Exploring LSTM Variants for Multivariate Time Series Dataset

Mini Project - Sem V

Supervisor: Dr. Ranjana Vyas

Group Members

Shreyas Gupta	IIT2019102
Sarthak Maheshwari	IIT2019117
Garvit Chittora	IIT2019142
Mrinal Bhave	IIT2019152
Varun Tiwari	IIT2019154

Introduction

What is the Stock Market?

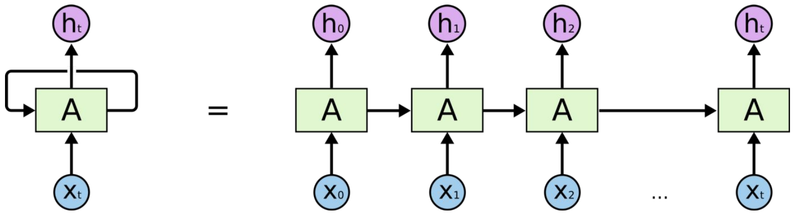
- ➔ A public marketplace where shares of public companies are bought and sold everyday
- ➔ Provides a way for the public to invest their money in equity
- ➔ Stock prices vary according to the changes in market, they are largely determined by the forces of demand and supply.

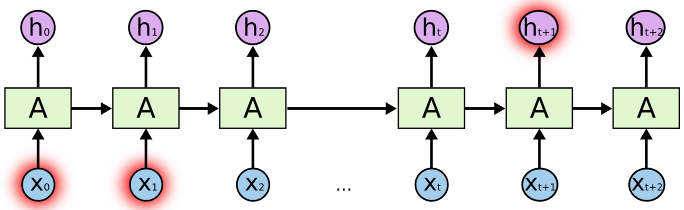
Time Series Analysis

- A specific way of analyzing a sequence of data points collected over an interval of time is termed as Time Series Analysis
- It can be useful to see how a given asset, security, or economic variable changes over time

Recurrent Neural Networks

- Many models have been developed for TSA, the most basic being Recurrent Neural Network
- RNN feeds on sequential data and produces output based on output from the previous cell





Why LSTM?

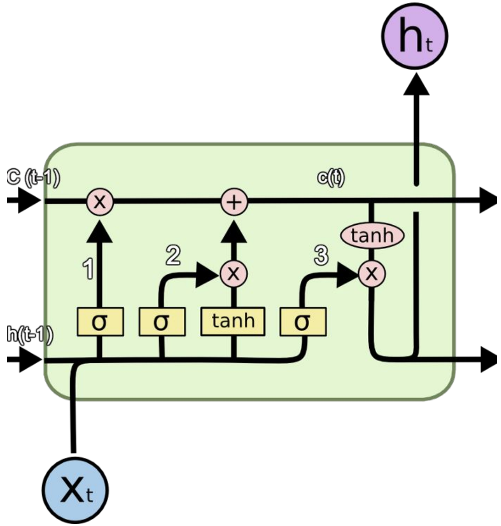
Unfortunately, RNNs don’t work for long term information because of its sequential nature

In the long term, memory can be easily corrupted due to constant large updates

To solve this issue, LSTM networks were explicitly designed to learn long term dependencies

LSTM - Long Short Term Memory

- LSTM networks also have a chain like structure like RNNs but have design differences used to keep track of the long term dependencies
- An LSTM chain not only keeps track of the output from last module(h(t-1)) but also the long term cell memory (c(t-1))
- A basic LSTM module has three layers instead of one, namely they perform the following functions:
    - Decide whether we want to keep or forget previous cell memory
    - Decide what new information to update in cell memory
    - Decide what to output for this module



Related Work

We reviewed previous efforts to predict stock market prices using different machine learning models and how their results compare when applied on different types of data and observed that:

- LSTM was the most used model, and it was proven to be better than ARIMA (Autoregressive Integrated Moving Average) model for stock price prediction
- The bidirectional and stacked LSTMs had better performance for short term prices opposed to the long term prediction results
- Supervised learning classifiers can be used to forecast stock price movement based on financial index data
- One paper observed that techniques like random forest, support vector machines were not exploited fully. It examined the use of the prediction system in real-world settings and issues associated with the accuracy of the overall values given
- An improvement in prediction of stock price movements was achieved when including a sentiment analysis for both LSTM and autoregressive integrated moving average with exogenous variables (ARIMAX) models

Problem Statement

Taking the past 20 years of price statistics from Nifty-50 as our source of data, we must build and compare the results for the stock price prediction models for the three variants of Long Short Term Memory - Classic, Stacked and Bidirectional.

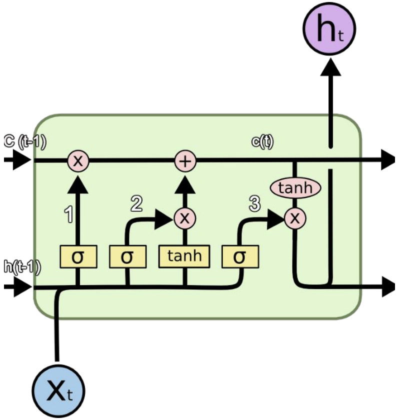
LSTM Variants

LSTMs and their variants can remember short-term memories for a very long time. A number of modifications to the original LSTM architecture have been suggested over the years, forming different variants, three of which are -

- **Classic**
- **Stacked**
- **Bidirectional**

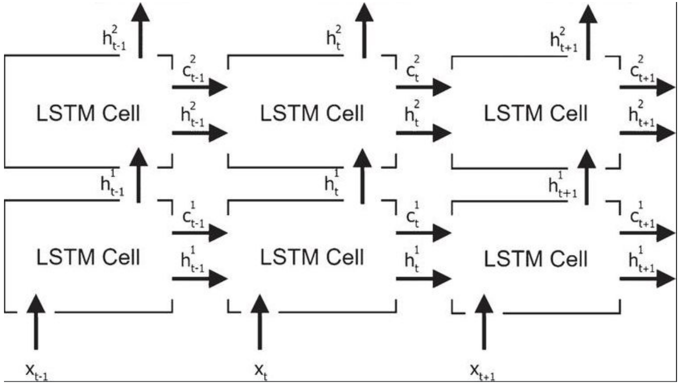
Classic LSTM

Already discussed - the most basic form of LSTM



Stacked LSTM

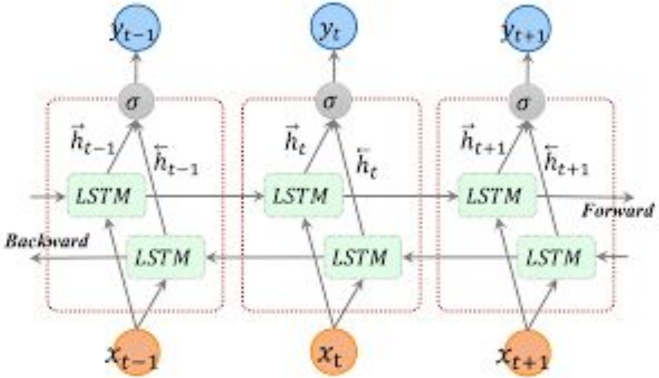
In Stacked LSTM, multiple LSTM cells are vertically stacked and each input in the sequence goes through a series of hidden layers to decide the output. This makes the model deeper and lets it account for more complex input patterns and accommodate more features.



Bidirectional LSTM

In Bidirectional LSTM, the LSTM chain is duplicated and the input sequence is processed in reverse in the second chain.

This makes it so that each cell now has both past and future information to make a decision instead of only past context.

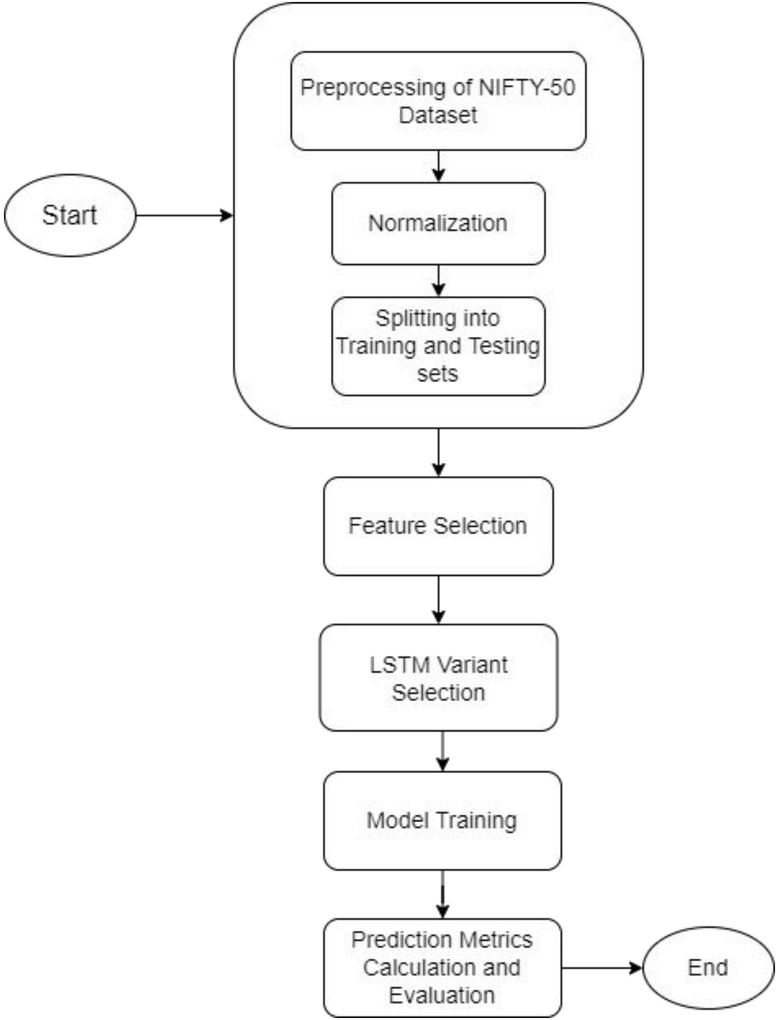


Methodology

Major steps which will be followed in this project:

1. Data Acquisition
2. Data Preprocessing
3. Details about RNN-based models
4. Evaluation metrics

Control Flow Diagram



Dataset

- The data is the price history of the fifty stocks in the NIFTY 50 from 1st January, 2000 to 30th April, 2021.

- The data set consists of the following columns :
  - Date
  - Stock Symbol
  - Type of Security (Equity)
  - Previous day’s close price
  - Open price of the day
  - Highest price of the day
  - Lowest price of the day
  - Last traded price of the day
  - Close price of the day
  - Volume

Normalization

- Data will be normalized using a python library sklearn
- Feature scaling is a method to normalize the range of independent feature variables to a given range, which is 0 to 1 in our case
- Data is then split into a training and a testing set at approximately 80:20 ratio

Training and Testing Models

- Data will be used to train the models with all the combinations of feature sets possible for all the three variants. This will be done for each stock individually and then an average of the results will be taken for all the stocks for the final accuracy results.

- Finally results will be compared for all the feature set combinations for all the variants to conclude the best feature set and variant for stock price prediction.

Training and Testing Models

- Coming to the specific parameters for the models, the neurons will use a rectified linear unit activation function and each model will have a dense output layer that makes a single value prediction

- The Classic LSTM model will have a single input layer.

- The Stacked LSTM model will have two input layers.

- The Bidirectional LSTM model will have two classic lstm layers, one forward and one backward whose results will be concatenated before passing on to the dense output layer

- Each model is trained over 200 epochs

Accuracy

- Evaluation of the methods will be made using root mean squared loss of the model.
- Then final accuracy will be calculated using the sklearn metric *r2_score* which calculates the coefficient of determination or R^2 score, i.e. the amount of the variation in the output dependent variable which is predictable from the input independent variables

Results

Features	Classic LSTM	Stacked LSTM	Bidirectional LSTM
Open, Volume	0.03252	0.02609	0.03609
Low, Close	0.01682	0.01696	0.01418
Low, Volume	0.02798	0.02728	0.02555
Open, High, Low	0.0147	0.01465	0.01506
Open, Close, Volume	0.01601	0.01589	0.01615
High, Low, Close	0.01305	0.01337	0.01235
Low, Close, Volume	0.01635	0.01549	0.01518
Open, High, Close, Volume	0.0149	0.01435	0.01545
High, Low, Close, Volume	0.01377	0.0126	0.01319
Open, High, Low, Close, Volume	0.01371	0.01349	0.01338

Conclusion

- This study employed the use of various LSTM variants to predict the prices of stocks from the NIFTY-50 index
- The three variants explored were: Classic, Stacked and Bidirectional
- Overall, **Bidirectional LSTM** produced the best accuracy with the least deviation of **0.01235** for the feature set {High, Low, Close}.
- **Stacked LSTM** was at a close second with the least deviation of **0.0126**
- While **Classic LSTM** produced the worst result with least deviation of **0.01305**
- The outcomes from this research study hold practical importance because these models can be used for real time prediction of stock prices by HFT firms or hedge funds

Future Scope

- Expanding the models compared: As of now, we are using 3 types of LSTM (Classical, Stacked and Bidirectional) to predict the stock market prices. We can further expand our observations by training other models like ARIMA, ARIMAX and Facebook's Prophet on our data and analysing their performance against each other in different scenarios.

- Adding Granularity: In this study, we only looked at the daily stock price data. As a result, our parameters were calculated over weeks. This is different from actual hedge funds and quantitative trading institutions, who study on the order of minutes. We can further this study by looking at intraday data trends in addition to closing price data, by which, we can create more robust models that can predict sudden changes in momentum during intra-day trading more accurately.

References

[1] K. A. Althelaya, E. M. El-Alfy, and S. Mohammed. Evaluation of bidirectional lstm for short-and long-term stock market prediction. In 2018 9th International Conference on Information and Communication Systems (ICICS)

[2] Raut Sushrut Deepak, Shinde Isha Uday, Dr. D. Malathi, “Machine Learning Approach In Stock Market Prediction”, IJPAM 2017

[3] Tao Xing, Yuan Sun, Qian Wang, Guo Yu. "The Analysis and Prediction of Stock Price", 2013 IEEE International Conference on Granular Computing (GrC).

[4] Siami-Namini, S., & Namin, A. S. (2018). Forecasting economics and financial time series: ARIMA vs. LSTM. preprint, arXiv:1803.06386

References - Contd.

[5] Sharma, A., Tiwari, P., Gupta, A. & Garg, P. (2021). Use of LSTM and ARIMAX Algorithms to Analyze Impact of Sentiment Analysis in Stock Market Prediction. Intelligent Data Communication Technologies and Internet of Things: Proceedings of ICICI 2020, 377–394.

[6] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270, 654–669.

[7] Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal- "Stock Market Prediction Using Machine Learning Algorithms", IJEAT, 2019.

[8] Gavriel, Stylianos (2021) “Stock Market Prediction using Long Short-Term Memory”

Thank You