FINANCIAL TIME SERIES FORECASTING

Sakthiprian S

November 2023

Aim and Objective

Time series data refers to the various attributes collected over an extended period of time. Time series forecasting refers to the ability to predict the values of these attributes with minimal to no error. This is one of the many problems which can be tackled using machine learning tools.

In this study, we choose the stock of Axis bank and analyse the time series data of its stock price obtained from Kaggle. The main objectives of this study are to forecast the opening, closing , highest and lowest prices of the stock given we have the information on the history of variation of these attributes .

Literature review

Many attempts have been made for multivariate time series forecasting. The earliest methods include the ARIMA method by George Box and Gwilym Jenkins. Other early methods include SARIMA (Seasonal Auto-Regressive Integrated Moving Average) and SARIMAX (Seasonal Auto-Regressive Integrated Moving Average with exogenous factors) algorithms. However, use of these algorithms have often been associated with overfitting and inability to learn complex patterns. Thus, an alternative neural networks approach is more viable since neural networks learn non-linear relationships.

Recurrent neural networks (RNNs) are Deep Learning models capable of retaining memory of patterns in historic time series data. Many research articles [1] have used RNNs to reduce loss associated with the forecasting. RNNs perform very well with short term data and fail to perform well in long term data with more complex patterns [2]. Thus, an architecture for long term memory is needed [3].

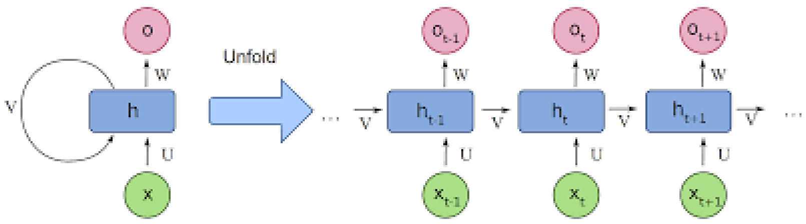


fig 1: Recurrent Neural Networks

source:https://miro.medium.com/v2/resize:fit:1100/format:webp/1*SKGAqkVVzT6co-sZ29ze-g.png

A modified version of RNNs is the LSTM (Long Short Term Memory) cell. Comparison with ARIMA and RNNs have also been studied [4]. Unlike the RNN, LSTM has two separate pathways to let both the short term patterns and the long term patterns flow [5].

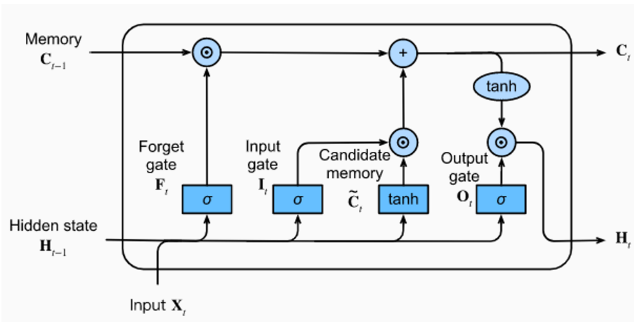


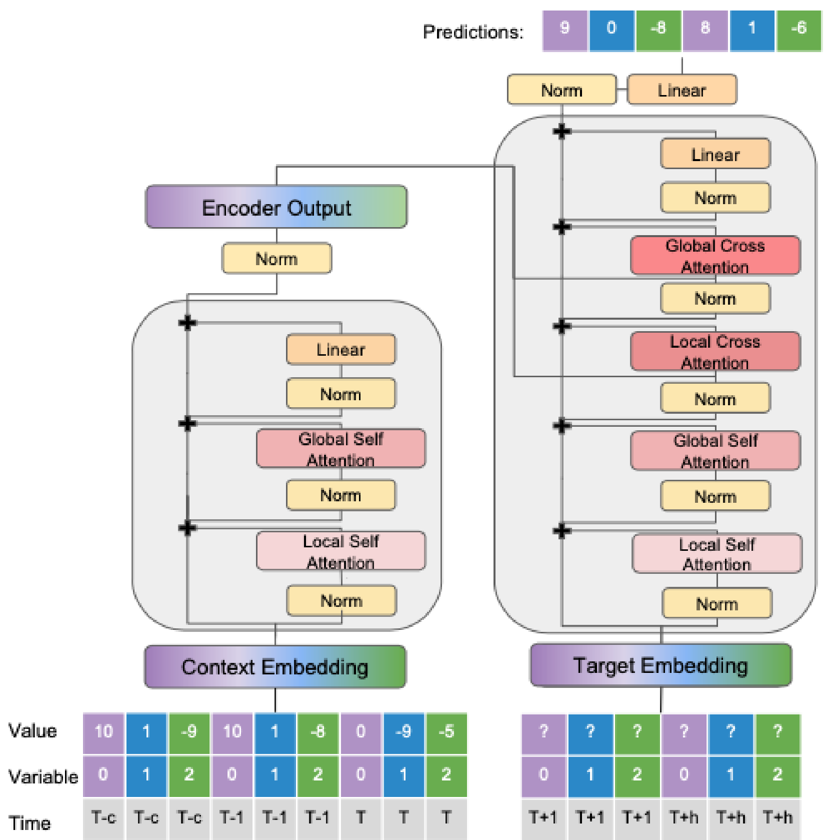
fig 2 : Lstm cell

source:https://miro.medium.com/v2/resize:fit:1100/format:webp/1*Mb_L_slY9rjMr8-IADHvwg.png

Several variants of the LSTM are used to better forecast multivariate time series forecasting. These include ConvLSTM (A CNN- LSTM hybrid) and Temporal Convolution Networks also provide a method of forecasting time series data [7].

A Convolutional Neural Network (CNN) is employed to identify the spatial relationships of features which are static with respect to time. This network when combined with the LSTM network (which specialises in temporal data patterns) can be used to identify hidden underlying patterns in multivariate uses, where each point in time, a set of features can be observed.

A LSTM (Bidirectional Long Short Term Memory) network has been used along with encoder-decoder to make time predictions. This architecture when combined with the attention mechanism [8] can be used to forecast time series data with lower loss [9]. This method is shown to perform better on spatiotemporal data.



Attention mechanism based model for multivariate time series forecasting. source:https://github.com/QData/spacetimeformer/blob/main/readme_media/spacetimeformer_arch.png?raw=true

Widely used in NLP (Natural Language Processing), attention mechanisms work on the basis of assigning attention levels to various features in the input word embeddings. In our case, word to vector embeddings are not required and the feature vector can be input directly. In the output of the encoder, attention is given to the most important feature. This encoder output (context vector) is then fed into the decoder which also has the target embedding (in our case, just the target vector) as the ground truth. The weights and biases of the encoder and the decoder circuits are then trained using backpropagation. Unlike NLP based applications, we need not worry about including the positional embeddings since the order of the features importance is not a factor to consider.

Methodology

The Axis bank data is loaded into a Google colab notebook and the following steps were performed.

1. Different features of the dataset are observed.
2. Data Cleaning performed by conversion of date column into a datetime datatype and irrelevant rows are dropped.
3. Data visualisation: The different time series dependencies are analysed by plotting graphs between the columns and date.
4. This cleaned data is then scaled using the Standard Scaler to make it computationally more efficient due to the presence of large data like volume columns.
5. This scaled data is then made into a series of 15 elements (window size) as features and the next element as the target.
6. A Stacked LSTM model is then built and Dropout and dense layers are added. Dropout would significantly improve the training time as it only takes 80 percent of the hidden neurons randomly during one epoch of the training.
7. The training loss and validation loss are analysed graphically.
8. The model is then made to predict for 45 days and the output has been plotted.

Conclusion

Implementation of stacked LSTM mechanism in the multivariate time series forecasting of financial information relating to the stock price of Axis Bank has been performed . The impact of machine learning models in financial forecasting given the previous information on the stock price has been analysed and understood.

References

[1] - Ashour, M.A.H. and Abbas, R.A., 2018, February. Improving time series' forecast errors by using recurrent neural networks. In *Proceedings of the 2018 7th International Conference on Software and Computer Applications* (pp. 229-232).

[2] - Manaswi, N.K. and Manaswi, N.K., 2018. Rnn and lstm. *Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition With TensorFlow and Keras*, pp.115-126.

[3] - Ludwig, S.A., 2019, June. Comparison of time series approaches applied to greenhouse gas analysis: ANFIS, RNN, and LSTM. In *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1-6). IEEE.

[4] - Siami-Namini, S., Tavakoli, N. and Namin, A.S., 2018, December. A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE international conference on machine learning and applications (ICMLA)* (pp. 1394-1401). IEEE.

[5] - Staudemeyer, R.C. and Morris, E.R., 2019. Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

[6] - Wan, R., Mei, S., Wang, J., Liu, M. and Yang, F., 2019. Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. *Electronics*, *8*(8), p.876.

[7] - Wan, R., Mei, S., Wang, J., Liu, M. and Yang, F., 2019. Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. *Electronics*, *8*(8), p.876.

[8] - Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, *30*.

[9] - Du, S., Li, T., Yang, Y. and Horng, S.J., 2020. Multivariate time series forecasting via attention-based encoder–decoder framework. *Neurocomputing*, *388*, pp.269-279.

[10] - Hao, Y. and Cao, H., 2020, July. A new attention mechanism to classify multivariate time series. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*.