NASDAQ Stock Analysis

Github: https://github.com/morpheu513/NASDAQ_stock_analysis

Suhas Vasisht
Dept of Computer Science
PESU
Bengaluru, India
suhas.vasisht@gmail.com

Amogh RK
Dept of Computer Science
PESU
Bengaluru, India
amoghrameshk@gmail.com

Chirag Kashyap
Dept of Computer Science
PESU
Bengaluru, India
chiragkashyap15@gmail.com

I. Introduction

Our problem statement pertains to stock market analysis. Now first off, why are stock markets important? Stock markets are a gauge of how the economy of a country is doing. It enables companies to be traded publicly and to raise their capital. But more importantly, for the purpose of this paper/project, we are interested in the investment aspect of it. Suppose you invest in a stock and its value rises, this results in a profit. On the other hand, suppose the stock value drops, this would mean a loss of money.

This brings up some very important questions. Which stock should you invest in? Will it rise over time or will it fall? There used to be more traditional methods to predict stocks with parameters such as how popular the company is, how viable that sector of the industry is, and many other factors. These kinds of points were noted by experts to try to find trends in the market.

These techniques, even though done by experts, still have a level of error in which cases a lot of money is lost when a lot of money is invested in a stock that drops in value. The goal of different stock market analysis methods is to have as little error as possible/be right as much as possible.

This is where modern data analytics comes in. Although what the experts are doing in a sense is data analytics, at the end of the day they cannot crunch numbers like a computer. Modern methods of data analytics involve building a model based on some logic/algorithms, training the model with data, and using the model to make future predictions of stock.

This method of predicting stocks has proved to be much more effective, with much lesser error rates which results in more value on investment. Now this brings up some important questions yet again. What exact logic should my model use? What makes one model better than the other? There are no perfect answers to these questions. The goal of stock market analysis is to predict the stocks as correctly as possible, without errors. There is no one model that is the best, different models might work better in different situations, they also might work better based on what data is given to them.

This is what we hope to do in the current situation of the stock market crash caused due to the pandemic brought forth by Covid-19. In times like this we cannot simply look at the stock market a few months ago and predict future stock values. Real world events like this shake the stock market so much that it cannot pick things back up like nothing happened. Some stocks will come back up, some will not and we are attempting to predict what exactly will happen to the stock market in the coming months/years.

We attempt to do this by comparing our current situation to the 2008 financial crisis. In 2008 the world saw the biggest drop in stock market history. This happened due to a lot of factors in the US, the key among them being the bursting of the United States housing bubble and excessive risk taking by banks. The stock market hit such a low that the effects of it are felt even now, 12 years later.

We understand that all the factors related to the 2008 financial crisis and our current pandemic are not completely related. We hope to gain important insights from what happened in the financial crisis and use those insights to build our model to predict what will happen to the stock market in the coming months/years after the stock market crash due to the covid-19 pandemic. More on the specifics of how we will approach our problem will be discussed in later sections.

II. Related works

Before looking to solve our problem, we must first look at related works in the field, so that we can learn and use knowledge gained from them in our model.

A. *Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model - Hyeong Kyu Choi*

In this paper the price correlation of two assets is predicted, and the methodology behind it is discussed in great detail(time series data). It is done using a hybrid ARIMA-LSTM model.The data used is S&P 500 stocks from the year 2008 to 2017. The paper talks about how there have been many different models to predict correlations such as Full Historical Model, Constant Correlation Model and several others.

Here in particular, the Autoregressive Integrated Moving Average(ARIMA) model and Long Short-Term Memory(LSTM) model are chosen. LSTM is a recurrent neural network model. The time series data is assumed to be composed of a linear portion and a non-linear portion. So to account for both these portions, the ARIMA model captures and filters out linear trends while the LSTM model captured non-linear trends. When used in combination as a hybrid model it is able to capture both aspects of the time series data.

The paper then goes into the exact specifications of training the ARIMA model and training of the LSTM model. Methods to prevent overfitting are used such as dropout(randomly turning off neurons to prevent interdependence) and regularization(Making sure that no weights in each layer become too large) for LSTM. The train dataset's MSE and development dataset's MSE started to converge after around 200 epochs(passes). MSE, RMSE, and MAE values of the prediction were calculated to choose the appropriate epoch. This paper then concludes that the ARIMA-LSTM model is a very well performing model as a correlation coefficient predictor for stocks.

The takeaway from this is that we can combine two models together to get even better results, which opens up even more possibilities as to how we build our model

B. *Coronavirus: Impact on stock prices and growth expectations -Niels J. Gormsen, Ralph S. J. Koijen*

This paper is aimed at predicting stock market trends post the covid-19 pandemic using data upto June 2020(stock market data and dividend futures). Although the exact mathematical model used here is not the most complex, it is still an important paper because it compares the current pandemic to the 2008 stock market crash which is something we want to be able to use in our model prediction.

It also compares our current situation to the H1N1 virus pandemic and the H2N2 virus pandemic. An assumption made in this paper is that a vaccine will not be available for another two years during which there will be more economic consequences.

Dividend futures are used to predict future stocks and lower bound of the expected growth of the stocks are taken using the 2008 stock market crash as a reference. The paper then gets into details of the liquidity of future dividends and gdp growth which is not something we are using to make predictions nor are we planning to predict these criteria, so we will not be using that information

What we takeaway from this paper is that there are many factors which affect stock prices, especially in times of a pandemic, and looking at previous stock market crashes and pandemics help us gain insights into the future growth of stocks.

C. *A Prediction Approach for Stock Market Volatility Based on Time Series Data-Sheikh Mohmmad Idrees*

They use a method which results in high profit prediction for a stock based on the combination of two very prominent formulae and logics AR and MA.

This uses the method of Box-jenkins, which is an integration of AR and MA , but it checks whether the following time series dataset is stationary or not and whether there is any significant seasonality that needs to be modelled.
This paper gives us insight on how to handle data which is stationary and which has some seasonality and predict its future based on the given parameters in the dataset.

D. *Forecasting on Stock Market Time Series Data Using Data Mining Techniques - A.Subashini , Dr. M. Karthikeyan*

This paper proposes a novel method for forecasting stocks using already prevalent data mining techniques along with some other prediction models. They have used the autoregressive integrated moving average(ARIMA) model to predict future stock prices.

The novel method introduces the use of Data Mining techniques along with the ARIMA model to handle updated predictions for the stock market. They use data mining as it can discover patterns in large data sets and has wide applications in the field of statistics. Data mining techniques are devised to address forecasting problems by providing a reliable model with data mining features.
The ARIMA model uses a combination of autoregression, a moving average model and differencing. In this context, integration is the opposite of differencing. Differencing is useful to remove the trend in a time series and make it stationary.

The main claims of this paper is to provide an efficient and optimized model which allows for the prediction of stock market data.

This paper gives us an insight on how to handle time series data and the different types of techniques which can be used to obtain inferences from this type of data.

We plan to implement the Data mining and visualization techniques employed by this paper in our model.

E. HATS: A Hierarchical Graph Attention Network for Stock Movement Prediction

In this paper, they have studied how to effectively utilize graph-based learning methods in stock market prediction, assuming it to be more effective than linear functions.

HATS method selectively aggregates information on different relation types and adds the information to the representations of each company. But there are other general methods used by technicians to predict the movement of a stock, without considering real world events, such as moving average, relative strength index, stochastics, momentum oscillator, and commodity channel index.

Main claims can be considered as taking several other relational factors into consideration ,other than day's high ,low and volume etc. for predicting the movement of a particular stock, using HATS.

This paper gives us insight to the different kinds of approach, which can be taken to gather information regarding a company to predict it's share value, example: one of the basics to be considered is if the share's intrinsic value is less than fundamental, then it will definitely go higher.

F. Comparative study on Prediction of Support and Resistance Levels with k-Nearest Neighbor and Long Short-Term Memory Methods.

Predictive Analytics is a broad subject which uses mining techniques, artificial intelligence and statistics to analyze and predict for business organizations to make proactive decisions. In stock markets where money and time are very important aspects, time series analytical and forecasting methods, a part of predictive analytics is applied to make predictions that ultimately benefit the investors. The goal of this paper was to compare between different models for the efficient prediction of future stocks. They use machine learning and deep learning models to predict the support and resistance levels for the indication of entry/ exit price level in the stock market.

The proposed model in this paper, performs better then traditional models like ARIMA or exponential smoothing as it pimples current state of the art deep learning methods which perform better than these models. These DL models , although more compute intensive, provide exceptional accuracy improvements when compared to traditional methods like ARIMA or ARMA.

From this paper we can conclude that newer deep learning models which employ techniques such as LSTM perform better than techniques like KNN's or other traditional techniques.

This paper sheds a light into the comparison of traditional models versus the newer models which employ deep learning and the potential merits/demerits that come with them.

From this paper we can see the magnitude of increase in accuracy between the older models and the new ones which use Deep Learning. We have decided to take the architecture of the LSTM model to use in our model.

General takeaway from previous works

There seem to be many different techniques and models used to solve the problem of stock prediction. Out of these only one (B.) is on the impact of coronavirus using the stock market, but it uses a very simple mathematical model which has a lot of room for improvement.

One more important thing to note is that in all these models the more the number of days we try to predict, the lesser the accuracy of our predictions become.

With the insights of the models used in the other papers, in combination with the insights on the coronavirus impact (from paper B.), we should be able to build our own model which predicts future stocks.

III. Proposed problem statement

The Covid-19 pandemic has completely shaken the world. Due to the nature of the coronavirus, it is not safe for anyone to leave their homes. Since no one left their homes, no one was going to work.

For people like software engineers this wasn't much of a problem as they could work from home. But the majority of industries took a huge hit and the stock market hit an all time low.

The previous comparable market crash to this is the financial crisis of 2008. We hope to compare this crisis to the current pandemic to predict future stocks.

We cannot completely correlate both the situations to predict stocks. Some might follow similar trends, some might not. For example, since everyone is at home they tend to watch more movies and order more things home. This led to rise in stocks of companies like Amazon and Netflix.

The main problem we hope to tackle/ aim for our project is to predict future stocks after the stock market crash due to the Covid-19 pandemic. We plan to do this by leveraging stock market data from the financial crisis and using it to train our model to make good predictions on future stocks.

We also have to be careful of which sectors we choose and what data we infer from them. As previously discussed companies like Amazon and Netflix have been doing well so their stocks have increased. So in our model we have to take care of predicting the sectors which have been doing well and predict them differently as compared to sectors which have been doing bad such as Travel and Tourism.

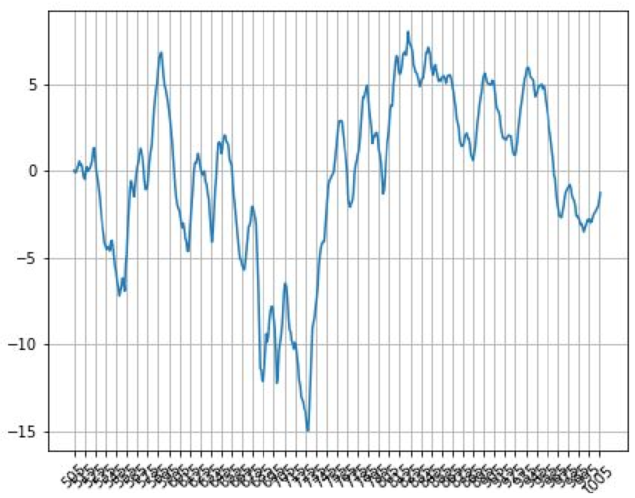
We will talk about the specifics of our model and how exactly we choose to solve the problem in a later section..

IV. Exploratory data analysis

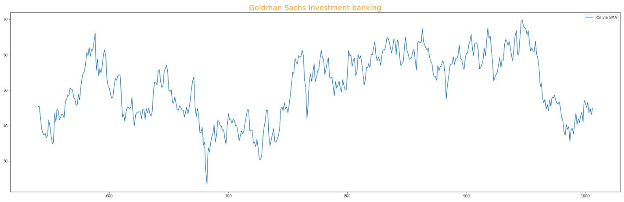
MACD is a trend-following momentum indicator that shows the relationship between two moving averages of a security's price..

The **RSI** or relative strength index tells us the strength or goodness of a particular stock. If the stock has a lot of positive closes(opening price > closing price) then the RSI will increase. It is a good indicator of how good the stock performs. Higher RSI implies better returns if we invest in that particular stock.

MACD



RSI



These two plots help us study the future trend of the stock in the market by drawing a trend line to indicate it's bearish/ bullish activity during the next few days/months/years based on the time/momentum period.

These are just meagre methods to read the trend of the stock considering everything about the company is functioning normally, so the physical/outside news about the company plays a bigger role in deciding it's trend too.

In 2008 Lehman Brothers, a financial services firm, filed for bankruptcy. At the time it was one of the largest investment banks, this caused the US economy to almost collapse. As an immediate effect of this the stock market also crashed. From the above graphs we can see the exact time when the stock prices of other investment sector companies, like JP Morgan Chase and Goldman Sachs, plummeted. This period is also referred to as a recession as it caused widespread economic loss. This affected not only the financial sector but other sectors as well.

The stock displays a huge plummet because the funds and investments made to the companies were taken out of fear, due to huge drop in the market, as a result money flow decreases and the share price decreases.

Harley:Davidson motorcycle manufacturer 2008

The same reason Boeing Aerospace share price dropped, the funding and investment decreased ,leading it to reduce manufacturing rate and difficulty in acquiring materials, resulting in a drop in manufacturing and future projects.

American Express Financial services 2020

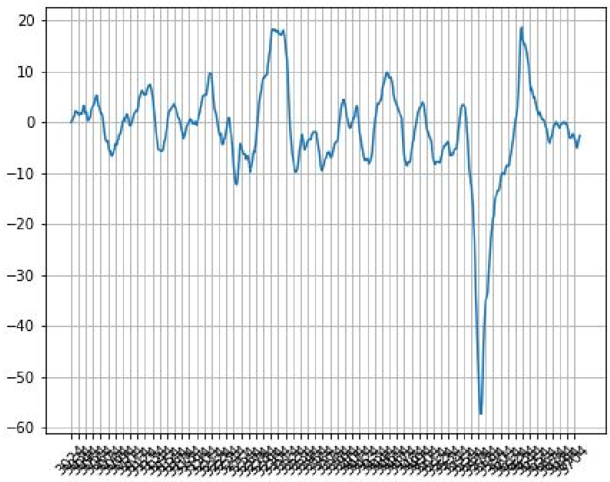
There is a huge plummet, due to pandemic corona, but the major reason is due to investment made by banks and financial sector companies in companies all over the world and in today's world all the countries are interdependent (import/export) which lead to loss because of lockdown of certain countries during the pandemic, which is stoppage in production of materials.

Yum! Brands Fast food 2020

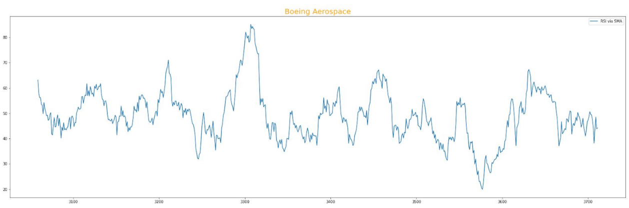
This is an American fast food corporation which is the parent organization for brands such as Taco Bell, Pizza Hut, KFC, etc. We all know that due to the pandemic the revenue of these fast food chains has reduced thereby having lesser investors and reduced stock prices.

Boeing Aerospace 2020

MACD



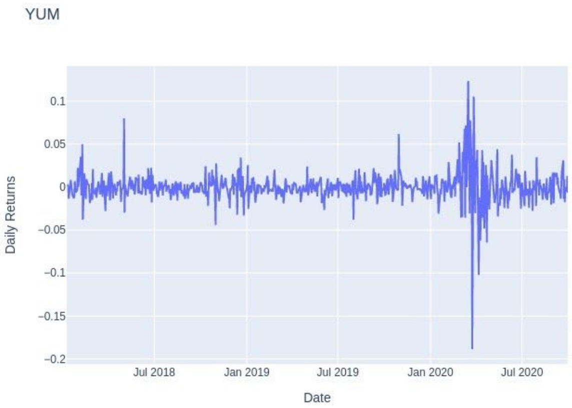
RSI

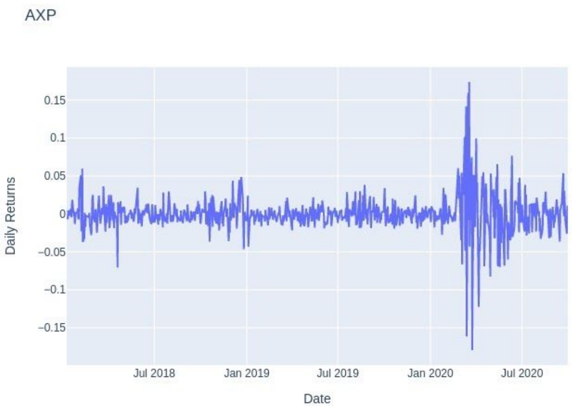


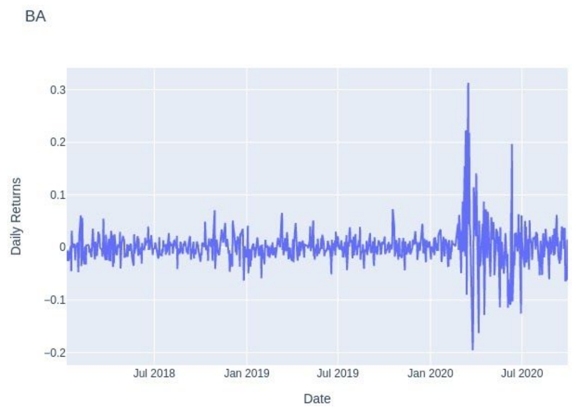
Raytheon Technologies Corporation Aircraft manufacturing 2020

Both these stocks belong to the same sector and we can see which highly relates to automobile manufacturing and tourism , since most of the parts are imported from different countries, the availability of parts reduced and the manufacturing process slowed down, another reason is the flow of money from financial sectors also reduced leading to huge depreciation.

Observing Daily Returns:







On plotting the daily returns we can see how it varies over the course of the year and how the pandemic has affected the returns. At the beginning of all of these graphs we can see that the daily returns do not show much variations which indicate the stability of the market. Once the pandemic becomes more widespread we can see that the variations of these daily returns increases indicating stock instability.

V. Our approach

As seen in the previous section(IV) there are many notable trends in the data which we would like to look into while building our model.

When it comes to our approach to solve the problem we have two important factors that we need to consider. The model and the data for the model itself.

It's important to choose relevant data for our specific problem statement. If it was just ordinary stock market prediction then we would require previous year data and that would be done. As previously discussed, in our current situation that would lead to incorrect results. We solve this.

The financial crisis of 2008 is in some ways similar to the Covid-19 pandemic in terms of the stock market crash. So if we were to take this data along with the data from a few years before 2020, we should have a very good dataset to work with to build our model.

We have 50 stocks varying from large to small cap and they are from different sectors , which helps us in learning about the sectors affected most as well as the least.

Now that we have the data for the model figured out we must now begin to approach the kind of model we want to build for our problem statement.

As seen in previous works, some popular models used were the ARIMA model(Paper D) and the ARIMA-LSTM model(Paper A).

While these models are effective to an extent, we feel that we can get better results with an RNN model, specifically more accuracy and the ability to predict over longer periods of time.

RNN stands for Recurrent Neural Network. To put it simply Recurrent Neural Network(RNN) are a type of Neural Network where the outputs from the previous step are fed as inputs to the current step. The main and most important feature of RNN is the Hidden state, which remembers some information about a sequence.

An RNN model alone might not be able to capture all the trends in the data, so we will also build an ARIMA(Auto-regressive Integrated Moving Average) model to see if it can capture any trends that the RNN model missed out on. How the ARIMA model works will be discussed in a further section.

How does our model stand out from other models? What value are we adding to the solution? In our model we plan to tune hyperparameters in order to allow for more efficient processing and hopefully better accuracy in predictions. We also hope to use the ARIMA and the RNN models to build a hybrid model which uses a combination of the two models to make even more accurate predictions.

DATASET

The dataset has been generated using **pandas_datareader,** which scrapes yahoo finance/ google finance website for details regarding the stock mentioned such as open,high,low ,close,volume. And based on the time period specified by giving the starting date and ending date.

This generates the stock's fluctuations only for the days the market is open, leading to not possessing Nan values in between. So the whole dataset is ready to be trained and tested directly, without any preprocessing or scaling of the dataset. The division has been done using skit.learn for the two time periods of economic recession and global corona pandemic.

ARIMA

ARIMA stands for Auto-Regressive Integrated Moving Average. There exist three main parameters which one should set while defining an ARIMA model, they are called (p,d,q). To understand these parameters we need to fully understand what ARIMA is.

This can be done by splitting ARIMA into different parts i.e the AR part, which tells us about a model which will perform regression on the previous values of a changing variable.

The Integrated part refers to transforming the time series to become stationary. This can be done by changing the current values present in the data to values obtained by calculating the difference between the current values and the previous values.

Lastly we take the MA part which allows us to represent a smoothed out version of the average price of a stock by splitting the data into multiple sets and calculating a series of averages from these splits.

The parameters are indicators of each of these parts(AR,I,MA) and are thus defined as:

p -lag order (for AR part)

d - degree of differencing (for I part)

q - order of MA (for MA part)

We have used auto_arima from pmdarima to calculate the p,d,q value for the arima model. And the ARIMA function was called using statsmodel.

We have used this model to predict the price of stocks to compare the stocks while stabilizing ,after the two time periods from 2006-09 and 2019-20 to see which sector has been affected the most and which the least. This helps in predicting which sector stocks to invest in and till what time ,so that no loss or minimum loss is bared by the investor.

RNN (LSTM)

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feed forward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data.

The model has been trained separately for two time periods, which are the economic recession in 2007-08 and global Covid-19 pandemic, to check and verify which sector has been affected the most and accordingly decide about prioritizing the sectors and its companies to help them stabilize. As companies generate revenue, revenue contributes to development of a sector and development of the whole sector contributes to lifting the economy of the country.

As the model didn't require any preprocessing , it was directly trained and tested by splitting up the dataset for a few companies from each sector.

ARIMA-RNN HYBRID

This model is a combination of ARIMA and Recurrent Neural Network ,the speciality of this model is that we have imported a library TA-Lib which has inbuilt overlap studies of Moving Average functions ,SMA, EMA, WMA, DEMA, TEMA, TRIMA etc. and the model is used to forecast the low volatility time series.

The moving average period is optimized such that the resulting output obeys a normal (mesokurtic) distribution. A normal distribution is defined as a kurtosis K value of 3. For $K > 3$, the data is leptokurtic and for $K < 3$, the data is platykurtic.

This is calculated by finding the kurtosis value of each function for the dataset given, so the MA function varies based on the fluctuation displayed by the stock over a given period of 60 days The length of data used for deciding on the MA function being used has a direct effect on the kurtosis value. With longer data sets, K is less than 3 for all values of moving average period length.

An LSTM (Long Short-Term Memory) recurrent neural network is used to forecast the high volatility time series. The neural network consists of a four neuron LSTM tensor, a two neuron LSTM tensor, and a single one neuron dense tensor.

The high volatility time series is pre-processed with a scalar function to adjust the feature range to between 0 and 1.

The LSTM model is fit with early stopping enabled to minimize potential over-fitting. The predicted low and high volatility time series are summed to generate the predicted closing prices.

The accuracy of the ARIMA-LSTM model is evaluated in several ways. MSE, RMSE, and MAPE are utilized to assess the error between the predicted closing prices and the actual closing prices. Two simulations are conducted to assess the model's predictive accuracy.

WMA

$$\text{WMA} = \frac{\text{Price}_1 \times n + \text{Price}_2 \times (n-1) + \cdots \text{Price}_n}{\frac{n \times (n+1)}{2}}$$

where:

$n = \text{Time period}$

TRIMA

$$trima_t = \frac{1}{\sum w} \sum_{i=0}^{n-1} w_i in_{t-i}$$

The calculation can also be expressed in terms of two Simple Moving Averages with adjusted periods n.

$$trima = \begin{cases} ma^{\frac{n}{2}+1}(ma^{\frac{n}{2}}(in)) & n \text{ is even} \\ ma^{\frac{n+1}{2}}(ma^{\frac{n+1}{2}}(in)) & n \text{ is odd} \end{cases}$$

EMA

$$\text{EMA} = \text{Price}_t \times k + \text{SMA}_y \times (1-k)$$

where:

$t = \text{Today}$

$k = \frac{2}{\text{Number of days in period} + 1}$

$\text{SMA} = \text{Simple Moving Average of closing price for the number of days in the period}$

$y = \text{Yesterday}$

DEMA

DEMA = (2 * EMA(n)) - (EMA(EMA(n))), where n= period

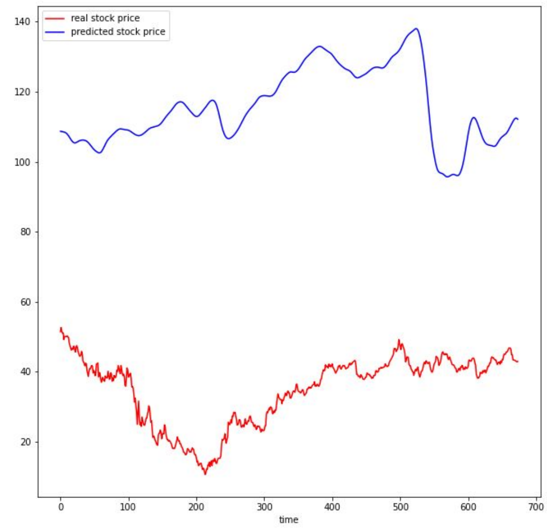
VI. RESULT & CONCLUSION

We have built 3 models- RNN model, ARIMA model and ARIMA-LSTM hybrid model. We will now look at the results for each of these models

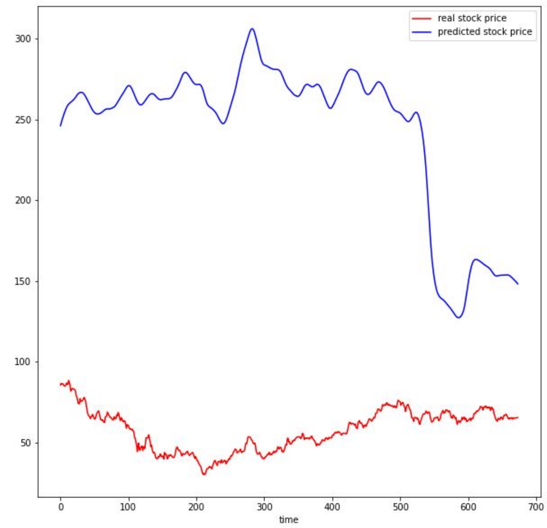
RNN Model

Our RNN model gave promising results for some companies while it gave less promising/failed for others

AXP(American Express Financial services) 2008

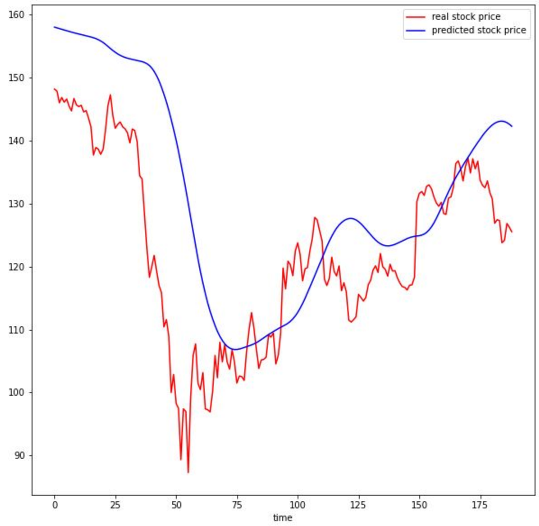


BA(Boeing Aerospace) 2008

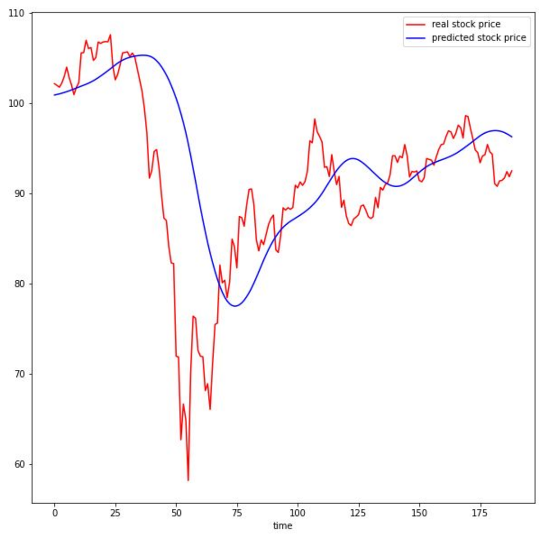


For both the companies seen in the 2008 period the predictions are off because the model cannot account for the severity of the 2008 crisis. We can look at these graphs to see how the stocks of these companies would have been if the 2008 crisis had never happened.

UTX (Raytheon Technologies) 2020



YUM (YUM Brands Fast food) 2020

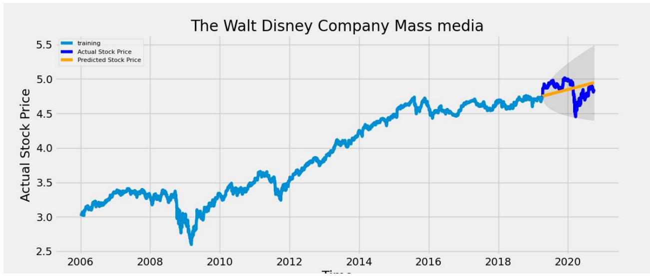


For both the companies seen in the 2020 period, stock prices dropped down and then climbed up which is predicted well by the RNN model. The prediction curve looks like it is a smoothened out version of the actual curve.

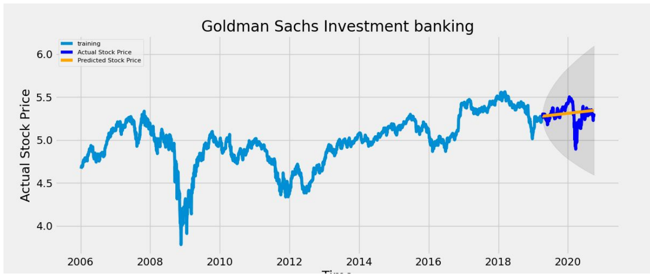
ARIMA Model

The ARIMA model captured linear trends in the data very well and in our results we can see the predicted prices as a straight line and the actual prices are varying along this predicted line

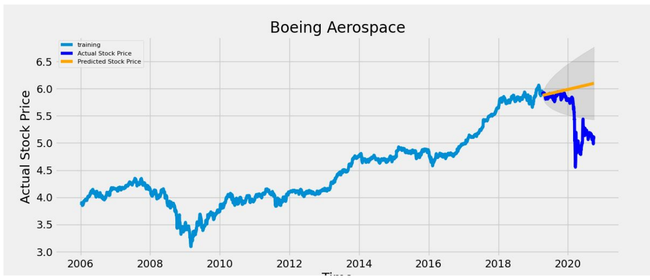
DIS(Walt Disney)



GS(Goldman Sachs)



BA(Boeing Aerospace)



The predictions for DIS and GS were good because the companies were able to bounce back up after the effects of the pandemic. However BA was affected so badly that the predicted and actual prices vary a lot. Again, from this we can see how the stock prices of BA would have been if the Covid-19 pandemic did not happen

ARIMA-LSTM Model

The ARIMA-LSTM hybrid model was supposed to use ARIMA to capture the linear trends and LSTM to capture the nonlinear trends. However the model did not work as well as we expected it to.

```
1
Prediction vs Close:         51.0% Accuracy
Prediction vs Prediction:    46.22% Accuracy
MSE:     323.5499349640634
RMSE:    17.987493848895777
MAPE:    5.490833127018244

2
Prediction vs Close:         52.99% Accuracy
Prediction vs Prediction:    47.81% Accuracy
MSE:     659.8241498287745
RMSE:    25.687042450013088
MAPE:    8.411223935964577

3
Prediction vs Close:         50.2% Accuracy
Prediction vs Prediction:    47.01% Accuracy
MSE:     162.87786625652433
RMSE:    12.76236131194084
MAPE:    4.206384989658252

4
Prediction vs Close:         50.6% Accuracy
Prediction vs Prediction:    46.61% Accuracy
MSE:     316.95786750860594
RMSE:    17.803310577210237
MAPE:    5.238793308460417

5
Prediction vs Close:         52.59% Accuracy
Prediction vs Prediction:    48.61% Accuracy
MSE:     274.3950291643885
RMSE:    16.564873351897035
MAPE:    6.367435051312706
```

The accuracy ranged from 45-55% which was very low and the MSE value was very high. These are all unfavourable results so we have chosen not to use this model for plotting out predictions as we will not get good results.

Conclusion

The ARIMA and RNN models performed well, they predicted stock prices well in the 2008 period as well as the 2020 period.The ARIMA-LSTM results were underwhelming as we had expected better results from it. This goes to show that there are more complicated factors involved when combining two different models together which are hard to catch. Where one model fails to predict the stock prices of a company correctly, others make up for it by predicting them right. No one model works perfectly, but they are all still good indicators of what could happen. Perhaps if we had sector specific data included for the companies we could make more accurate predictions and get consistent results for all companies. Overall we are happy with the results of our models.

VII. BIBLIOGRAPHY

1. Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model - Hyeong Kyu Choi https://arxiv.org/pdf/1808.01560v5.pdf

2. Forecasting on Stock Market Time Series Data Using Data Mining Techniques - A.Subashini , Dr. M. Karthikeyan http://www.ijesi.org/papers/NCIOT-2018/Volume-4/2.%2006-13.pdf

3. HATS: A Hierarchical Graph Attention Network for Stock Movement Prediction - Raehyun Kim Et al. https://arxiv.org/pdf/1908.07999v3.pdf

4. CORONAVIRUS: IMPACT ON STOCK PRICES AND GROWTH EXPECTATIONS - Niels J. Gormsen, Ralph S. J. Koijen https://www.nber.org/papers/w27387.pdf

5. A Prediction Approach for Stock Market Volatility Based on Time Series Data - Sheikh Mohammad Idrees Et al. https://ieeexplore.ieee.org/abstract/document/8626097

6. Comparative study on Prediction of Support and Resistance Levels with k-Nearest Neighbor and Long Short-Term Memory Methods - Thambusamy Velmurugan , Thiruvalluvan Indhumathy http://www.joics.org/gallery/ics-4035.pdf

7. Industries Hit Hardest by the Recession https://www.cnbc.com/2012/06/01/Industries-Hit-Hardest-by-the-Recession.html

8. A guide to the financial crisis — 10 years later https://www.washingtonpost.com/business/economy/a-guide-to-the-financial-crisis--10-years-later/2018/09/10/114b76ba-af10-11e8-a20b-5f4f84429666_story.html

VIII. Contribution:

Suhas: Literature Survey, Data Retrieval, Pre-processing, Built Arima Model

Amogh: Literature Survey, EDA, Built Hybrid ARIMA+LSTM Model

Chirag: Literature Survey, EDA, Built RNN Model