Stock price prediction

Applied data science

Phase 4 submission document

Register no:-961721106307

Name:- Madhu suthanan M

Efficiency analysis of stock price prediction

Abstract

Accurate prediction of stock market returns is a very challenging task due to volatile and non-linear nature of the financial stock markets. With the introduction of artificial intelligence and increased computational capabilities, programmed methods of prediction have proved to be more efficient in predicting stock prices. In this work, Artificial Neural Network and Random Forest techniques have been utilized for predicting the next day closing price for five companies belonging to different sectors of operation. The financial data: Open, High, Low and Close prices of stock are used for creating new variables which are used as inputs to the model. The models are evaluated using standard strategic indicators: RMSE and MAPE. The low values of these two indicators show that the models are efficient in predicting stock closing price.

INTRODUCTION

Since 1950's people are buying and exchange stocks in the stock market. Before, it was hard to buy one because you need to have a broker and then call them and state how many stocks you would buy. Nowadays, people just open a brokerage account online and buy stocks without having to call a stock broker. Stock market is where people buys shares or stocks of a company that is public listed. A public listed company are companies that is listed in the stock exchange market. Meaning, a public listed company is public owned, no one owns a 100% stake of it thus it has no owner except for the person that owns the most stake of the company. Most of the time it's the corporate stock investors and sometime from retail investors that buys the stocks. Stocks or shares is a piece of a company paper that you can buy and own a tiny proportion of that company. Say you buy one unit of stock from Apple Inc., you own the company by one unit of their entire value. The company valued at 2Trillion USD, so if you buy say 100units, you own about 0.00000075% stake of Apple Inc. based on their current value at November 2021.

Corporate investors and retail investors are different from each other. Corporate investors are like banks, hedge funds, and any company that buys stocks, while retail investors are individual investors like me, you, your friends and more. This final year project can benefit those two entity whether you are corporate or retail investors.

The targeted user for this project will be the investors and also can be for the finance college student in practicing how to invest. Stock price prediction analysis in this project is the act of predicting the future price or value of a certain company stock traded on the stock exchange/market. An accurate prediction of a stock's future price could bring significant profit to investors. The efficient-market hypothesis or (EMH) suggests that stock's current prices reflect all currently available information and any price changes that are not based on newly revealed information thus it is inherently unpredictable (Fama & Eugene, 1970, p. 383). Others differ and those with this perspective in this title such as myself, have few techniques and innovations in technology which purportedly permit the investors to acquire future value data.

This final year project aims to predict the stock price in the stock market for future profits. The way the investors earn money, whether corporate or retail, is from capital gains. A capital gains is profits that you earn from selling stock that you bought. For example, you bought a stock for RM5 per share, then the stock goes up at RM5.50 per share. Then, you sell the stock, so your capital gains are RM0.50 per share.

Predictions on stock market exchange prices are an incredible test because of the way that it is an enormously perplexing, tumultuous and dynamic environment. There are many research from different regions meaning to take on that test and Machine Learning approaches have been the focal point of large numbers of them. There are numerous instances of Machine Learning calculations had the option to arrive at good outcomes while doing that sort of expectation. This final year project concentrates on the utilisation of LSTM networks on that situation, to anticipate future patterns of stock prices dependent on the value history, close by with specialised investigation pointers. For that objective, a prediction model was constructed, and a progression of experiments were executed and theirs outcomes broke down against various measurements to survey in case this kind of calculation presents and enhancements when contrasted with other Machine Learning techniques and venture procedures. The outcomes that were acquired are promising, getting up to a normal of 55.9% of exactness while foreseeing in case the cost of a specific stock will go up or not sooner rather than later.

1.1 Problem Statement

Predicting how the stock market price will perform is one of the most difficult things to do even without the use of Machine Learning (ML). There are so many factors involved in the prediction physical factors vs. psychological, rational and irrational behaviour, etc. All these aspects combine to make share prices volatile and very difficult to predict with a high degree of accuracy.

Stock investment nowadays is a type of income that most people have, especially the wealthy. It is an income that is called as portfolio income. An income received from investments, dividends and capital gains. Unlike earned income, is a type of income that received from jobs or company that you worked for.

Predicting the stock price can ease the work of investors in investing their next company's stock of choice. Now, investors of way in predicting the future price of a stock is by doing bunch of analysis using formulas and they need to calculate by themselves. The analysis could be technical analysis, P/E Ratio, earnings per share and more. By developing a Machine Learning model, all of that can easily be automated and calculated by a computer algorithm without the need of having a human supervision.

- If you realise, the stock market price appears in the news every day, specifically at the bottom of your screen tv. We hear about it every time when the price reaches an alltime high or a new high or a new low. The percentage of investment and business opportunities in the Stock exchange market can increase if an efficient algorithm could be devised to predict the short term price of an individual stock (Joish Bosco, 2018).
- In this final year report. we will see if there is a possibility of devising a model using an LSTM Neural Network which will predict stock price with a less percentage of error in the final output. We will also see how reliable and efficient this model will be.

1.2 Objective and Scope of Study

1.2.1 Objective

This web-based or interface-based stock prediction analysis dashboard will be developed to ease the work of targeted users like entrepreneurs, investors, businesses, and more in making their future decisions in investing in the stock market. The stock price data that gathered for training and predicting the future price on the web-based dashboard will be useful for the users to make analysis as the data is a real-time data. The Machine Learning model that this project develop extracting stock price data using scraping method which is extracting the data directly from website, specifically from Yahoo Finance.

- To design an AI based technology with Machine Learning algorithm in assisting the users in making predictions for their stock investment.

- To develop a dashboard platform using Data Scrapping techniques in gathering the data with python technology and predicting the price using the data using ML algorithm called as LSTM Neural Network.
- To develop a user-friendly interface and dashboard to output the result of prediction analysis of stock price using the ML algorithm.

1.2.2 Scope of Study

This project will help the targeted users to improve their future decision in making more effective and more accurate and to solve the limitation issues as well in the stock market industry. This study will be focusing on the investors, financial analyst, economist, businesses, individuals and as well as finance college student as the users. Hence, the research and development process of this study will consist of both targeted users and the variety of stock investing platforms and their features. The details of the scope can be classified into two categories below:

- Targeted Users

  Users that will mainly involve with this project's platform are the users in the business world, critics, and financial and economy analyst. The users will make use of this project's dashboard to make analyst to make their own future decision making process. This dashboard-based prediction model will have a feature where some finance strategies and decision making will be done by Artificial Intelligence (AI) and Machine Learning (ML). This will ease the work of the analyst out there in making their task for future decision making.

- Development Tools

  The proposed development tools for this study will be a Machine Learning algorithm using the LSTM (Long-Short Term Memory) Neural Network for the analyst of the stock price data and a simple data scraping technique for extracting the data from the website (yahoo.finance.com). All the technologies involved will be integrated with each other to make this study a success. The implementation of the AI and ML technology on the dashboard will be developed via the backend development of the dashboard with machine learning model and integrated with the data scraping techniques for the stock data for the ML model to make analysis.

CHAPTER 2

LITERATURE REVIEW

LSTMs Neural Network is widely used around the world for its sequence prediction problems. The method also have proven to be extremely effective. The particular reason for circumstance they work so perfectly well is because the LSTM Neural Network is able to store past information that is important, and forget the information that is not (Stock Prices Prediction Using Machine Learning and Deep Learning Techniques (, n.d.). LSTM has three gates for its implementation which are:

1. **The Input Gate:** The input gate adds in the information to the cell state.
2. **The Forget Gate:** This gate eliminates the information that is no more required by the ML model.
3. **The Output Gate:** This gate in this algorithm which is LSTM selects the information to be shown as output.

2.1. LSTM Neural Network for Predicting Stock Market's Price Movement

Forecasts or predictions on securities exchange (Stock Market Exchange) costs are an incredible test because of the way that it is an enormously perplexing, tumultuous, and dynamic climate. There are many investigations from different regions meaning to take on that test and Machine Learning approaches have been the focal point of a considerable lot of them. There are numerous instances of Machine Learning calculations had the option to arrive at good outcomes while doing that sort of expectation. This literature concentrates on the use of LSTM networks on that situation, to anticipate future patterns of stock costs dependent on the value history, close by with specialised examination markers. For that objective, a prediction model was constructed, and a progression of trials were executed and theirs outcomes dissected against various measurements to survey assuming this kind of calculation presents and enhancements when contrasted with other Machine Learning techniques and venture systems. The outcomes that were acquired are promising, getting up to a normal of 55.9% of exactness while foreseeing assuming the cost of a specific stock will go up or not soon (David M. Q.

Nelson, n.d.).

2.1.2. Background and Research

With regards to stock market exchanges, notwithstanding its inborn complexity and dynamism, there has been a steady discussion on the prediction consistency of the stock returns. (Malkiel, 1970) presented the Efficient-Market speculation that characterizes that the current cost of a resource consistently mirrors generally past data accessible for it instantly. On the other hand, there are different authors who guarantee that, truth be told, stock market prices can be predicted essentially somehow (MacKinlay, 1999). Furthermore, an assortment of strategies for predicting and displaying stock behaviour have been object of investigation of various disciplines, like financial aspects, insights, physical science and software engineering. Furthermore, an assortment of strategies for predicting and displaying stock behaviour have been object of investigation of various disciplines, like financial aspects, insights, physical science and software engineering. It's worth focusing on that in 2012, it was assessed that roughly 85% of exchanges inside the United States stock exchanges were performed by calculations (M. Glantz, 2013). A well-known strategy for modelling and predicting the stock market exchange is technical analysis, which is a technique dependent on historical data from the market, primarily price and volume. It follows a few suspicions which are, the prices are characterised solely by the supply and demand connection. Then, at that point, stock prices change following inclinations, changes on the supply and demand cause tendencies to invert. Changes on relation of the market can be recognised on graphs and examples on patterns will quite often repeat (Dahlquist, 2006). As such, technical analysis do not consider any outer variables like political, social or full scale prudent.

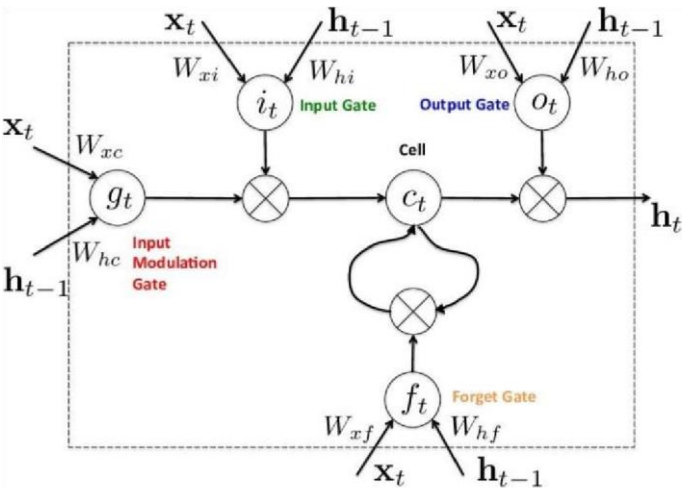


Figure 1: Long Short-Term Memory (LSTM) Gates Diagram

Long Short Term Memory (LSTM) neural network (Figure 1), which are utilised in this task are a profound and repetitive model of neural network. Recurrent network contrast from the customary feed-forward networks as if they don't just have neural associations on a solitary bearing, all in all, neurons can pass information to a past or a similar layer. In which case, data doesn't stream on a single way, and the down to earth impacts for that is the presence of short term memory, long term memory that neural network as of now have in outcome of preparing. LSTM were presented by (Schmidhuber, 1997) and it focused on a superior presentation by handling the evaporating angle issue that recurrent network would endure when managing long data arrangements. This project propose a model based on LSTM to foresee stock market price exchange movements developments utilising an info that did not depend on text, it is not something that has been broadly study. This predictions project intends to utilise a wide scope of specialised pointers to do as such, and the aim is to evaluate the use of such technique that is something usually utilised on investment strategies. Also, we need to test the theory that the short term memory capacity can introduce better outcomes contrasted with traditional feed forward networks.

2.2. Prediction using the Stock Market Analysis ways.

Stock market exchange are one of the most entrancing innovations within of our time. They essentially affect numerous regions like business, academic, occupations, technology and accordingly on the economy (Hiransha, 2018). Throughout the years, investors and analysts have been keen on creating and testing models of stock price conduct or behaviour (Fama, 1995). However, analysing stock market movements and price behaviours is extremely challenging because of the markets dynamic, nonlinear, nonstationary, nonparametric, noisy, and chaotic nature (Abu-Mostafa, 1996). According to (Zhong, 2017) , stock market exchanges are impacted by many profoundly interrelated elements that incorporate economic, political, mental, and organisation-specific factors. Fundamental and technical analysis are the two main ways to deal with analysis of the stock market sectors. To put money or invest into stocks and accomplish high profits with low risk, investors have utilised these two significant ways to deal with settle on choices in the financial markets (Arévalo, 2017).

Hu et al. (2015), said that, Fundamental analysis of stock prices is primarily founded on three fundamental angles (I) macroeconomic analysis like Gross Domestic Product (GDP) and Consumer Price Index (CPI) which examinations the impact of the macroeconomic climate on the future benefit of a firm or company, (ii) industry analysis which assesses the worth of the organisation dependent on industry status and prospect, and (iii) organisation analysis which examines the current activity and monetary status of an organisation to assess its inward value. Distinctive valuation approaches exist for fundamental examination. The normal development estimate method contrasts Stock-A and different stocks in similar class to all the more likely get valuations, i.e., accepting two companies have a similar development rate, the one with the lower Price-to-Earnings (P/E) proportion is viewed as better. Thus the reasonable cost is the profit times target P/E. The P/E technique is the most ordinarily utilized valuation strategy in the stock business industry.

Numerous new technologies and strategies have been proposed throughout the years to attempt to predict the stock prices by means of numerous roads, thanks to the difficult and always changing scene of stock market exchanges (Chen, 2016). In this final year project, will be focus around two themes, in particular, stock analysis and stock prediction. This project checks out the research's past, however principally focus around current techniques, featuring a portion of the primary difficulties they posture and late accomplishments for stock analysis and prediction.

2.3 Taxonomy of Stock Market Price Analysis Approaches.

Ongoing progressions in stock analysis and prediction fall under four classifications which are statistical, pattern recognition, AI or machine learning (ML), and sentiment analysis. These classes for the most part fall under the more extensive classification of fundamental analysis. Be that as it may, there are some ML methods which likewise consolidate the more extensive classes of technical analysis with the fundamental analysis way to deal with predicting the financial or stock market exchanges. Figure 2 below shows a scientific categorisation of famous stock prediction techniques. These strategies have acquired prominence and have shown promising outcomes in the field of stock analysis in the new past.

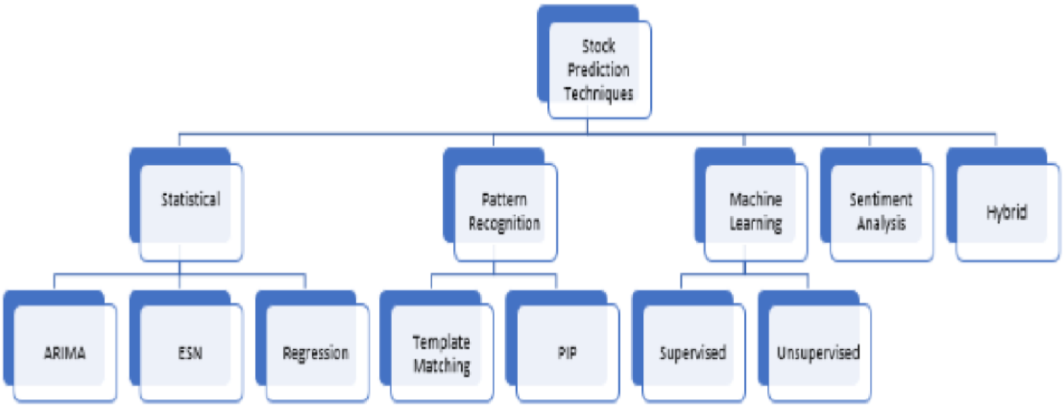


Figure 2:
Taxonomy of Stock Prediction Techniques.

Prior to the appearance of machine learning strategies, statistical procedures which frequently accepts linearity, stationarity, and ordinariness gave a method for breaking down and predict stock prices. Time series in the stock or financial exchange analysis is a sequential assortment of observations, for example, day by day deals aggregates and prices of stocks. As indicated by Zhong and David Enke (2017), one gathering of measurable methodologies which fall into the class of univariate examination, because of their utilisation of time series as information factors, are the Auto-Regressive Moving Average (ARMA), the Auto-Regressive Integrated Moving Average (ARIMA), the Generalised Autoregressive Conditional Heteroskedastic (GARCH) instability, and the Smooth Transition Autoregressive (STAR) model. The ARIMA model is a broadly utilized strategy for financial exchange examination. ARMA consolidates Auto-Regressive (AR) models which attempt to clarify the energy and mean inversion impacts frequently saw in exchanging markets and Moving Average (MA) models which attempt to catch the shock impacts saw in time series. A vital limit of the ARMA model is that it doesn't consider unpredictability bunching, a vital experimental peculiarity in numerous monetary time series.

There are a lot of approaches in predicting the stock prices whether its human way like fundamental analysis or machine way which is machine learning (ML) or AI. This literature explains the approaches. However, in this final year project paper, we will cover the machine learning approach that has been proposed which is Machine Learning (ML) technique. Machine learning in AI has been broadly read up for its possibilities in the predictions of stock market business sectors (Shen, 2012). Machine learning assignments are extensively grouped into supervised and unsupervised learning. In supervised learning, a bunch of named input data for preparing the calculation and noticed result information are accessible. Be that as it may, in unsupervised learning, just the unlabelled or noticed result data is accessible. The objective of supervised learning is to prepare a calculation to naturally plan the info information to the given result information. At the point when prepared, the machine would have figured out how to see an input data relevant item and predict the expected result. The objective of unsupervised learning is to prepare a calculation to track down an algorithm, relationship, or group in the given dataset. It can likewise go about as an precursor for supervised learning assignments (Bhardwaj, 2015). A few calculations have been utilized in stock price bearing prediction. Less difficult procedures, for example, the single decision tree, discriminant analysis, and naive Bayes have been supplanted by better-performing calculations like Random Forest, logistic regression, and neural networks. With nonlinear, information-driven, and simple to-sum up attributes, multivariate analysis using profound Artificial Neural Networks (ANNs) has turned into a predominant and famous examination apparatus in the stock market analysis (Zhong, 2017). As of late, profound nonlinear neural organization geographies are starting to stand out in time series prediction.

CHAPTER 3

METHODOLOGY/PROJECT WORK

3.1 Development Methodology.

In developing this proposed project, the development cycle or strategy that will be use is the Waterfall Model Approach. This methodology is the soonest SDLC (Software Development Life Cycle) approach in programming or web advancement. The waterfall approach is one of two well-known strategies to handle programming projects. The other technique is known as Agile. It tends to be more obvious cascade when you analyse it excessively Agile. Waterfall and Agile are two totally different task the management methodologies, however both are similarly legitimate, and can be valuable relying upon the undertaking.

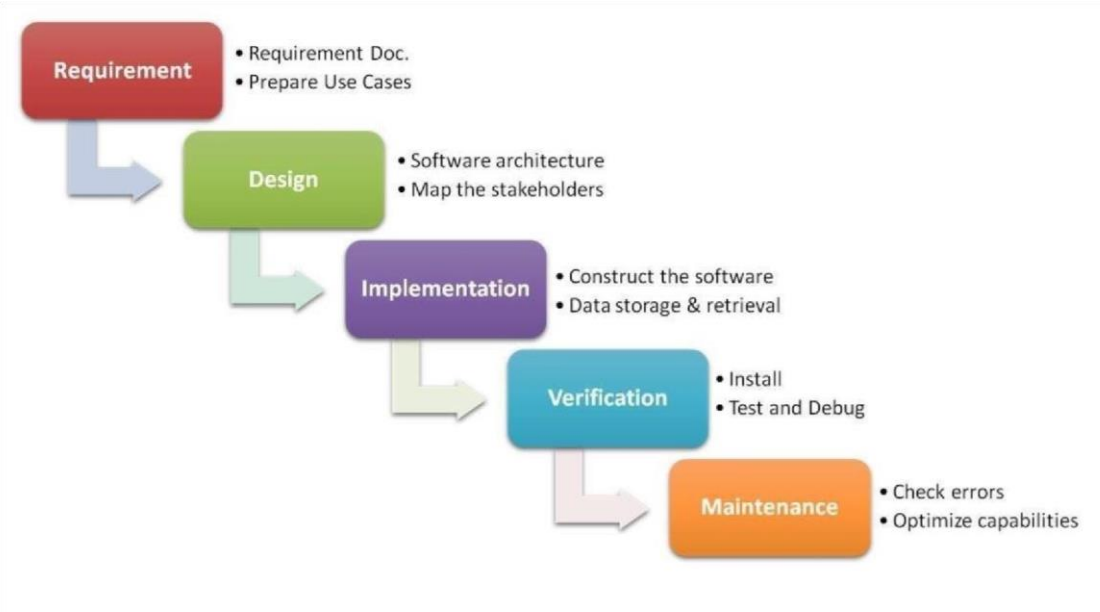


Figure 3: Waterfall Methodology

- Requirements

  The Waterfall methodology relies upon the conviction that all project necessities can be accumulated and perceived forthright. The project supervisor gives a valiant effort to get a definite comprehension of the task support's prerequisites. Composed prerequisites, normally contained in a solitary archive, are utilized to depict each phase of the venture, including the expenses, presumptions, hazards, conditions, achievement measurements, and courses of events for consummation.

- Design

Here, programming engineers plan a specialized answer for the issues set out by the item prerequisites, including situations, formats, and information models. Initial, a more elevated level or coherent plan is made that portrays the reason and extent of the task, the overall traffic stream of every part, and the reconciliation focuses. When this is finished, it is changed into an actual plan utilizing explicit equipment and programming advancements.

- Implementation

  When the plan is finished, specialized execution begins. This may be the most brief period of the Waterfall interaction on the grounds that careful examination and configuration have as of now been finished. In this stage, software engineers code applications dependent on project necessities and details, with some testing and execution occurring too. In the event that critical changes are needed during this stage, this might mean returning to the plan stage.

- Verification

  Before a project can be delivered to clients, testing should be done to guarantee the item has no blunders and every one of the necessities have been finished, guaranteeing a decent client experience with the product. The testing group will go to the plan archives, personas, and client case situations provided by the item director to make their experiments.

- Deploy and Maintenance

  When the product has been conveyed on the lookout or delivered to clients, the upkeep stage starts. As deformities are found and change demands roll in from clients, a group will be allocated to deal with updates and discharge new forms of the product.

3.2 Development Tools

The development tools that will be used for this improvement of this undertaking are the Microsoft visual studio. It will be utilized as the IDE for the development of the machine learning prediction model. It is suitable and it is open source, so it is free. Not just that, Microsoft Visual Studio is likewise viable to most programming dialects out there. This is an extraordinary IDE as that the improvement of this undertakings utilizes different programming dialects to begin coding. The programming dialects includes are Python for machine learning model advancement, and Dash framework from python for the User Interface (UI) of this model.

Figure 4: Microsoft Visual Studio Logo

3.3 Prototype/Machine Learning model

To visualized how this project is going to be developed, a graphic user interface UI/UX is created to start the development process for this ML model platform. It is important for development process that the UI/UX needs to be done first before starting the implementation phase. This study is going to be develop model-based/dashboard based, so the user interface for this project is a website-based UI. There will be sections for this UI. The model will be like dashboard type and there will be sections where the ML model predicts the stock price, and a section only for analysis of stock price movements for blue chip stock/big companies like Facebook, Apple, and others.

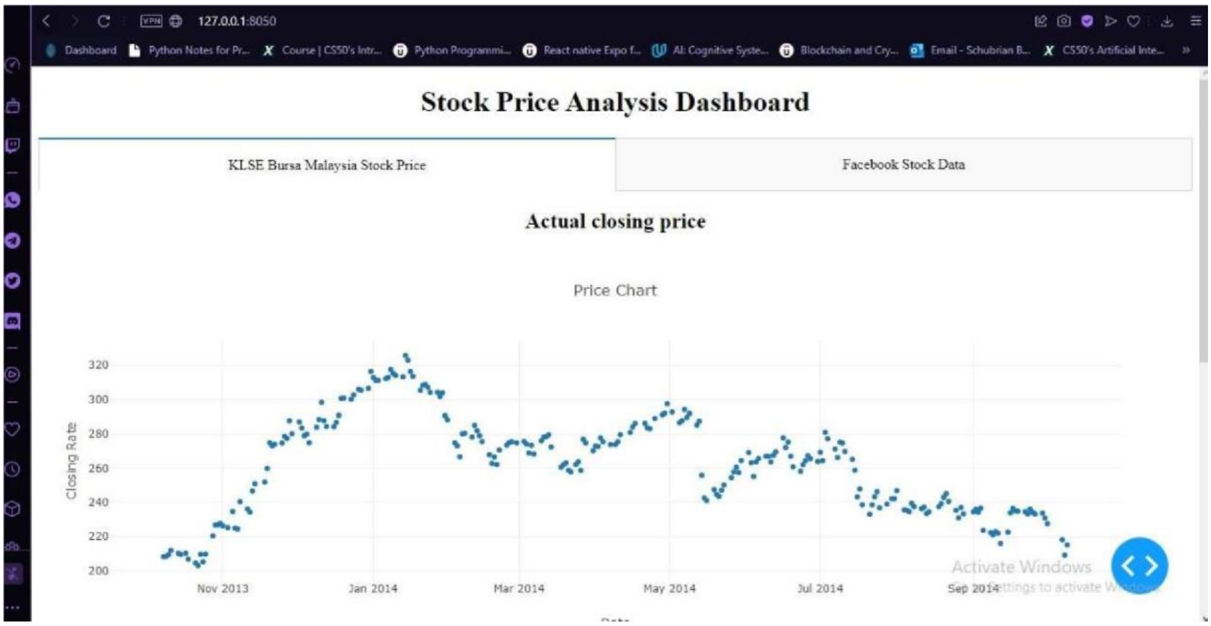


Figure 5: ML model for prediction dashboard.

In this part, this model will construct a dashboard to examine stocks. Dash is a python framework that gives deliberation over flask and react.js to assemble logical web applications. Composed on top of Plotly.js and React.js, Dash is great for building and conveying data applications with altered (UI). It's especially appropriate for any individual who works with information and data.

Through several straightforward examples, Dash abstracts away the innovations in general and conventions that are needed to assemble a full-stack web application with intuitive information perception. Dash is straightforward enough that you can tie a UI to your code in under 10 minutes. Dash applications are delivered in the internet browser. You can send your applications to VMs or Kubernetes groups and afterward share them through URLs. Since Dash applications are seen in the internet browser, Dash is innately cross-stage and portable prepared.

There is a ton behind the structure/framework.

Dash is an open-source library delivered under the tolerant MIT license. Plotly creates Dash and furthermore offers a stage for composing and conveying Dash applications in an enterprise climate.

3.3.1 Data Set and data exploration

In this project, the dataset used is a variety of stocks or company data to make use for analysis and prediction. The dataset consists of stock data from Malaysia stock which is Kuala Lumpur

Stock Exchange (KLSE) and other blue chip company stocks like Facebook, Apple, and Tesla. The prediction model will only make use of the KLSE data and other stocks data will make use for analysis and performance purposes like viewing the trends and movements.

```
df = web.DataReader('^KLSE;1=9', data_source='yahoo',
                    start='2018-01-01', end='2021-10-13')
```

Date	High	Low	Open	Close	Volume	Adj Close
2018-01-02	1783.479980	1772.000000	1783.099976	1782.699951	82478600.0	1782.699951
2018-01-03	1795.869995	1785.209961	1785.790039	1792.790039	167780400.0	1792.790039
2018-01-04	1803.449951	1795.560059	1795.560059	1803.449951	161150500.0	1803.449951
2018-01-05	1817.969971	1803.810059	1804.790039	1817.969971	180760700.0	1817.969971
2018-01-08	1832.150024	1813.030029	1819.739990	1832.150024	176714600.0	1832.150024

The dataset consists of data that are available to extract and download, like the KLSE data, new or recent data are difficult to find as it has insufficient sources and other stocks. The data extracted are from the recent years 2014-2017 data to make use for analysis. The data head for KLSE stock can be shown below. The method for below in Figure 5 shows how this project extract the data which is directly from a site using the Pandas library from python as it has “webReader” command where user ask the machine to extract the data.

Figure 6: ML model for prediction dashboard.

Next, we will be visualizing the data for prediction using the matplotlib library from python. This is to ensure the data extracted can be visualize, hence, can make a prediction with the visualized data.

The figure 6 below shows the closing price for KLSE (Bursa Malaysia Stock Price) according to the date shown above for extraction. The data that is taken is using web-scraping technique where the data received is real-time and directly from website which is yahoo finance.

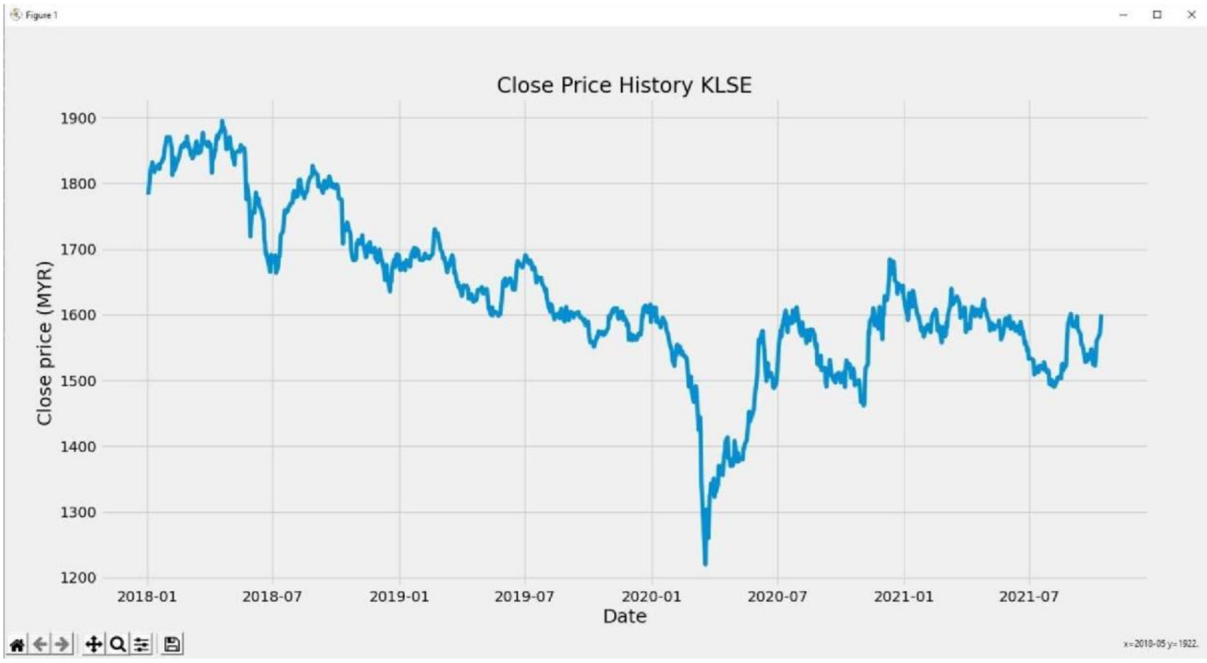


Figure 7: KLSE Price real-time dashboard

This chart as of now expresses a lot of things. The explanation I picked this stock over others is that this diagram is overflowing with various practices of stock price after some time and in Malaysia, this stock is important as its behaviour determines the other stock behaviour in the market, specifically in the KLSE (Kuala Lumpur Stock Exchange). This will make the learning more powerful just as give a change to test how great the prediction is for an assortment of situations.

3.3.2 Data Abstraction

The data abstraction is a step of finding the resource to the very best in categorized the above datasets and learning the best out of it. It is to customize the dataset in finding the best constraints to consider and the unwanted resources are the dump which will be dumped.

Datasets are cleared on this progression for the starting system. The significant dataset is the set that carries the worth to the dataset for a superior agreement and gives a superior yield and creation by assessing something very similar.This is a component deliberation model to remove the including of the dataset. This is a component model cycle where every one of the plausible assets is arranged and a similar will be being used for the featuring.

3.2.3 Training and test the dataset for prediction

The following stage is to split the dataset into training and test sets to keep away from overfitting and to have the option to examine the speculation capacity of the model. Officially, overfitting alludes to the circumstance where a model learns the data yet, in addition, the noise that is important for preparing data to the degree that it contrarily impacts the presentation of the model on new unseen data.

In other words, the noise mentioned above (for example random fluctuations) in the preparation set is learned as rules/patterns by the model. In any case, these loud learned portrayals don't matter to new unseen data and consequently, the model's presentation (for example accuracy, MSE, MAE) is adversely affected. MSE, (Mean Squared Error), MAE, (Mean Absolute error).

The "'target" value to be predicted in the model will be the "Close" price stock price value. It's a smart thought to standardize the dataset before model fitting. This will support the performance.

While Implementing any LSTM, we ought to consistently reshape our X train in threedimensional, add 1 the purpose for is the time step and the 1 is given to the LSTM. Code:

```
X_train =X_train.reshape(X_train.shape[0],X_train.shape[1] , 1)

X_test = X_test.reshape(X_test.shape[0],X_test.shape[1] , 1)
```

Then, at that point, import required modules for the stacked LSTM.

Sequential for initializing the neural network, Dense for adding a densely connected neural network layer, LSTM for adding the Long Short-Term Memory layer, Dropout for adding dropout layers that prevent overfitting

The modules are tensorflow.keras.models, tensorflow.keras.layers, tensorflow.keras.layers import LSTM. Then, utilize a successive model and add the layers of the LSTM as said, in the above sentence.

Then, will add the LSTM layer and later add a few Dropout layers to prevent the datasets from overfitting.

When characterizing the Dropout layers, we determine 0.2, implying that 20% of the layers will be dropped. From there on, we add the Dense layer that determines the result of 1 unit.

After this, I accumulate the model utilizing the well-known Adam streamlining agent and set the misfortune as the mean_squarred_error. This will process the mean of the squared blunders. Then, fit the model to run on 100 ages with a group size of 32. It is contingent upon the specs of the PC, this may require a couple of moments to complete the process of running.

```
model=Sequential()
model.add(LSTM(50,return_sequences=True,input_shape=(100,1)))
model.add(LSTM(50,return_sequences=True))
model.add(LSTM(50))
model.add(Dense(1))
model.compile(loss='mean_squared_error',optimizer='adam')
```

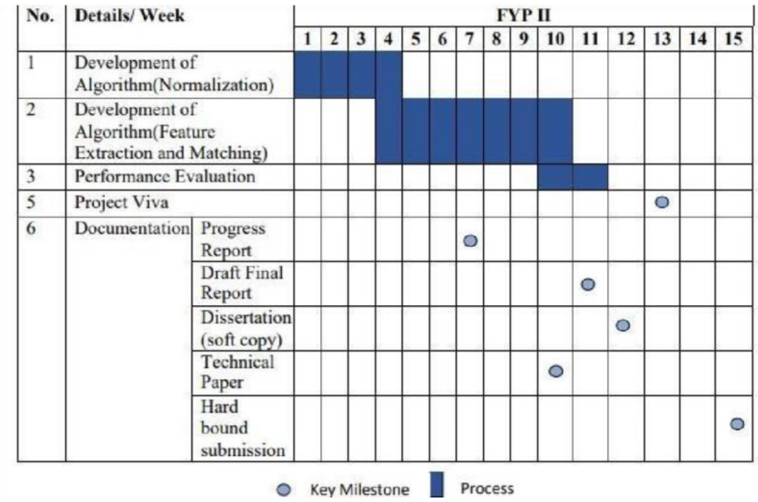
Lastly is to fit the X_train and the y_train.

3.4. Gantt Chart

Table 1: FYP1 Gantt Chart

No.	Details/ Week	FYP 1													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Background Study	■	■												
2	Literature Review	■	■	■	■	■	■								
3	Extended Proposal	■	■	■	■	■	■								
4	Proposal Defense									○					
5	Data Acquisition						■	■							
6	Development of Algorithm (Localization)								■	■	■	■	■		
5	Interim Report (Draft)													○	
6	Interim Report (Final)														○

Table 2: FYP2 Gantt Chart



CHAPTER 4

RESULT AND DISCUSSION

4.1. Predicting the stock market prices using the test set.

To predict the stock price, we really want to do several things in the wake of stacking in the test set:

1. Combine the training dataset and the test set on the axis = 0.
2. Then, set the time step as 60.
3. Use the model which is MixMaxScaler to transform the new dataset.
4. Reshape the dataset.

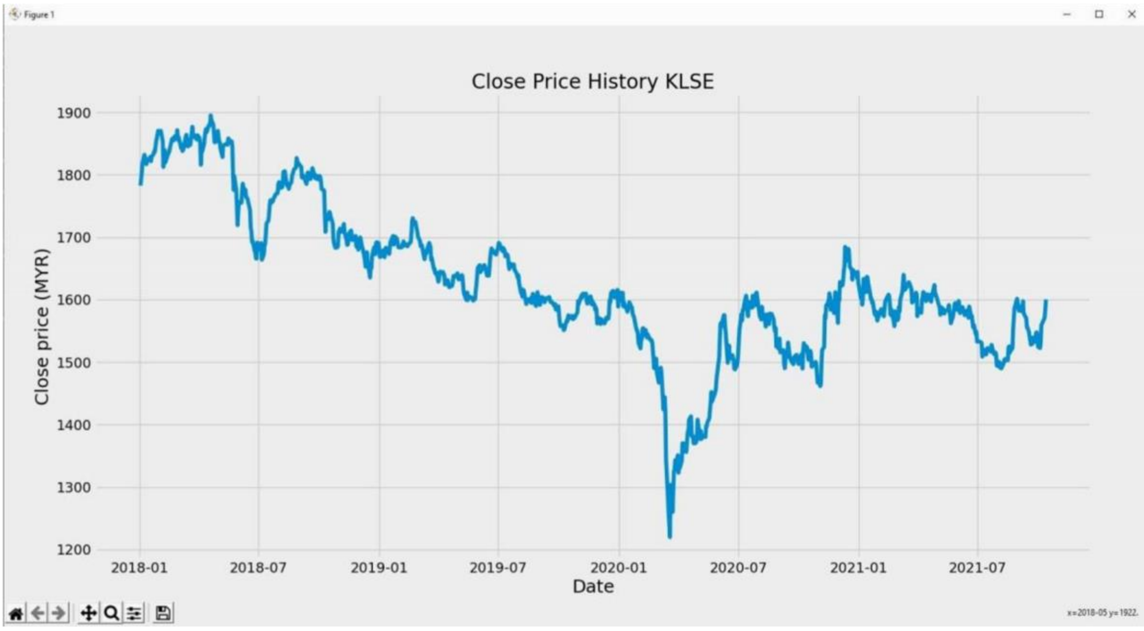


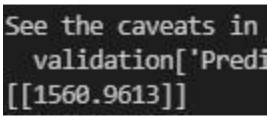
Figure 8: Before prediction KLSE closing price

The above graph shows the real time data of KLSE stock price or other words is, current data. You can see, at the end of the graph, the graph shows the price at RM1600 per unit of KLSE's stock. Next, you will see the machine predicts the price and the outcome is almost quite accurate as the real-time data which makes the LSTM model accurate.

As for observation, you can see the price had a major drop in price somewhat around March 2020. This is because, due to Covid-19 Lockdowns and investors are afraid of losing their investments, hence they sell, hence the huge drop in price.

Figure 9: After prediction closing KLSE stock price

On the figure above, you can see that the model predicted the price as accurate as the output in figure 7. The machine accomplishes its objective in predicting the price and the accuracy for the predictions is at good percent.



The result above shows the exact data of stock price for the prediction at 8/10/2021. Below is the actual price for 8/10/21

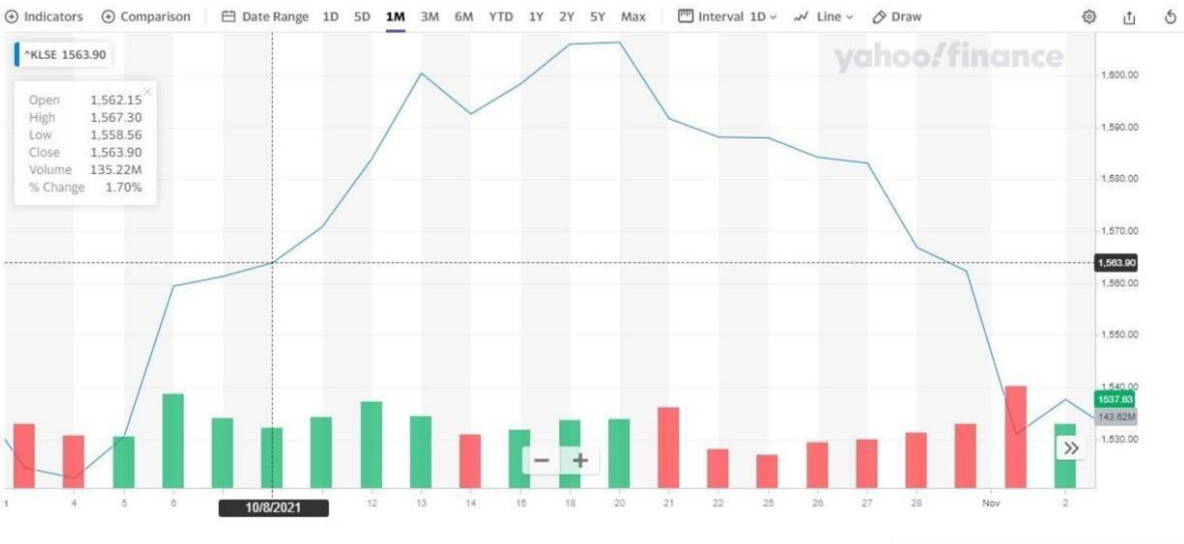


Figure 9. Actual price KLSE stock. Which is 1562.57.

4.2. Prediction on a dashboard using Dash Framework to visualize on a browser.

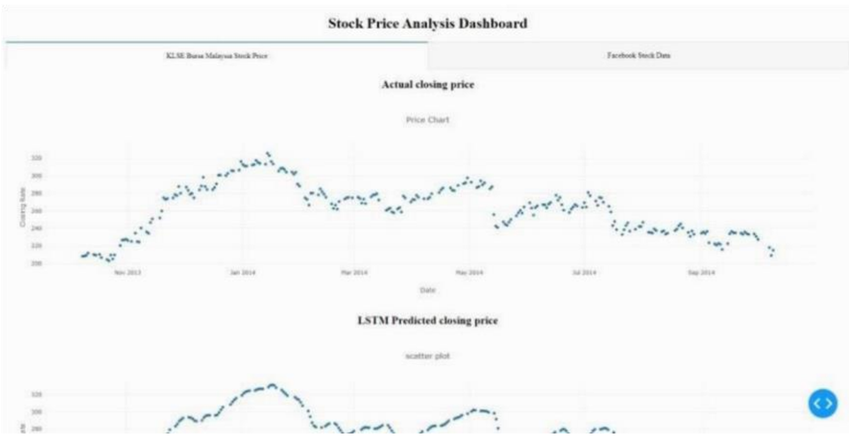


Figure 10: Prediction Dashboard

Figure 11: Analysis dashboard for investment decision

The figure above is where the output for predictions to be displayed and for user to interact with the chart and make their own analysis. For the first page, the KLSE Bursa Malaysia Price page is where the model makes it predictions on the stock price. The other page is where the extracted data from big companies like Facebook, Apple and Tesla is displayed. The reason is to display the big company's stock its low and high price for investment future investment decision. The page has a feature where user can compare the company stock and view it as they like.

4.3. Discussion

Stock price movements varies from a lot of aspects. That includes economic, politics, academic, company situation, and more. The prediction is just one of the ways that can be used to know the stock's price based on current movements.

This is to ease the work of investors, individual, analyst and maybe politicians. Stock market performance can affect how companies perform. If the price of a certain stock goes up, we can conclude that said company is performing at its best at that current time.

Dashboard that this project has created is interactive for users to use. It has the feature for users to view the predicted price for the stock chosen. Although, the predicted stock in this development project is the one that personally chosen, but the behaviour of the chosen stock can determine how other stock in the exchange perform. The reason is to that, is because I chose the stock as it is an Index fund where all the stock in the exchange (KLSE) combined.

So, technically the model that is developed is predicting all the stock in the exchange.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

To sum up everything that has been stated, stock market price is difficult to predict still. The reason is to that, it is a curve which keeps changing and turning on prices from low to high and high to low. Stock market is manipulated. What it means by that, is that nowadays, you can manipulate the market by doing simple things like trending on social media.

Recent example, when a group of young people went on Reddit (Social media platform) and ask their followers to buy a stock from the New York stock exchange so the price of that stock will go up and eventually they sell all, then the stock is back to its low prices. This is an example how you can manipulate the stock market.

However, if the stock market is manipulated regularly, a machine can learn from their pattern and can still predict the stock prices or outcome. Hence, this project's objectives can still be achieved as it a machine that learns throughout time.

5.2. Recommendations.

Stock market business sectors give an exceptional stage to exchanging and contributing, where exchanges can be executed from any gadget that can associate with the Internet. With the coming of stock market exchanges, individuals have the chance to have numerous roads to make their speculations develop. That, however, it likewise brought about various sorts of assets like common assets, speculative stock investments, and list assets for individuals also, establishments to put away cash as per their danger hunger. State-run administrations of most nations contribute a piece of their medical care, work, or retirement assets into financial exchanges to accomplish better returns for everybody. Web-based exchanging administrations have effectively upset the way individuals purchase, what's more, sell stocks. The monetary business sectors have advanced quickly into a solid and interconnected worldwide commercial center. These headways deliver new freedoms, and the information science methods offer many benefits, yet they likewise represent an entire arrangement of new difficulties. In this paper, we propose a scientific categorization of computational ways to deal with securities exchange examination and forecast present a nitty-gritty writing investigation of the cutting-edge calculations and techniques that are regularly applied to stock market predictions and examine a portion of the proceeding with difficulties in this space that require more consideration and give freedoms to future turn of events and exploration. In contrast to conventional frameworks, the securities exchanges today are assembled utilizing a mix of various advances, for example, ML, master frameworks, and enormous information which speak with each other to work with more educated choices. Simultaneously, worldwide client network on the web has delivered the financial exchange helpless to client opinions, less steady because of creating the news, and inclined to vindictive assaults. This is the place where further examination can assume a significant part in making ready how financial exchanges will be dissected furthermore, made more vigorous later. A promising exploration course is to investigate different calculations to assess whether they are adequately amazing to foresee for the more extended term since business sectors act like gauging machines over the long haul having not so much commotion but rather more consistency. Half breed draws near that join factual and ML procedures will presumably end up being more helpful for the stock predictions.

References

(n.d.). Retrieved from Stock Prices Prediction Using Machine Learning and Deep Learning Techniques (: https://www.analyticsvidhya.com/blog/2018/10/predictingstockpricemachine-learningnd-deep-learning-techniques-python/

Abu-Mostafa, Y. S. (1996). Introduction to financial forecasting. *Applied Intelligence 6*, 20513.

Arévalo, R. J. (2017). A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting. *Expert Systems with Applications 81*, 177–92.

Bhardwaj, A. Y. (2015). Sentiment analysis for Indian stock market prediction using Sensex and nifty. *Procedia Computer Science 70*, 85-91.

Chen, T.-l. a.-y. (2016). An intelligent pattern recognition model for supporting investment decisions in stock market. *Information Sciences 346*, 261-74.

Dahlquist, C. K. (2006). Technical Analysis: The Complete Resource for Financial Market Technicians. FT Press.

David M. Q. Nelson, A. C. (n.d.). Retrieved from Stock Market's Price Movement Prediction With LSTM Neural Networks: https://www.researchgate.net/profile/AdrianoPereira14/publication/318329563_Stock _market%27s_price_movement_prediction_ with_LS TM_neural_networks/links/5a4d5545aca2729b7c8b3b7d/Stockmarketspricemovement-prediction-with-LSTM-neural-networks.pdf

Fama, E. F. (1995). Random walks in stock market prices. *Financial Analysts Journal 51*, 75–80.

Hiransha, M. E. (2018). NSE stock market prediction using deep-learning models. *Procedia Computer Science 132*, 1351-62.

Hu, Y. K. (2015). Application of evolutionary computation for rule discovery in stock algorithmic trading. *Applied Soft Computing 36*, 534-51.

M. Glantz, R. K. (2013). Multi-asset Risk Modeling. *Academic Press*.

MacKinlay, A. L. (1999). *A non-random walk down Wall Street.* Retrieved from http://gso.gbv.de/DB=2.1/CMD?ACT=SRCHA&SRT=YOP& IKT=1016&TRM=ppn+249613484&sourceid=fbw bibsonomy

Malkiel, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 383–417.

Schmidhuber, S. H. (1997). Long short-term memory. *Neural Comput., vol. 9, no. 8, pp*, 1735–1780.

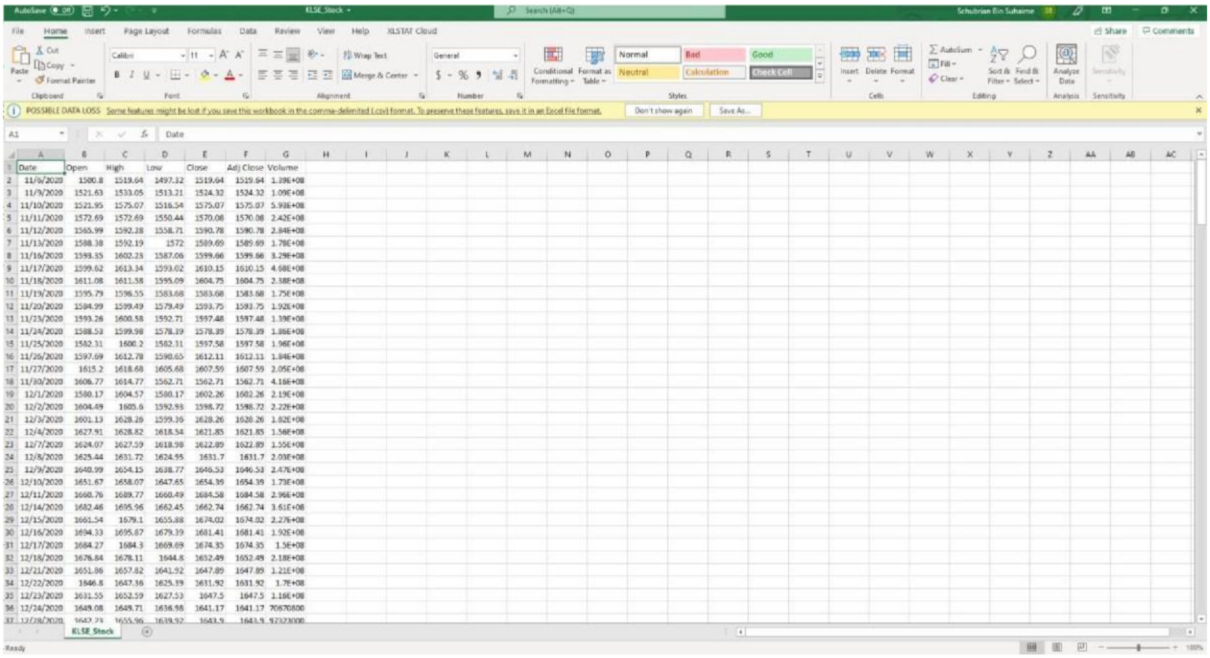
Shen, S. H. (2012). Stock Market Forecasting Using Machine Learning Algorithms. *Stanford: Department of Electrical Engineering, Stanford University*, 1-5.

Zhong, X. a. (2017). Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications 67*, 126-39.

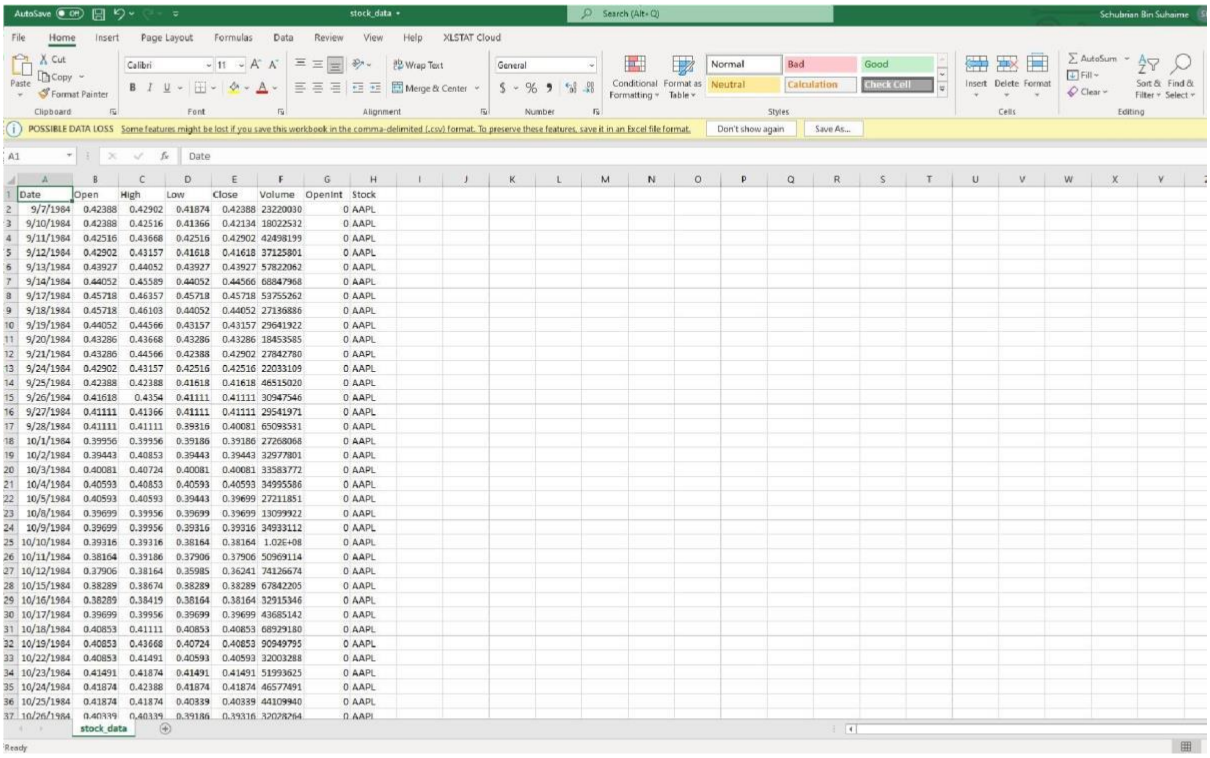
Fama, Eugene (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". Journal of Finance. 25 (2): 383–417. doi:10.2307/2325486. JSTOR 2325486

Joish Bosco (2018). "Stock Market Prediction and Efficiency Analysis using Recurrent Neural Network" <https://www.grin.com/document/419380>.

Appendix



KLSE Downloaded data



Big Companies, FB, APPL and TSLA data

```
[array([0.83347631, 0.84841441, 0.86419612, 0.8856926 , 0.90668578,
        0.89898722, 0.89302105, 0.88407894, 0.89265093, 0.89744765,
        0.8976253 , 0.90147449, 0.89106673, 0.90177051, 0.90816625,
        0.91540579, 0.91392531, 0.926983  , 0.93891569, 0.9634915 ,
        0.96061929, 0.96343222, 0.93765715, 0.87752037, 0.91339237,
        0.9174783 , 0.88843144, 0.90375448, 0.90797379, 0.91080154,
        0.91576109, 0.94394916, 0.94198019, 0.9452077 , 0.94061809,
        0.95013761, 0.94803527, 0.96488306, 0.94229103, 0.94919009,
        0.94209857, 0.92218628, 0.930699  , 0.9151985 , 0.91774487,
        0.92411097, 0.94972303, 0.95388325, 0.9435644 , 0.92610957,
        0.92776769, 0.93006232, 0.94257242, 0.95650371, 0.97289246,
        0.95564492, 0.94778371, 0.95154399, 0.94476349, 0.94251314])]
```

Array of training and test dataset for predictions

```
df = web.DataReader('^KLSE;1=9', data_source='yahoo',
                    start='2018-01-01', end='2021-10-13')
```

Data retrieve from applying pandas_reader library for web-scraping

Predicted price dashboard.

```
# Create the testing dataset, Create a new array containing scaled values from index 703 to the end of the dataset
test_data = scaled_data[training_data_len - 60:, :]
# Create the data sets x_test and y_test
x_test = []
y_test = dataset[training_data_len:, :]
for i in range(60, len(test_data)):
    x_test.append(test_data[i-60:i, 0])

# Convert the data into numpy array
x_test = np.array(x_test)

# Reshape the data bcs the data is 2D and LSTM expecting 3D output
x_test = np.reshape(x_test, (x_test.shape[0], x_test.shape[1], 1))

# Get the model's predicted price values
predictions = model.predict(x_test)
predictions = scaler.inverse_transform(predictions)

# Evaluate the model: getting the root mean squered error (RMSE)
rmse = np.sqrt(np.mean(predictions - y_test)**2)

# Plot the data
train = data[:training_data_len]
validation = data[training_data_len:]
validation['Predictions'] = predictions
```

Sample code for LSTM model in predicting the prices.