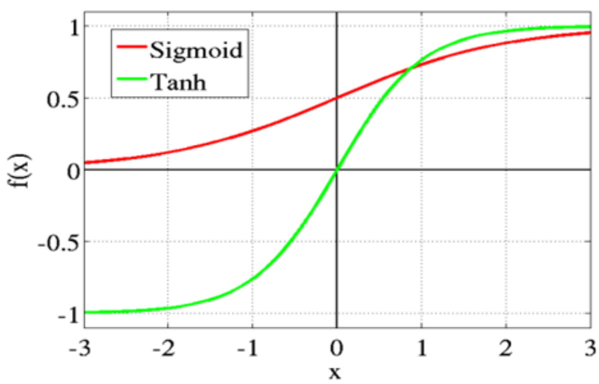



Figure 7: Comparison of Sigmoid and tanh activation functions.

## 4. Results and discussion

The LSTM method is employed to forecast Pakistan stock market volatility. Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to handle sequential data like time series and in our study Pakistan Stock Market volatility (Open Prices, Daily data). They use three gates—forget, input, and output—and a cell state to manage information flow, allowing them to retain both short-term and long-term dependencies, which is a common issue with traditional RNNs. This architecture helps overcome the vanishing and exploding gradient problems common in standard RNNs. We use `Python` (van Rossum & Fred L. Drake, 2009) with `Spyder` (Team, 2025) to analyze the data. In our study, research variable is Daily Open Prices for the period from January 3, 2000 to May 30, 2023. The independent variables are Gold, Crude Oil, Crop Production, Money Supply, Exchange Rate, GDP Growth, Unemployment. A sample data is displayed in Table 1. We employ log likelihood function for variable selection

Table 1: A sample set of original data used.

| Index | Daily | Open price | Monthly | Gold | Crude Oil | Years | Crop production | Inflation | Money supply | Exchange Rate | GDP Growth | Unemp-loyment |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1/3/2000 | 1422.90 | 1/1/2000 | 4740 | 1310.78 | 2000 | 73.43 | 4.14264 | 1.32e+12 | 49.3007 | 3.9 | 3.9 |
| 1 | 1/4/2000 | 1457.07 | 2/1/2000 | 5000 | 1409.95 | 2001 | 66.98 | 4.36666 | 1.48e+12 | 53.6482 | 2.0 | 7.0 |
| 2 | 1/6/2000 | 1479.70 | 3/1/2000 | 4770 | 1424.02 | 2002 | 67.75 | 3.14826 | 1.69e+12 | 61.9272 | 3.1 | 8.3 |
| 3 | 1/12/2000 | 1531.28 | 4/1/2000 | 4660 | 1215.56 | 2003 | 71.90 | 3.29034 | 1.99e+12 | 59.7238 | 4.7 | 8.3 |
| 4 | 1/13/2000 | 1570.84 | 5/1/2000 | 4590 | 1408.20 | 2004 | 75.09 | 2.91413 | 2.34e+12 | 57.7520 | 7.5 | 7.7 |
| 5 | 1/14/2000 | 1581.59 | 6/1/2000 | 4760 | 1534.07 | 2005 | 80.73 | 7.44462 | 2.70+e12 | 58.2579 | 9.0 | 7.2 |
| 6 | 1/17/2000 | 1626.09 | 7/1/2000 | 4690 | 1459.49 | 2006 | 78.57 | 9.06333 | 3.24e+12 | 59.5145 | 5.8 | 6.2 |
| 7 | 1/18/2000 | 1661.12 | 8/1/2000 | 4570 | 1515.42 | 2007 | 85.11 | 7.92100 | 4.93e+12 | 60.2723 | 5.5 | 9.2 |
| 8 | 1/19/2000 | 1640.05 | 9/1/2000 | 4560 | 1661.47 | 2008 | 90.86 | 7.59868 | 5.44e+12 | 60.7385 | 5.0 | 5.2 |
| 9 | 1/20/2000 | 1718.36 | 10/1/2000 | 4590 | 1658.77 | 2009 | 90.13 | 20.28610 | 5.79e+12 | 70.4080 | 0.4 | 5.5 |
| 10 | 1/21/2000 | 1719.02 | 11/1/2000 | 4850 | 1833.39 | 2010 | 87.96 | 13.64780 | 6.81e+12 | 81.7129 | 2.6 | 5.6 |
| 11 | 1/24/2000 | 1852.45 | 12/1/2000 | 5070 | 1464.27 | 2011 | 94.60 | 12.93890 | 7.81e+12 | 85.1938 | 3.6 | 6.0 |
| 12 | 1/25/2000 | 1789.64 | 1/1/2001 | 5050 | 1535.58 | 2012 | 90.59 | 11.91610 | 8.79e+12 | 86.3434 | 3.8 | 6.0 |
| 13 | 1/26/2000 | 1685.70 | 2/1/2001 | 5040 | 1626.65 | 2013 | 98.95 | 9.68235 | 1.03e+13 | 93.3952 | 3.7 | 6.0 |
| 14 | 1/27/2000 | 1711.42 | 3/1/2001 | 5123 | 1512.80 | 2014 | 98.53 | 7.69216 | 1.17e+13 | 101.6290 | 4.1 | 6.0 |
| 15 | 1/28/2000 | 1729.08 | 4/1/2001 | 5120 | 1570.33 | 2015 | 90.47 | 7.18938 | 1.30e+13 | 101.1000 | 4.1 | 5.9 |
| 16 | 1/31/2000 | 1760.79 | 5/1/2001 | 5300 | 1692.98 | 2016 | 102.00 | 2.52933 | 1.46e+13 | 102.7690 | 4.6 | 5.9 |
| 17 | 2/1/2000 | 1776.84 | 6/1/2001 | 5519 | 1711.39 | 2017 | 105.71 | 3.76512 | 1.66e+13 | 104.7690 | 4.6 | 5.8 |
| 18 | 2/2/2000 | 1753.10 | 7/1/2001 | 5470 | 1576.52 | 2018 | 103.13 | 4.08537 | 1.83e+13 | 105.4550 | 6.1 | 5.6 |
| 19 | 2/3/2000 | 1803.63 | 8/1/2001 | 5620 | 1657.18 | 2019 | 104.08 | 5.07806 | 2.01e+13 | 121.8240 | 3.1 | 6.9 |

when estimating GM models. The log likelihood value serves as an indicator of the models' goodness of fit. The higher value of log likelihood, better the model will be and also find the Akaike information criterion (AIC) for every GM model. AIC is employed to compare different possible models to choose the best model. Results are given in Table 2 below. First column lists the macroeconomic variables considered in this study.

In Table 1 sample original data is displayed with dates. To apply different statistical procedures and to normalize the data, we ignore the columns of Daily, Monthly and Years from Table 1.

We have 5786 total observations. In order to apply long short term memory method our input should be of type (batch size, time steps, Features). The batch size denotes the number of samples contained within each batch, time steps signify the count of time steps within each sequence, and features represent the quantity of input variables. Train $X$ in Table 3 is training data set which include all variables

Table 2: Comparison of different GM models.

| GM models | LLF | AIC | BIC | RMSE |
|---|---|---|---|---|
| Crude oil | 21640.2 | -43268.4 | -43228.4 | 8.157 $e^{-05}$ |
| Gold prices | 21843.7 | -43675.4 | -43635.4 | 8.159$e^{-05}$ |
| Crop production | 20882.7 | -41753.4 | -41713.5 | 8.175$e^{-05}$ |
| Inflation | 21777.1 | -43542.3 | -43502.3 | 8.155$e^{-05}$ |
| Money supply | 21746.5 | -43481.0 | -43441.0 | 8.164 $e^{-05}$ |
| Exchange rate | 21967.7 | -43923.5 | -43883.5 | 8.154$e^{-05}$ |
| GDP growth | 21905.3 | -43798.6 | -43758.6 | 8.159$e^{-05}$ |
| Unemployment | 21777.1 | -43542.3 | -43502.3 | 8.155$e^{-05}$ |

normalized data that have the order (5772, 14, 9). If we plus 5772 and 14 we will get the same number of observations that is 5786. The '14' means fourteen working days and '9' shows the total number of variables.

Now we explain how LSTM works to predict matrix $Y$. We have 5786 data points, LSTM model requires the information in matrices form of the order (5772, 14, 9). This explains that there are 5772 matrix of order 14 by 9, where 14 represents number of days and 9 represents the number of variables. As LSTM works the resultant matrix which named as Train Y of order (5772,1). The Train $Y$ matrix is given in Table 4. The predicted scaled values are given in Train $Y$ matrix also called $Y$ predicted scaled values which are shown in Table 5.

Table 3: Train $X$ scaled data.

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | -1.22352 | 0.220079 | 0.554097 | 12.2122 | 6.68779 | 1.38948 | 7.47467 | 13.0064 | 9.28462 |
| 1 | -1.22138 | 0.249547 | 0.609037 | 11.1339 | 7.05247 | 1.56399 | 8.10592 | 6.64219 | 18.8757 |
| 2 | -1.21997 | 0.190536 | 0.617743 | 11.2626 | 5.86909 | 1.79303 | 9.38597 | 10.3268 | 20.0898 |
| 3 | -1.21674 | 0.221473 | 0.500578 | 11.9564 | 5.30038 | 2.12023 | 9.73051 | 15.6861 | 20.0898 |
| 4 | -1.21427 | 0.215693 | 0.608853 | 12.4897 | 4.69797 | 2.50196 | 8.80751 | 23.005 | 18.0329 |
| 5 | -1.21359 | 0.229730 | 0.679600 | 13.4326 | 12.0629 | 2.98186 | 8.99876 | 30.0894 | 18.6329 |
| 6 | -1.21081 | 0.223950 | 0.637681 | 13.0715 | 14.6979 | 3.48356 | 9.11396 | 19.3707 | 14.9987 |
| 7 | -1.20861 | 0.214042 | 0.669117 | 14.1648 | 12.8385 | 4.99959 | 9.18506 | 18.5658 | 12.5625 |
| 8 | -1.20868 | 0.213216 | 0.751206 | 15.1261 | 12.3137 | 5.88303 | 10.5568 | 16.692 | 12.5625 |
| 9 | -1.20504 | 0.215693 | 0.749668 | 15.0041 | 12.967 | 6.26476 | 12.5774 | 1.28283 | 13.2910 |
| 10 | -1.20500 | 0.237162 | 0.847835 | 14.6413 | 22.1697 | 7.37724 | 12.9072 | 8.65195 | 13.5339 |
| 11 | -1.19665 | 0.255327 | 0.640568 | 15.7513 | 21.0063 | 8.46791 | 13.0821 | 11.0016 | 14.5051 |
| 12 | -1.20058 | 0.253676 | 0.689448 | 15.081 | 19.5418 | 9.53676 | 13.0056 | 12.6715 | 14.5051 |
| 13 | -1.20708 | 0.252850 | 0.731635 | 16.4786 | 15.7056 | 11.1837 | 14.1554 | 12.3365 | 14.5051 |

LSTM model proceeds with data and produce the structure information in Table 6. This structure represents the LSTM layers. Firstly, lstm_1 denotes the layer whose names as lstm_1. The outer shape means that for each input sequence, it produces the output sequence of length 14 and each element in the sequence has the dimension 64. It has 18944 parameters. The other lstm layer named as lstm, the output shape is (None, 32) which means that it produces a single vector of dimension 32 as output. This typically indicates that the LSTM layer is set

Table 4: Train $Y$ scaled data.

| | Train $Y$ |
|---|---|
| 0 | -1.28547 |
| 1 | -1.20437 |
| 2 | -1.20238 |
| 3 | -1.20238 |
| 4 | -1.20286 |
| 5 | -1.19970 |
| 6 | -1.20049 |
| 7 | -1.19892 |
| 8 | -1.19818 |
| 9 | -1.19646 |
| 10 | -1.20068 |
| 11 | -1.20079 |
| 12 | -1.28465 |
| 13 | -1.19968 |
| 14 | -1.19812 |
| 15 | -1.19774 |
| 16 | -1.19435 |
| 17 | -1.19081 |
| 18 | -1.18877 |
| 19 | -1.18587 |

Table 5: Predicted $Y$ scaled data.

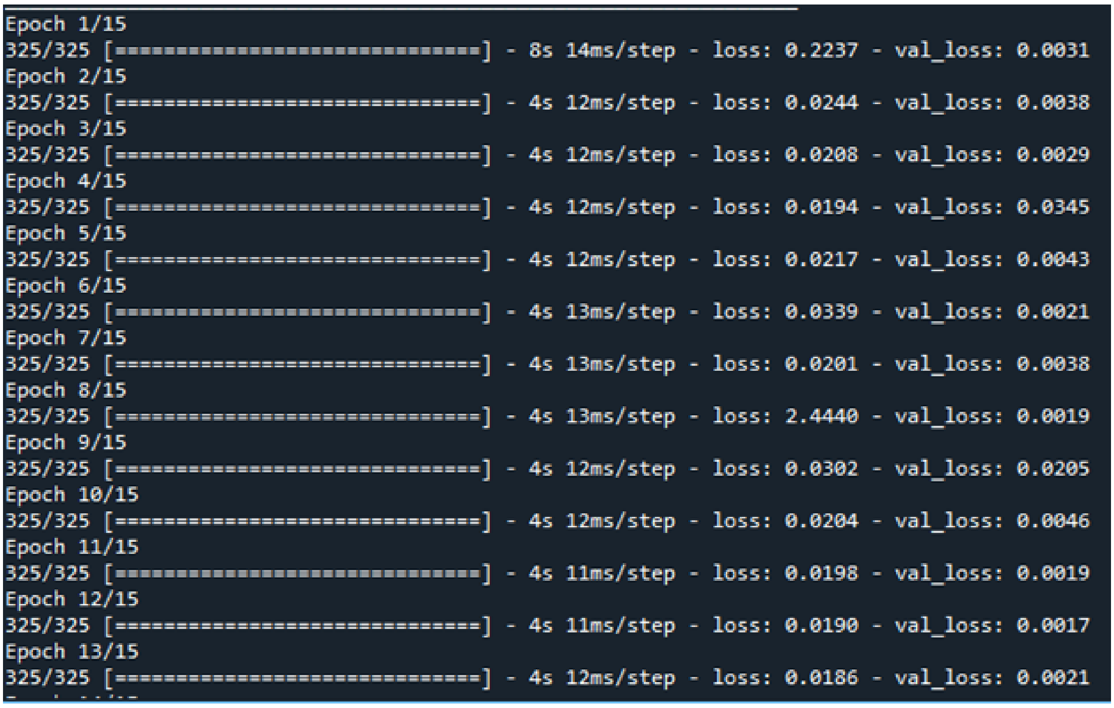
| | Predicted $Y$ ($\hat{Y}$) |
|---|---|
| 0 | -1.20176 |
| 1 | -1.31467 |
| 2 | -1.71544 |
| 3 | -1.34504 |
| 4 | -1.31295 |
| 5 | -1.31660 |
| 6 | -1.28851 |
| 7 | -1.26820 |
| 8 | -1.41761 |
| 9 | -1.46593 |
| 10 | -1.35944 |
| 11 | -0.94289 |
| 12 | -1.06619 |
| 13 | -1.56949 |
| 14 | -1.24170 |
| 15 | -1.11742 |
| 16 | -1.12451 |
| 17 | -1.14549 |
| 18 | -1.12390 |
| 19 | -1.14334 |

to return only the final hidden state of the sequence. It has 12416 parameters. The next layer is dropout layer which is used to prevent over fitting. The output shapes remains the same(None, 32) which indicates after applying dropout layer there are no new parameters are produced as parameters is 0. The next layer is known as flatten layer which is used to convert a multi-dimensional input into a one dimensional vector. The output shape remain the same and produce zero parameter. The final layer, known as dense layer, is a fully connected layer in a neural network that adjusts output dimensionality from preceding layer. Dense layers help to define the relationship between the data values in which the model is working. The output shape is (None, 1) indicating a single scalar value as output. It has total 33 parameters. Thus, as a result we have total 31,393 and all are trainable. The main purpose to train the data is to obtain correct forecast which is possible when the training loss is minimum. The results in Figure 8 depict how the LSTM works to reduce the loss. The outputs are displayed in Figure 9.

In Figure 9, epochs represent number of iterations over the entire data set, basically back propagation is used to reduce loss function. Now we explain the learning curves, both curves represent the loss. One curve is known as validation loss curve while other is described as training loss curve. In the process of dividing data into training and test sets in machine learning, the primary distinction lies in their purposes: training data, a subset of original dataset, is utilized to train model, whereas test data is employed to evaluate the model's accuracy. Typically, the majority of data is allocated for training, while a smaller portion is reserved

Table 6: LSTM structure.

| Layer (type) | Output Shape | Param # |
|---|---|---|
| lstm_1 (LSTM) | (None, 14, 64) | 18,944 |
| lstm (LSTM) | (None, 32) | 12,416 |
| dropout (Dropout) | (None, 32) | 0 |
| flatten (Flatten) | (None, 32) | 0 |
| dense (Dense) | (None, 1) | 33 |
| Total params | | 31,393 |
| Trainable params | | 31,393 |
| Non-trainable params | | 0 |



```
Epoch 1/15
325/325 [==============================] - 8s 14ms/step - loss: 0.2237 - val_loss: 0.0031
Epoch 2/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0244 - val_loss: 0.0038
Epoch 3/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0208 - val_loss: 0.0029
Epoch 4/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0194 - val_loss: 0.0345
Epoch 5/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0217 - val_loss: 0.0043
Epoch 6/15
325/325 [==============================] - 4s 13ms/step - loss: 0.0339 - val_loss: 0.0021
Epoch 7/15
325/325 [==============================] - 4s 13ms/step - loss: 0.0201 - val_loss: 0.0038
Epoch 8/15
325/325 [==============================] - 4s 13ms/step - loss: 2.4440 - val_loss: 0.0019
Epoch 9/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0302 - val_loss: 0.0205
Epoch 10/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0204 - val_loss: 0.0046
Epoch 11/15
325/325 [==============================] - 4s 11ms/step - loss: 0.0198 - val_loss: 0.0019
Epoch 12/15
325/325 [==============================] - 4s 11ms/step - loss: 0.0190 - val_loss: 0.0017
Epoch 13/15
325/325 [==============================] - 4s 12ms/step - loss: 0.0186 - val_loss: 0.0021
```

Figure 8: Training loss and validation loss.

for testing. As looking on Figure 9, the validation loss is better than training loss, usually in training loss probably high. The main thing to describe that the difference between both curves should be minimum. It is clear from Figure 9 that as number of epochs increases, the difference between curves remains minimum.

Using inverse transformation, the forecasted values of $Y$, $\hat{Y}$, are listed in Table 7, which are very close to observed values $Y$.

## 5. Summary and conclusion

Forecasting stock market volatility has always been discussed on popular media around the globe. Many researchers are working on it since many years. Investors are also keen towards the volatility of stock market as they want to maximize

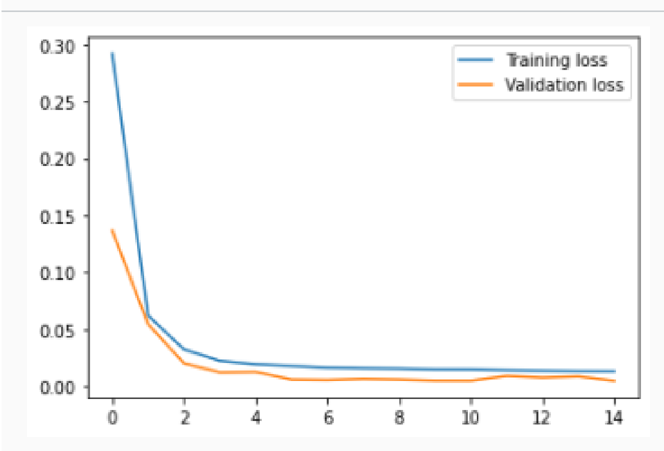



Figure 9: Graphical representation of training loss and validation loss (y-axis) against the number of epochs or training steps (x-axis).

Table 7: Forecasted Open price using LTSM model.

| Index | Day | Open Price ($\hat{Y}$) | Open Price ($Y$) |
|---|---|---|---|
| 0 | 2023-06-01 | 39625.76 | 41388.20 |
| 1 | 2023-06-02 | 39269.45 | 41258.43 |
| 2 | 2023-06-03 | 38902.03 | 41390.47 |
| 3 | 2023-06-04 | 38662.14 | 41752.64 |
| 4 | 2023-06-05 | 38434.26 | 41975.36 |
| 5 | 2023-06-06 | 38368.47 | 42223.62 |
| 6 | 2023-06-07 | 38498.27 | 41735.08 |
| 7 | 2023-06-08 | 38357.45 | 41915.66 |
| 8 | 2023-06-09 | 39205.29 | 41887.60 |
| 9 | 2023-06-10 | 39395.02 | 41594.97 |
| 10 | 2023-06-11 | 39645.13 | 41631.57 |
| 11 | 2023-06-12 | 39827.39 | 41467.91 |
| 12 | 2023-06-13 | 39915.28 | 41411.76 |
| 13 | 2023-06-14 | 39972.79 | 40638.86 |
| 14 | 2023-06-15 | 40119.03 | 40693.22 |
| 15 | 2023-06-16 | 40345.04 | 40231.16 |
| 16 | 2023-06-17 | 40555.21 | 40175.95 |
| 17 | 2023-06-18 | 40179.51 | 40334.05 |
| 18 | 2023-06-19 | 41068.66 | 41431.86 |

their profits. There are so many techniques for the prediction of volatility of the stocks. The primary reason is to minimize the loss and making sound prediction. For this purpose there are so many methods which researchers are using for stock prediction. To assess impact of financial indicators on stock market variability. The technique which is used in this thesis is LSTM method. LSTM model is employed to learn, process, and classify sequential data due to its ability to capture long-term dependencies between data time steps. LSTM technique is design in such a way that an additional module neural network that learns when to remember and when to forget information. The long short term memory method has feedback connections which allow it to process entire sequences of data. This will make LSTM model highly effective in understanding and predicting the time series data. LSTM model achieves superior prediction accuracy.

## References


Abuein, Q. Q., Shatnawi, M. Q., Aljawarneh, E. Y., & Manasrah, A. (2024). Time series forecasting model for the stock market using LSTM and SVR. *International Journal of Advance Soft Computing Applications*, *16*(1). https://doi.org/10.20944/preprints202409.1005.v1

Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R., & Khatri, R. K. (2022). Predicting stock market index using LSTM. *Machine Learning with Applications*, *9*(2), 100320. https://doi.org/10.1016/j.mlwa.2022.100320

Bosworth, B., & Collin, S. (1999). Capital flows to developing economies: Implications for saving and investment. *Brookings Papers on Economic Activity*, *1999*(1), 143–180.

Chen, Z. (2024). S&p 500 stock price prediction using LSTM. *Highlights in Science, Engineering and Technology*, *88*, 57–63.

Gupta, P., Malik, S., Apoorb, K., Sameer, S. M., Vardhan, V., & Ragam, P. (2024). Stock market analysis using long short-term model. *EAI Endorsed Transactions on Scalable Information Systems*, *11*(2), 1–6. https://publications.eai.eu/index.php/sis/article/view/4446

Khan, W., Ghazanfar, M., Azam, M., Karami, A., Alyoubi, K., & Alfakeeh, A. (2020). Stock market prediction using machine learning classifiers and social media, news. *Journal of Ambient Intelligence and Humanized Computing*, *13*, 1–24.

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, *103*, 25–37.

Parmar, I., Agarwal, N., Saxena, S., Arora, R., Gupta, S., & Dhiman, H. (2018). Stock market prediction using machine learning. *First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 574–576.

Rather, A., Agarwal, A., & Sastry, V. (2015). Recurrent neural network and a hybrid model for prediction of stock returns. *Expert Systems with Applications*, *42*(6), 3234–3241.

Stiglitz, J. (2000). The contributions of the economics of information to twentieth century economics. *Quarterly Journal of Economics*, *115*(4), 1441–1478. https://doi.org/10.1162/003355300555015

Team, S. D. (2025). Spyder: The scientific Python development environment [Accessed: 2025-06-05].

Vaisla, K., & Bhatt, A. (2010). An analysis of the performance of artificial neural network technique for stock market forecasting. *International Journal on Computer Science and Engineering*, *2*(6), 2104–2109.

van Rossum, G., & Fred L. Drake, J. (2009). *Python 3 reference manual*. Python Software Foundation. https://docs.python.org/3/reference/

Wang, Y. (2024). Research on stock price prediction based on LSTM modeling–A stock market as a case study. *Highlights in Business, Economics and Management*, *24*, 1634–1640.

You, Y. (2024). Forecasting stock price: A deep learning approach with LSTM and hyperparameter optimization. *Highlights in Science, Engineering and Technology*, *85*, 328–338. https://doi.org/10.54097/vfa8fe80

Zhang, J., Xiang, Y., Zou, Y., & Guo, S. (2024). Volatility forecasting of Chinese energy market: Which uncertainty have better performance? *International Review of Financial Analysis*, *91*, 102952. https://doi.org/10.1016/j.irfa.2023.102952