IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 9, Issue 4, Aug - Sep 2021
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

Stock Index Prediction using Deep Learning (LSTM - RNN) with reference to NSE NIFTY50

K. ANKUSH AGARWAL, M. RAMESH SHARMA

[1] **Dean (Academics & Administration)**, MIT, Pune, Maharashtra.

[2] **Assistant Professor – Business Analytics**, Department of Business Administration, MIT, Pune, Maharashtra

Abstract:-

Financial Analytics plays a very important role interms of Prediction and analysis of stock market data. Financial Analytics has become very important for Better Risk Management and Risk Analysis**.** Stock Index price prediction needs timely information that helps the traders to take important decisions. Time Series are extremely nonlinear in nature and hence, accurate stock price forecasting has been a challenge. Accurate Prediction of Stock Prices and the direction of index price movements is also essential for a stock trader to trade profitably as well as hedger to design risk management strategies. Deeplearning approach to stock price forecasting is been presented in this study. Deep learning models based on LSTM RNN (Long-Short Term Memory – Recurrent Neural Networks) using Keras open source wrapper with TensorFlow library in the backend are designed and empirically evaluated on the index of NIFTY 50 for the time period 2007 to 2019**.** Long Short-Term Memory (LSTM) networks are a modified version of Recurrent Neural Networks, which makes it easier to remember past data in memory. Prediction of next 30 days and plot of the output indicates that a downward movement is expected. Results indicate that deep learning models proposed in this study can generate stock index price forecasts which can be used for basis of Risk Management.

Keywords:- Finanical Analytics, Deep Learning, LSTM RNN, Keras, training,testing

1. Introduction:

Analytics has become very important in the Industry of Finance, which is used for Better Risk Management and Risk Analysis. Stock Market is a place where stocks of a company will be

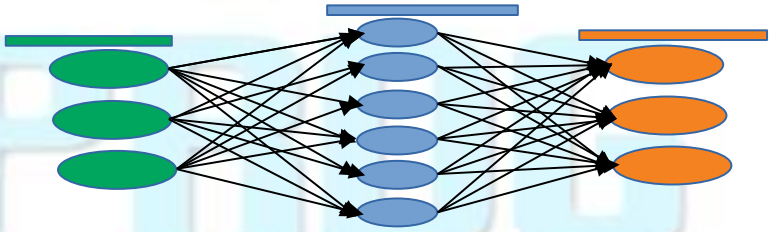
bought and sold. Predicting the returns of the stock market is one of the most difficult to analyse because of involvmenet of many factors which include both micro and macro factors. Stock Market can be categorized in to Primary Market (IPO) and Secondary Market. Primary market is where the new issues are introduced to the market through the process of IPO's (Initial Public Offerings). Secondary market is where investors trade securities (buy or sell). Stock market is having a highly volatile and non-linear time series data. A time series is a set of data measured over time to know the status of volatality. Linear models such as AR, ARMA, ARIMA etc have been used for stock market returns forecasting. The problem exists with these models are they work only for a model identified for a particular index and it won't perform well for another. Stock Market forecasting takes higher risk and it is one of the most important reason for the difficulty in stock market prediction. Here is where the application of deep-learning models in financial forecasting comes in to the picture.

Deep Neural Networks got its name due to the use of the architecture of neural network in Deep Learning models and it is also called as Artificial Neural Network(ANN). ANNs are capable to learn and generalize from their experience. Practical application of Artificial Neural Network's in forecasting problems is very successful.

Figure 1.1.1: Artificial Neural Network

In general, an Artificial Neural Network (ANN) consists of three layers:

1) Input layer 2) Hidden layers 3) Output layer

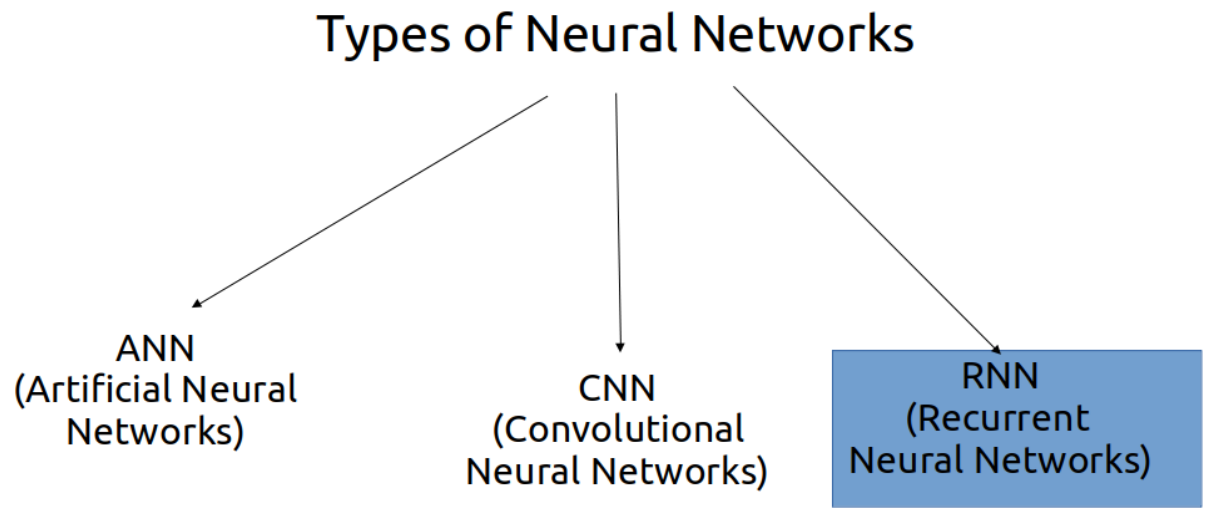


1.1 ARTIFICIAL NEURAL NETWORK

ANN (Artificial Neural Network) is a computational structure which mimics to that of biological neurons. Artificial Neural Network has capability to learn from the underlying patterns from the data. It is designed to identify an underlying trend from a data and to generalize from it. Artificial Neural Network's are considered as non-linear statistical data tools. The underlying relationship between outputs and inputs can be modeled using Artificial Neural Network (ANN). Non-linear activation functions are used in all the nodes in hidden as-

well-as output layers excluding input layer. Each node in the input layer is connected to each neuron in the succeeding hidden layer followed by output layer.

Fig 1.1.2 Types of Neural Networks



1.2 RECURRENT NEURAL NETWORK (RNN)

Recurrent Neural Networks (RNN) will take inputs from two sources the present and from the past data. Information from these two sources are used to decide how they react to the new set of data. This is done with the help of a feedback loop where output at each instant is an input to the next moment. Here we can say that the recurrent neural network has memory. Each input sequence has plenty of information and this information are stored in the hidden state of recurrent networks. Recurrent Neural Network cannot able to store long time memory, so the use of the Long Short-Term Memory (LSTM) based on memory line proved to be very useful in forecasting cases with long time data.

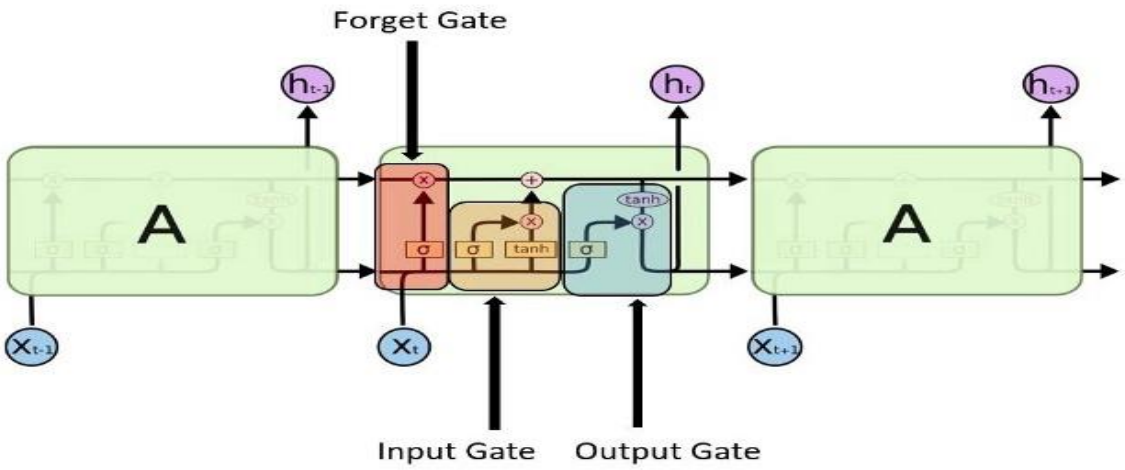
LSTM have a different structure compared to other neural networks. Conventional RNN has a very simple neural network with a feedback loop but LSTM consists of a memory block or cells instead of a single neural network layer.

2. LONG SHORT TERM MEMORY

LSTM is a special type of RNN. These networks are proficient in learning about long-term dependencies. It was introduced by Hochreiter and Schmidhuber in 1997. These networks are clearly designed to evade the long- term dependency problem, but remembering information

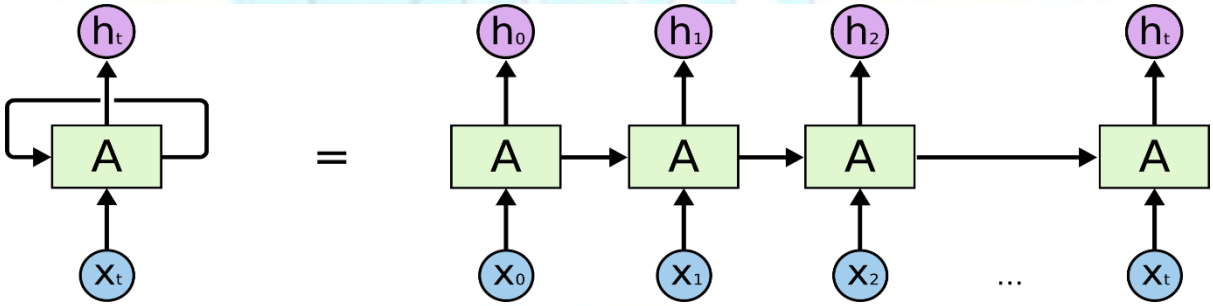
for a long time period back is their normal behavior. LSTM have a different structure compared to other neural networks.

Internal Structure of LSTM



Source: https://towardsdatascience.com/understanding-rnn-and-lstm-f7cdf6dfc14e

LSTM RNN



Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Long Short-Term Memory (LSTM) networks are a modified version of Recurrent Neural Networks, which makes it easier to remember past data in memory. LSTM is an added extension in the RNN that makes it possible to increase the RNN's Memory. LSTM enable the RNN's to remember the input and the output for a long time due to the memory. Long Short-Term Memory (LSTM) networks are a modified version of Recurrent Neural Networks, which makes it easier to remember past data in memory. LSTM is an added extension in the RNN that makes it possible to increase the RNN's Memory. LSTM enable the RNN's to remember the input and the output for a long time due to the memory.

3. METHODOLOGY & DATA

The data in the present study consists of the daily closing prices of NIFTY50 index in the National Stock Exchange extracted from yahoo finance and data series cover the period going from September 2007 to December 2019. The prediction involved for the next 1 month i.e, for 30 days. Feature Close data has been taken for the present study. Steps involved in the present study are Preprocessing of the data, creating the LSTM Model, predicting the test data, predict the test data and plot the output and Prediction.

LSTM Recurrent Neural Networks have proven their capability to outperform in the time series prediction problems. When it comes to learn from the previous patterns and predict the next pattern in the sequence, LSTM models are best in this task.

3.1 LSTM - RNN (Long Short Term Memory – Recurrent Neural Network) has been used and the percentage of training data is 65% and the other 35% of data is for testing.

Total : 2998

Training Data : 1948 (65% of Total Count)

Testing Data : 1050 (35% of Total Count)

3.2 LSTM are sensitive to the scale of the data so MinMaxScalar is been used to bring the closing price of Nifty index in the range of (0 to 1). For training mean squared error is used to optimize the model. Also, 50 Epochs are used for training the data.

3.3 Time Step of 100 has been choosen for the study.

3.1.1 Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 100, 50)	10400
lstm_1 (LSTM)	(None, 100, 50)	20200
lstm_2 (LSTM)	(None, 50)	20200
dense (Dense)	(None, 1)	51

Total params: 50,851
Trainable params: 50,851
Non-trainable params: 0

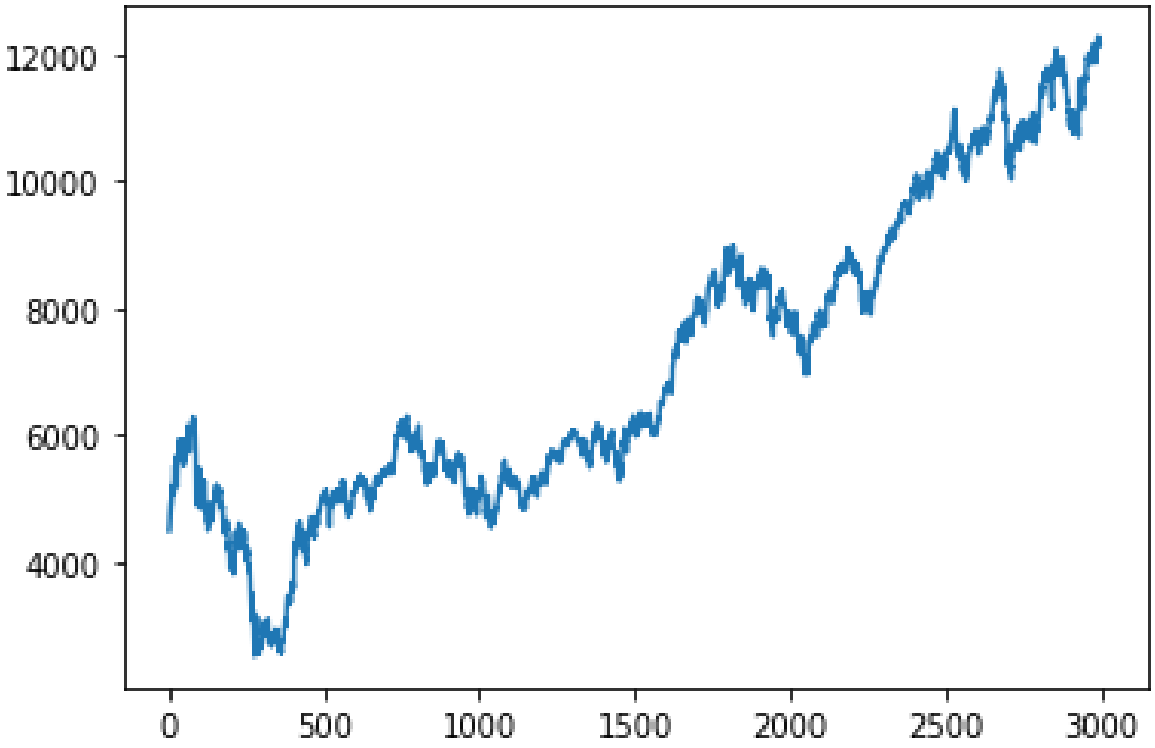
3.1.2 Descriptive Statistics of the Feature 'Close' of NIFTY50

Count	2998
Mean	7134.31
Std	2417.35
Min	2524.2
25%	5253.6
50%	6157.28
75%	8744.95
max	12271.8
`Name: Close,dtype:float64`	

4. Results and Discussion

After training the model it is observed that the result of the testing has shown different results based on number of epochs used as well as the size of the data taken for training and testing which is having significant impact on the result of testing.

4.1.1: The graph depicts the movement of Nifty 50 Stock Index from September 2007 to December 2019 in terms of the closing prices. Since 2007 it is been depicted that volatality is been existed in the stock index and also it can be observed the uptrend.



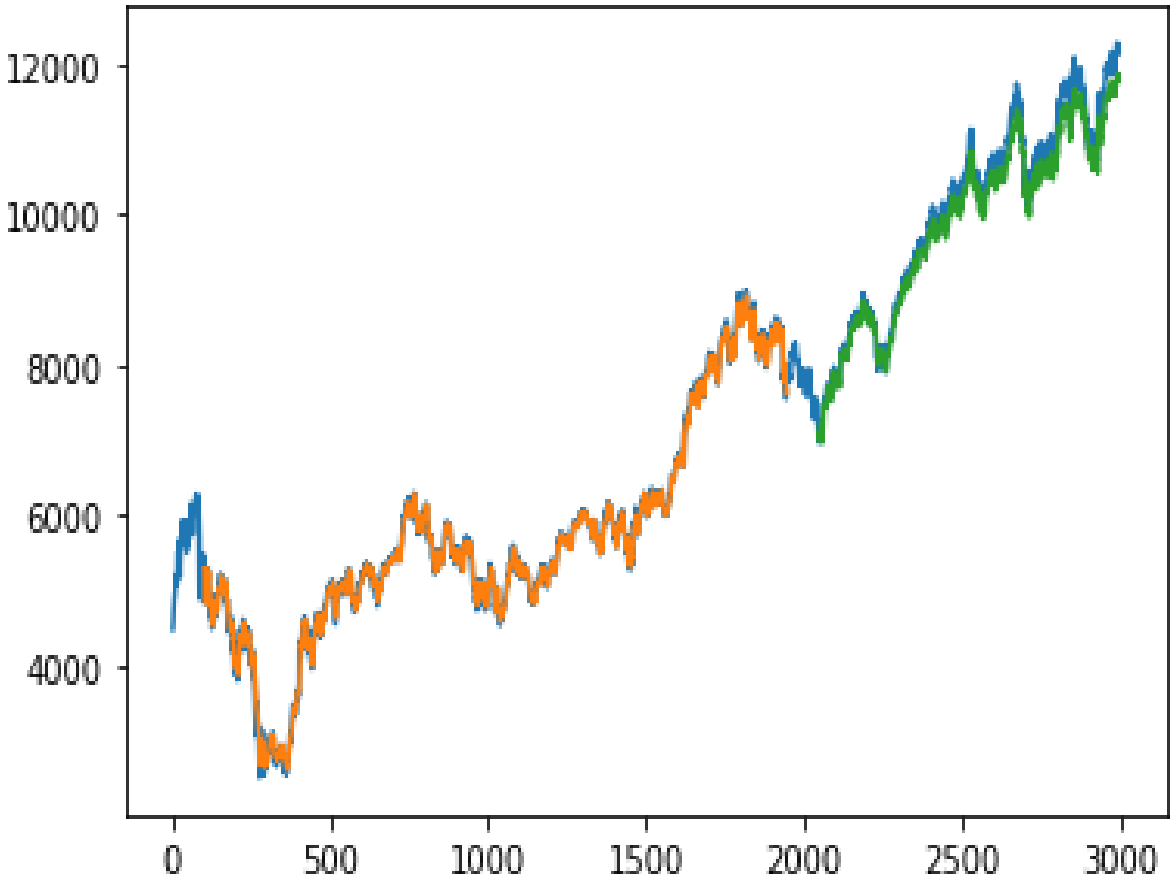
4.1.2 Graph Depicts the Training (Color - Orange), Testing (Color - Green) and Actual Movement (Color - Blue) of Nifty 50

Model has been run with Epochs = 50 which Train on 1847 samples, validate on 949 samples.

Epoch 50/50

1847/1847 [==============================] - 12s 7ms/sample - loss: 2.6896e-04 - val_loss: 0.0027

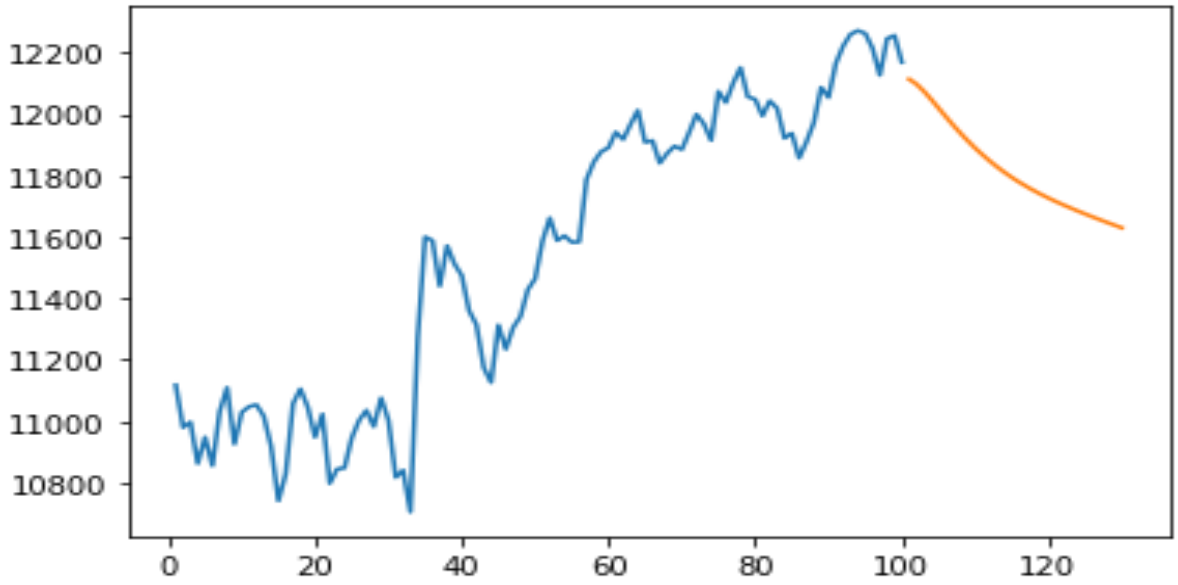
Epoch 50 has given a val_loss of 0.0027



As we can observe in the above plot the LSTM model has predicted almost the same sequence as it was expected.

Interpretation: Blue Color line represents the movement of the toal NIFTY50 data set for the time period 2007 to 2019. Green Color represents the Test Data (65%) Predicted output and it starts on a specific period. It can be observed that the actual closing data points of NIFTY20 represented in color blue and test data (35%) is represented in green color in the last part of the graph. The Training Dataset which is 65% of the total data set is been represented in the color Orange.

4.1.3 Predicting the output for the next 30days



The graph 4.1.3 depicts that there will be a downward movement in the coming 30 trading sessions according to the analysis. The prediction is purely based on the past price movements.

5. Conclusion:-

The present study proposes RNN based on LSTM built to forecast the movement of NIFTY 50 Stock Index of National Stock Exchange, the model has shown some significant results in terms of estimating the movement of the stock index. Movement of the Stock Index is based on so many factors like micro and macro factors. In the present study it is analysed the dependency of past time steps in the future performance of the Index. From the obtained results, it is observed that the present deep learning model (LSTM-RNN) is able to generate good forecast for financial time series.

References:-

[1] S. Selvin , R. Vinayakumar , E. A. Gopalakrishnan , V. K. Menon and K. P. Soman.(2017) ”Stock price prediction using LSTM, RNN and CNN-sliding window model.” International Conference on Advances in Computing, Communications and Informatics: 1643-1647.

[2] Rather A. M., Agarwal A., and Sastry V. N. (2015). ”Recurrent neural network and a hybrid model for prediction of stock returns.”Expert Systems with Applications 42 (6): 3234-3241.

[3] Zhang G., Patuwo B. E., and Hu M. Y. (1998). ”Forecasting with artificial neural networks: The state of the art.” International journal of forecasting 14 (1): 35-62.

[4] Heaton J. B., Polson N. G., and Witte J. H. (2017). ”Deep learning for finance: deep portfolios.” Applied Stochastic Models in Business and Industry 33 (1): 3-12.

[5] Jabin S. (2014). ”Stock market prediction using feed-forward artificial neural network”. growth 99 (9).

[6] Hamzaebi C., Akay D. and Kutay F. (2009). ”Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting.” Expert Systems with Applications 36 (2): 3839-3844.

[7] Rout A. K., Dash P. K., Dash R., and Bisoi R. (2015). ”Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach.”Journal of King Saud University-Computer and Information Sciences 29 (4):536-552.

[8] https://towardsdatascience.com/understanding-rnn-and-lstm-f7cdf6dfc14e

[9] https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Websites:

1. https://towardsdatascience.com

2.https://analyticsindiamag.com

3. www.analyticsvidhya.com