Introduction To Deep Learning(ECE-GY 9123)

Final Project

Stock Market Analysis and Prediction Using Deep Learning

Anindya Prusty(ag7120) || Akhand Pratap Singh(aps646)

GITHUB LINK: https://github.com/AkhandSingh96/Stock-Market-Prediction-Using-Deep-Learning-.git

1. Introduction

Due to the vulnerability of the stock market, this field has always drawn interest from traders, investors, academic researchers etc, because it is a challenging yet a very rewarding task. Predicting the stock market and designing suitable, easy yet effective Machine Learning models has become much more mainstream as compared to traditional ones. And stock price time series requires continuous and deep research.

Deep Learning Models can be trained and employed by following the specifications of neural networks and learning the performance of the stocks over past available data to predict the stock market prices.

Deep Learning models have become quite useful and pivotal in such areas where we have to do or find out data by the method of prediction, for these models have no fancy inbuilt layers that tend to make the task tedious , and thus make the traditional classification and regression predictions work well where many such data samples are available.

Thus an important aspect of this project was also to come up and select those models that fit really well and those that give comprehensive and superior predictions for time-data series. Hence, the models that proved quite beneficial for our project are **Recurrent Neural Networks (RNN)** , **Long-Short Term Memory (LSTM)** and **Gated Recurrent Unit (GRU).**

The project henceforth, taken up by us will deal with this agenda. Following Steps indicate the overall workflow of this project:

A. Get Stock Price Data which will help build the features for this model
B. Preparing Training and Testing Datasets
C. Fit the required neural network to make it learn from the training data
D. Backtesting the System
E. Predict stock prices

2. Literature Survey

Investing in a stock market is one of the most interesting domains whenever we think about financial wellness (earning, saving, investing and protecting). This involves a certain level of risk and this is what makes it even more intimidating. Stock market is a very complex field,

where one correct prediction can help investors yield a good amount of profit through investments. This is where investors can buy and sell securities, stakes in individual companies or ETFs.

People who study the behavior of the stock market using **indexes,** which work by measuring a weighted average value of a collection of securities. The security market index shows the performance of the stock market. S&P 500, NASDAQ are some popular ones.

So when an index decreases its value, it means that the average value of the stocks has decreased from its previous day of trading and otherwise.

By learning the terminology of the stock market and by applying various deep learning models, it has been seen that by properly implementing this, investors have been able to get a better vision and a broader understanding of how the market is performing.

By learning about various deep learning models, we can clearly say that **Convolution Neural Networks** work really well where we want to predict cases that involve numerous targets, but one other advantage is their ability to interpret what's being shown in the images considered.

While working with time-series data, the most common models that we use are **Recurrent Neural Networks (RNN)** , **Long-Short Term Memory (LSTM)** and **Gated Recurrent Unit (GRU).**

So before we have the idea of which model to choose, we need to have a fair share of idea of how the deep learning data is categorized as.

Data can be classified as:

I. **Uncorrelated Data**
II. **Serial Data**
III. **Image Data**

In Deep Learning these data can be addressed in the following three ways:

I. **Feedforward Neural Network**
II. **Recurrent Neural Networks (RNN)/Long-Short Term Memory (LSTM)**
III. **Convolution Neural Network (CNN)**

Among the above ways, it is generally not very fruitful to use **Feedforward Neural Network,** as it takes input data points as independent, without considering the correlations within.

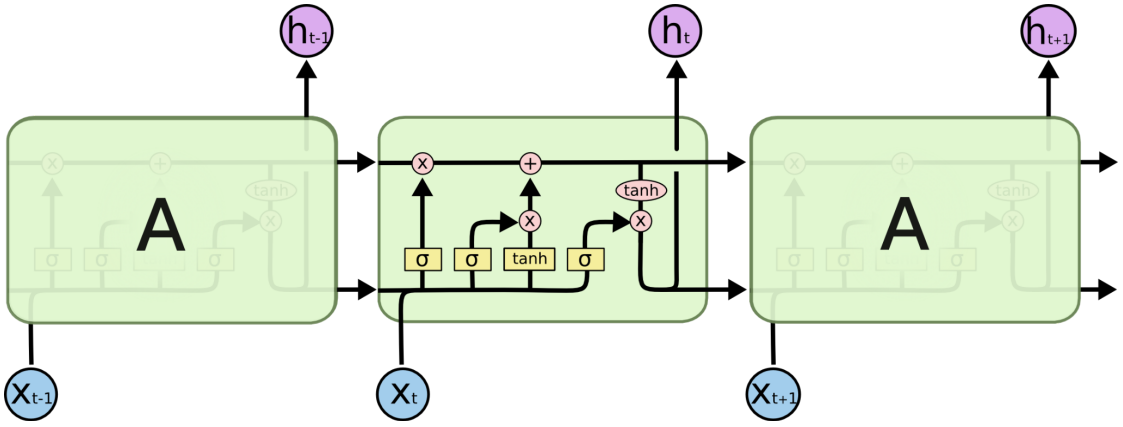
In our project, we have tried to predict the stock market prices, using three models- **LSTM, RNN and GRU.** We have used the stock market dataset from yahoo finance over the past years for this purpose. By implementing python libraries and fitting our models, we try to predict the future prices of the stock.

Long-Short Term Memory Recurrent Neural Network(LSTM)

Comprising a cell, input gate and output gate, this model has an advantage over other neural networks for it can process the entire sequence of data. Because of the feedback connections in its architecture, it is a recurrent neural network.

The cell stores the values over time intervals and the gates control the flow of information into and out of the cell. The cell is responsible for remembering the information regarding the dependencies between the input element.There at three gates in this model- **INPUT, FORGET and OUTPUT.**

Input gate handles the amount of new value that flows into the cell, forget cell controls the value that remains in the cell and the output one is used to calculate the value of the cell used to compute the output of the LSTM model.



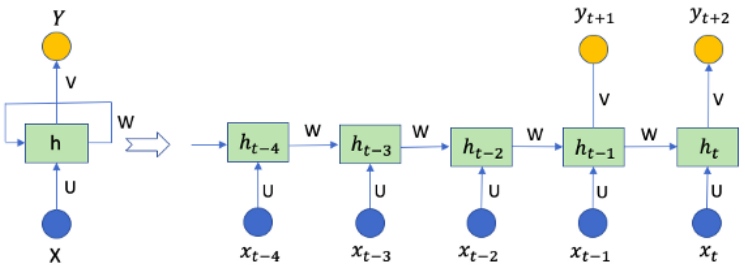
While designing our LSTM model to predict the stock market prices, we have acquired data and come up with the data for the following problem statements:

I. Change in stock price overtime.
II. Daily return of the stock on average
III. Moving Average of various stocks
IV. Correlation between different stocks.

Recurrent Neural Networks(RNN)

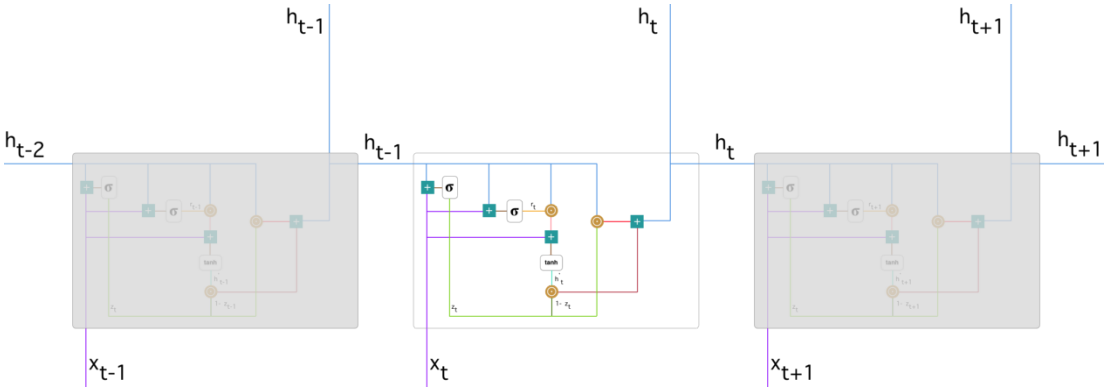
RNNs perform the same task for every sample, based on the result from its previous computations. RNN has a memory that passes information from one step to the other, which tells us that it can handle many steps. Thus, this makes it a suitable model and method to use for stock market prediction as information is passed from one point to another.

Thus the data concluded will have the impact of the results in the past. In this the optimizer has the first-order derivative of the loss function to search for the optimal values. RNN being recursive, the derivation process will make the value smaller, and eventually vanish, which is also known as **Gradient Vanishing.** This makes RNN not a good choice to retain memory.



Gated Recurrent Network (GRU)

The vanishing gradient drawback that we saw in the case of RNN, such a scenario is avoided when using GRU. Since it does not have a cell state and output gate, unlike in LSTM, it has fewer parameters than LSTM. It uses hidden layers for transporting the information, It has two gates- **RESET and UPDATE gates.**



3. Model Description

In this project we have used three models to make predictions based on the stock market data obtained from yahoo finance. The three models we used are Vanilla RNNs, LSTMs and GRUs. All the models mentioned above are sequential neural networks and are utilized extensively in the natural language processing field.

Recurrent neural network is a type of network architecture that accepts variable inputs and variable outputs, which contrasts with the vanilla feed-forward neural networks.

One of the appeals of RNNs is that they might be able to connect previous information to the present task, such as using previous video frames might inform the understanding of the present frame.

LSTM 's and GRU's were created as an alternative to short-term memory. They have internal mechanisms called gates that can regulate the flow of information. These gates can learn which data in a sequence is important to keep or throw away.

By doing that, it can pass relevant information down the long chain of sequences to make predictions. Almost all state of the art results based on recurrent neural networks are achieved with these two networks.

LSTM and GRU's are mostly found in speech recognition, speech synthesis, and text generation.

3.1 Loss Function and Hyperparameters

For Recurrent Neural Network we have used mean squared error as our loss function and rmsprop as our optimizer. In our model we have used 2 hidden layers which are sufficient enough for complex predictions. Each of these hidden layers consists of 32 neurons. We train the model for 100 epochs with a batch size of 150.

Furthermore, incase of LSTM we build a model with 3 hidden layers of 128,64 and 25 respectively. We train our model for 1 epoch and a batch size of 1. We use adam as our optimizer and mean squared error as our loss function.

Lastly, we utilize a sequential GRU model with 2 layers, with the first layer being added with a dropout regularization technique. Each of the two layers consists of 50 neurons each. Also an important thing to note is that we use tanh as our activation function.

We train the GRU model for 100 epochs with a batch size of 150. We also try to modify the architecture of GRU so that there is a dropout layer after every layer (regularization) in order to get better results.

3.2 Training Details

We train each of our models on the training dataset. We split the actual dataset (obtained from yahoo finance) into a ratio of 85:15 as train and test dataset incase of RNNs and GRUs. However in the case of LSTM we divide the dataset into a ratio of 80:20 as train and test dataset respectively.

The batch size and number of epochs (as described in previous section) we chose were the ones that gave the most optimal result on our training dataset.

4. Results and Analysis

Long-Short Term Memory Recurrent Neural Network(LSTM)

First we imported our dataset through yahoo finance and then set up end and start times for the stocks we have to choose to find data of (Tesla- TSLA, General Electric- GE, Apple- AAPL and Amazon- AMZN).

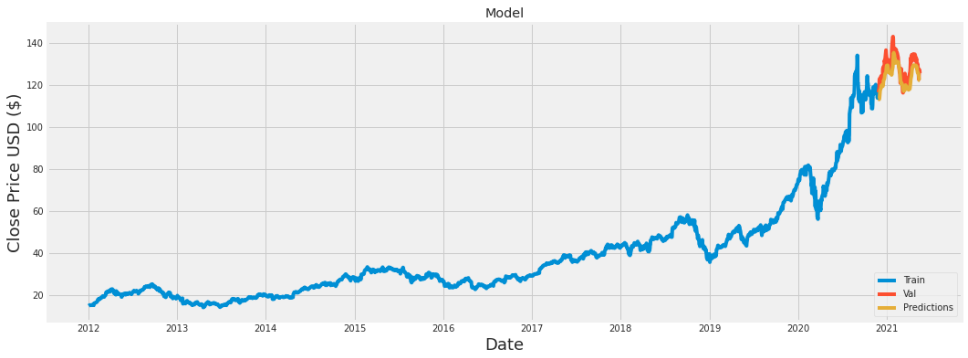
Now that we have pulled our data from the dataset, we can check the statistics of any stock we have chosen and see the various details of the stock such as the High, Low, Opening Price, Closing Price, the volume of stocks available in the market.

```
[ ]  TSLA.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 253 entries, 2020-05-18 to 2021-05-18
Data columns (total 7 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   High          253 non-null    float64
 1   Low           253 non-null    float64
 2   Open          253 non-null    float64
 3   Close         253 non-null    float64
 4   Volume        253 non-null    float64
 5   Adj Close     253 non-null    float64
 6   company_name  253 non-null    object 
dtypes: float64(6), object(1)
memory usage: 15.8+ KB
```

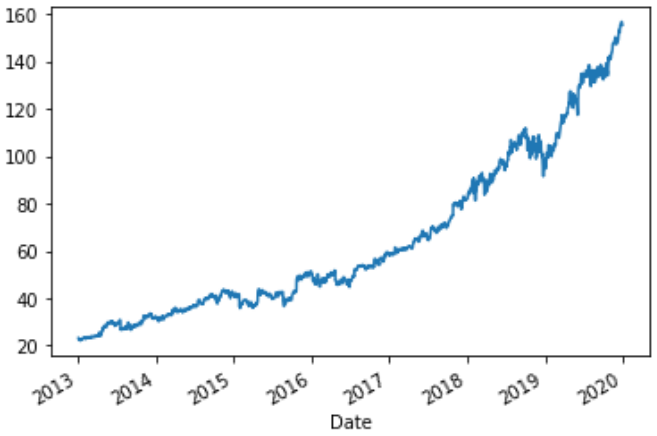
We can also visualize the plot of closing price vs year.

By using the stock quote data we try to predict the closing stock price of Tesla.

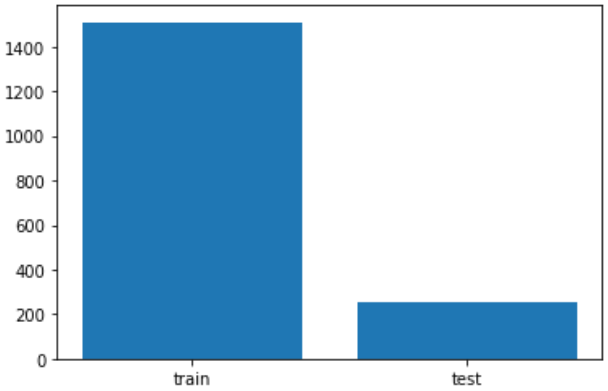


Gated Recurrent Unit (GRU)

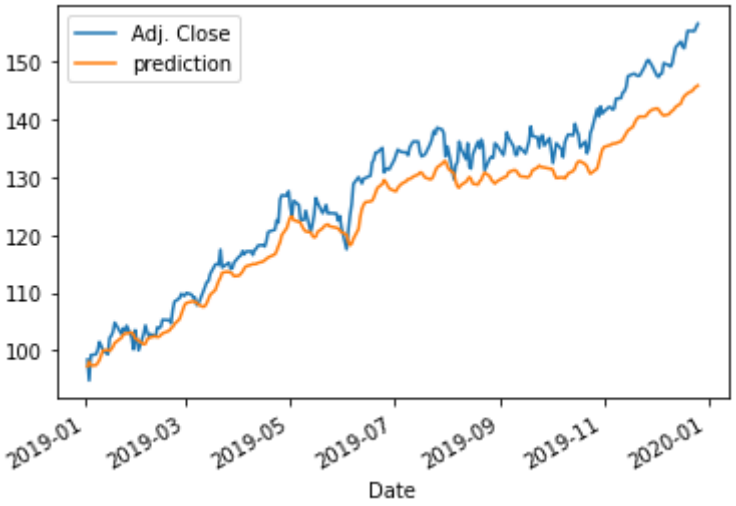
We have used the data from Yahoo Financials to prepare our dataset. We have used this data to predict the stock market prediction for Microsoft(MSFT). And then use this data to plot the graph of the closing market price data for Microsoft (MSFT).



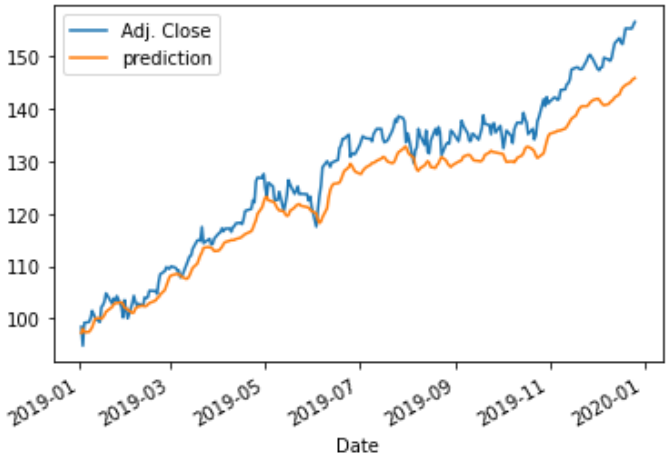
After that we compare the training and test sizes.



After this, we have split the testing data in the ratio of 80:20. And then first we have used the GRU without regularizer model, to compute.

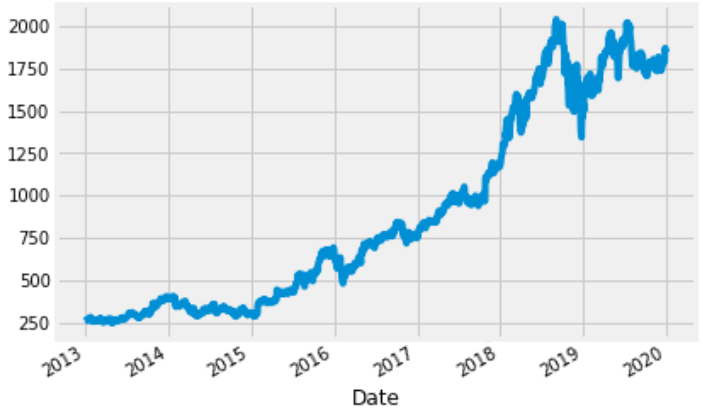


After this we have used the GRU with regularizer for our model training.

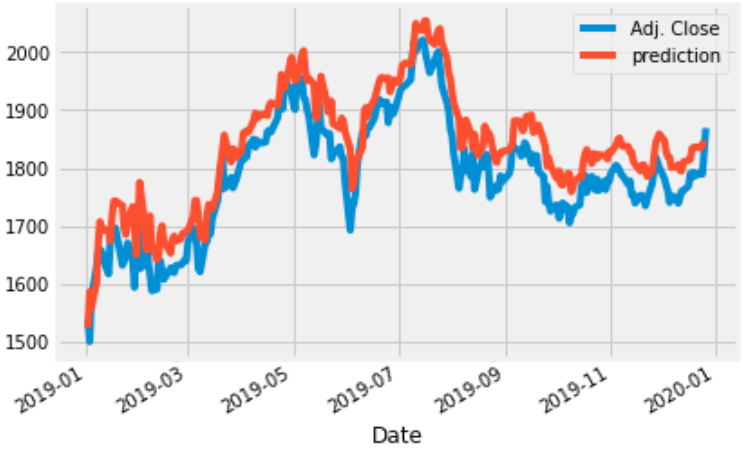


RECURRENT NEURAL NETWORK (RNN)

For this model training purpose, we have used the stock market dataset from Yahoo Financials. And then like we did before, we plot the closing market price data of Amazon.



After this we create the training and testing data samples, and then using this we fit our RNN model onto the data henceforth acquired.



5. Conclusion and Future Work

In conclusion, we strongly believe that sequential neural networks are by far the most accurate mathematical models when it comes to making predictions in the stock market.

Although, investors and stock traders believe that predicting the stock market trends is something that requires a higher degree of research, experience and pre-requisite financial knowledge.

As far as future research about involvement of neural networks in finance is considered, we expect more advanced sequential models to be introduced in future.

6. References:

[1] "Sentiment Analysis of Tweets to Gain Insights into the 2016 US Election" - https://journals.library.columbia.edu/index.php/cusj/article/download/6359/3023

[2] "Location-based Sentiment Analysis and Visualization of Twitter Election Data" - https://dl.acm.org/doi/fullHtml/10.1145/3339909

[3] "Understanding LSTM Networks "- https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[4.]https://medium.com/swlh/a-technical-guide-on-rnn-lstm-gru-for-stock-price-prediction-bce2f7f30346#_=_

[5] https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be

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

[7] https://www.aclweb.org/anthology/U17-1001.pdf