Acknowledgement

I would like to thank my supervisor, Rendon-Sanchez Juan, for encouraging and giving expert advice, always supporting, and providing sources which helped me in completing my research project. This could not have been completed without the guidance of my supervisor. I wish warmly thank you for being my supervisor and helping me out with the most important project of my master's degree.

Contents

APPENDIX

Proposal

The stock market plays a vital role in the monetary development of countries. Likewise, the development of other countries depends on the stock market. Stock market prediction is important in many ways, if the prediction is successful, it can give significant profits but if in case it doesn't it can lead to losses. A rising stock market would indicate strong economic growth for the nation but if it is not utilized properly stock market prices can fall by wrong assumptions and providing fake news outside which can also lead to economic downfall. Therefore, the process of the stock market is uncertain and unexpected. The effect of change in pricing in the stock market depends on many factors such as internal or external. Hence, forecasting the stock market prices based on the previous prices is important in finance and business. Forecasting is basically of two types: fundamental and technical which comes into account in the prediction of the stock market. Fundamental Analysis deal with the fair market value of the stock whereas technical analysis attempts to predict the future value of the stock by using useful information. Also, in the fundamental analysis, usually it looks after the difference from the real value whether it is higher or lower and in technical, they predict the starting and ending point for the traders.
As there are a lot of investors, investing a good amount in the stock market includes a high amount of risk, as said stock market prices may increase or decrease within no time. There is a situation where the prediction of the stock market is very important which is not an easy task. Prediction is based on multiple factors, but a prediction is based on the previous prices in one of them.

The aim and the objective of the project are to analyze and forecast future prices based on past prices by using different algorithms which will help managers and other top-level departments to make a firm decision and set their strategies accordingly.
For this project, we'll be considering the stock price data of BARCLAYS BANK PLC from 1st January 2018 to 1st January 2022 which is publicly available on Yahoo Finance.

Before, we discussed the investors investing in the stock market and their risks.
How their risk is high? In the stock market, every day there is some news floating that changes the price of stocks and in those some are fake. (3) Those fake news deliberately flowed into the market just to manipulate the stock prices. At those points in time, we can see the new high and low prices, and there is a high risk for the investor. Most of this fake news originated from the analysis of past behaviour of stocks. To escape from these, different algorithms came into action which will predict the future prices from the original prices. Previously, stock predictions were made based on fundamental and technical networks which had an error loss of 20% on average. New algorithms have less percentage of error which can predict almost the right amount of prices. Also, predicting day-to-day values will be much more difficult, rather than monthly. This was the one problem which make rapid change in the prices. Insider trading, lack of transparency, price manipulation, takeovers and mergers are some of the common problems which are faced by investors in making proper decisions.

Literature Overview:

In the literature, there are many techniques which are applied for the prediction of the stock market. Prediction or change in stock prices is based on multiple factors. From those, some factors can be classified into different groups based on their nature. For example, one of the variables can be financial reporting which includes the balance sheet, cash flow statement and income statements. Based on previous and current year statements and using that information we can predict the future outcomes of the stock prices. Furthermore, some of the macro-economic variables like gross domestic product (GDP), the unemployment rate and the inflation rate are also involved and considered important factors for the prediction. Not only macro-economic variables but the influence of social media like tweets, statements released from the head of companies and many more are considered under other variables which are also responsible for stock market studies.

Those are some of the overviews of the data source considered under stock market predictions. Further, machine learning techniques are trained to predict the movements of stock prices. These models give results based on historical data. Considering, that any of the models used results in 90% of accuracy in the first year doesn't mean that it will give the same in upcoming years. This is because of the change in market price, the number of employed people, and other factors affecting price movements which is why it is recommended to test the model again for confirmation and to give a solid reason to the investors. Some of the machine learning models which will be used in this project are mentioned in the below section, Methodology.

Methodology:

In our research project, our task is to predict the stock market trend. As the stock market is known for being uncertain, dynamic and nonlinear, predicting correct future trends is a challenging task as there are many micro and macro factors that depended on the change of trends or prices. Those all factors cannot be considered at one time to predict an accurate trend. But prediction is important; for that, there are some already existing methods which can be used for forecasting. From those models, we have selected the mentioned models which are suitable for future price prediction based on the past prices. Those machine learning algorithms are linear regression, polynomial regression, ARIMA (4) model and correlation between the trends. Linear Regression is known as a helpful tool for technical and quantitative analysis in the financial market which analyses two different variables and results in a single relationship. This regression can help the traders to check when the stocks are over brought or oversold and also, they can locate the price point like starting and endpoints. These are some important features of the linear regression model which we will be using in our report which will help traders and investors for making their decisions.

Hence, after applying linear regression, it will result in a straight line which will not closely relate. To solve this issue and make it more appropriate, Polynomial Regression comes into the analysis. When there is a dependent and independent variable and if we want to find the relation between them, their polynomial regression helps us out (6). It is used when there is no linear correlation between variables. It provides the best approximate relationship between dependent and independent variables. It also helps in getting a curvilinear relationship between variables. As it will result in the data set in better and if we compare linear regression with it, Polynomial will give us clear and better results. But it doesn't mean that we should skip linear regression and do only polynomial as polynomial is needed when there is no linear correlation.

Another model we will be using is the ARIMA Model, which stands for AutoRegressive Integrated Moving Average Model. This model is used for the prediction of future values based on only past values. Also, it is very well known and used to measure events of a specific period. When it comes to the analysis of data sets which is in time series like quarterly, monthly, daily, etc., the ARIMA model is known as the best analyzing time series model. Parameters of the ARIMA Model are: - AR(p) represents autoregression order, I(d) represent the degree of difference and MA(q) order of moving average. Combinedly, considering all the parameters results in time series data forecasting. While performing the model, we cannot rely on the forecasted model and make strategies and important decisions. For that, we need to check the accuracy of the predicted model, and whether it is reasonable to use or not. To check the accuracy, we will be performing another factor called MAPE; which stands for Mean Absolute Percentage Error. As its name defines the percentage of error should be less for more accuracy. If it results, in less than 10%, then our model is excellent and if it's less than 20%, it's good. If by chance it results in more than probably our model needs further improvements.

After applying those methods, we'll get to know the predicted price of the stock market. Those will be the only results based on the previous data, not considering all the factors. So, it might not be fully accurate. After getting the outcomes, we will check the accuracy through MAPE analysis.

Limitations:
As mentioned, we'll be using linear regression, polynomial regression, and Arima Model for the prediction. While running these models, we might come across some problems which can be considered a limitation of those models. Based on our stock data, some of the dataset files are too large and not properly structured which may not result in linear lines which can be difficult to analyze. Some of the limitations of the ARIMA model is not effective in the long run. If we have to forecast for 3-5 years, this model won't provide effective and accurate results. Also, it cannot be used for seasonal time series data, which is important in some conditions, for that another model called SARIMA can be used which gives results based on a seasonal basis. Furthermore, if the prediction requires a more detailed explanation of the forecast, the ARIMA model won't be the perfect model to choose as it won't result in the same. For getting more detailed and explanatory forecasting, exponential smoothing or the moving average method should be preferred.

Executive Summary:
After analyzing the outcomes, and considering all the factors, we will be providing a written summary of the report which will we be reflecting on all methods used, the importance of it and what outcomes they provided. The report will also include graphic representation (results) to prove the methods involved.

The Stock Market Prediction

Introduction:

The stock market serves as a platform for trading shares of all the listed companies. People who take part in such transactions of stocks and assets are the stock members or participants. These participants invest in different companies which are categorized under 11 sectors according to The Global Industry Classification Standard (GICS) namely energy, material, industrial, consumer discretionary, consumer staples, health care, financials, information technologies, communication services, utilities and real estate. As these companies are categorized, similarly participants are also categorized into different groups such as hedgers, speculators, arbitrageurs, and traders. Issuing stocks or going public is another way for companies or organizations to grow or raise their assets, also known as stock traders. Further, for avoiding losses, debts and interests, companies do offer shares instead of borrowing capital in cash. Also, participants or shareholders are purchasing shares for their profit. Either they borrow from the companies that pay regular dividends or by selling shares at higher prices which they purchased (Bustos, 2020). In this way, forecasting the stock market is of great interest.

Human instinct has forever been more inquisitive about what's to come in future. Determining a methodology of prediction, what is going to happen later based on what had happened and what's in present (G. González'Rivera, 2009). The stock exchange is known to be a complex versatile framework that is hard to anticipate because of the huge number of variables that decide the everyday price changes. Traders do need quick and correct information to make effective decisions. Throughout the long term, investors and specialists have been keen on creating and testing models for predicting stock prices and understanding the effects of certain price fluctuations. (Fama, 1970) (Fama, 1995). For these reasons, Machine Learning, Deep Learning, Neural Networks, Artificial Intelligence, etc is generally used in the financial sector to give a new mechanism that can assist investors in settling on better choices in both investment and management to accomplish better performance of their securities investment. Such frameworks and programming can empower investors to expect what will be happening in the organization, based on their previous behaviour and present condition to give them a head to settle on choices so they don't suffer from any huge losses.

Financial exchange value forecast is an issue that has the potential to be worth billions of dollars and is effectively explored by the biggest monetary corporations. Even after utilizing different AI methods, it's a huge issue since the solution is not clear yet.

Many different strategies have been utilized to foresee stock prices. Fundamental Analysis and Technical Analysis are two of the most common techniques to predict the stock market. Fundamental Analysis is the examination of the fair-trade value of stock and checking whether the prices are high or low while trading compared to the real value. Also, it helps in forecasting the right trade price for the securities by analysing from a macro to micro perspective. The Technical Analysis is focused on the future movements of the prices by reflecting traders when to make profits. This analysis is subjective to the chart patterns to identify the start and end point whereas in the fundamental process it prefers to be based on profit loss statement and balance sheet (Brockewell, 2000). Forecasting stock prices and return based on past data using time series models have been in application in past studies.

Some of those methods are the auto-regressive conditional heteroscedasticity model (ARCH), auto-regressive moving average model (ARMA), and autoregressive integrated moving average models (ARIMA), Linear Regression, Polynomial Regression are some of the popular techniques used in financial forecasting. Subsequently, these techniques address more advantageous and more inventive than other options, which makes them appealing to be embraced by analysts for monetary market estimating. Predicting is not the only part, after prediction we have to check the accuracy of the predicted time series data. For evaluating their performances some evaluation metrics are calculated as Mean Squared Error (MSE) or Mean Absolute Percentage Error (MAPE). For this report, we will be using MAPE metrics as it is compatible with both practical and theoretical point of view. Also, while using MAPE we can prove the minimization of the risk. Whereas MSE is only used for the comparison of the efficiency.

Context or Background:

A nation's core competitiveness is determined by its ability to manage its finances, and this sector of the economy has been growing annually (14). The stock market will play a crucial role in the capital markets framework that underpins the real economy and contribute to the core competitiveness of the nation (15). More and more institutions and people are actively participating in stock market transactions as a result of the nation's strong economic growth, strong policy backing, and progressive increase in the public's awareness of financial management. Stock price prediction has become a concern that professional analysts and investors attach enormous importance to because of the need for the related financial sectors that have also followed (Lee, 2017) (Jaike Li, 2021). While increasing the number of shareholders, the prediction of the stock market has become everyone's research topic. However, after many studies and implementation of various models for the best results, the results came unrelated. After using Artificial Intelligence (AI) in the market, results were positive but not fully accurate, hence saying, "Stock Market is Unpredictable".

All economic sectors and industries are represented by stock markets. As a result, they act as a gauge for the economic cycle as well as the desires and concerns of the people who drive prosperity and progress (Importance of Stock Market, 2020). Companies do raise their capital with public funds by issuing companies stock so that they can grow and expand their business for more profits, a part will be distributed among shareholders and the rest with companies. Further, this kind of investment leads to economic growth. Before that, stock market prediction is important for the shareholders to check and predict the prices and invest their money in the right stock. Stock Prediction matters to both, the shareholders as well as registered companies. The future prediction also affects the current business activities. As it is equally important for the business to make further strategies accordingly, for avoiding financial losses. Later on, it may lead to significant profits as well while assuming predicted prices to be true. The same applies to the market and shareholders, they might trade or buy shares more which improves the company to grow. In the end, it will automatically improve the economy. The stock market is extremely volatile. It is believed by most people that stock prices are determined by supply and demand. However, it is not strongly believed as there are many different factors.

Since there are numerous websites which can provide real-time data instantly. The vast amount of financial data that is accessible via digital media creates ideal circumstances for a data mining investigation. As there have been multiple algorithms and intelligence systems for forecasting based on complex conditions using previous data, different information providing sources like financial articles, social media news, or maybe using a hybrid system which includes multiple factors for forecasting accurately. However, having a volatile market and sharp price movements, the accuracy of using AI models is not perfect which means forecasting may take bad turns.

Research Problem:

For avoiding such turns in a volatile market, we will be researching and implementing the machine learning model with checking its accuracy with other metrics. Our research problem is concerned with the stock market pricing prediction and the models involved to make the predictions. The exploration helps a ton of new financial traders in choosing when to trade a specific stock. It additionally helps in understanding the experts and monetary news information more rapidly than doing likewise physically. Also, we will try to predict the stock market prices based on historical data and check the accuracy of the outcome.

We have also analysed the data and the flow of last 4 years.

Literature Review:

A lot of information in different organizations from many sources is managed by various monetary areas. The broad utilization of innovation has made the gigantic measure of information, known as Bid Data, effectively accessible and available. Organizations and states know about the colossal experiences that might be gathered from utilizing huge information, however, frequently miss the mark on time and assets expected to filter through its wealth of information. Organizations and states know about the tremendous experiences that might be gathered from utilizing huge information, yet frequently miss the mark on time and assets expected to filter through its extravagance of information. To get, process, communicate, and trade significant data from informational collections, different areas are utilizing Artificial Intelligence strategies (Bosco, 2018). One strategy for AI that is progressively used for large information handling is Machine Learning. A model from the finance industry will assist with the sense of how machine learning functions. Before, financial specialists, and analysts, have gone through a great deal of information from different organizations all over the planet to make effective investment determinations. Be that as it may, certain significant data probably won't be generally detailed by the media and could be known to a few individuals who enjoy the benefit of working for the organization or living in the country from which the data had been started collecting.

As mentioned earlier, we will be using the ARIMA model for forecasting and MAPE metrics for checking accuracy which will be explained further. While our research, we came across multiple machine learning algorithms and models which are previously used for forecasting and many different metrics for accuracy. We track down some of those methods and provide an overview below:

Forecasting has for quite some time been a well-known field in numerical science and finance, so there is a lot of related research in the field. Some of the common machine learning models and algorithms used by researchers are:

- Artificial Neural Networks (ANN)
- Support Vector Machine (SVM)
- Genetic Algorithms (GA)
- Long-Short Term Memory (LSTM)

An overview of the above-mentioned models and algorithms is mentioned below and specifies their accuracy rate or percentage. These are the models/algorithms which have been used by researchers for forecasting the price trend. Besides these, there are furthermore methods which are significantly used:

- Polynomial Regression (PR)
- ARIMA Model
- Fuzzy Algorithms (FA)
- Hybrid Approaches (HA)

The handling and analysis of information are accelerated by using mentioned techniques. In our research project, we will be analysing the ARIMA model and checking its accuracy. Further, predictive examination computations can plan on considerably larger informative indexes and do further analysis on a variety of aspects with small organisational alterations. Not only these but there are some common predictive models like Decision Tree: as the name suggests about the branches, it divides the dataset into branches according to similar categories which helps the reader to analyse the path (What does prediction mean in Machine learning, 2022). Linear and Logistic regression help in analysing the relationship between dependent and independent variables. Neural Network is also one of the common methods to identify or solve complex problems.
Following are some of the main techniques used by researchers for analysing/predicting the stock market and finding their accuracy:

Artificial Neural Networks (ANN):-

(Devadoss, 2013) Artificial Neural Networks are utilized for a wide assortment of a task in different fields. This model is profoundly adaptable that can plot any non-linear function as well. ANN is used for the classification and forecasting of data. They are capable of producing output with non-time series, time series and financial time series data. Some of the research studies (Yumulu, 2004) have proved ANN's method to be outperforming in the forecasting of market developments with an average failure of 3% in which 75% of the data is trained and the rest 25% is tested. The presence of the nonlinearity and unpredictability of the monetary market is propounded by numerous specialists and monetary examiners. ANNs are appropriate for issues whose arrangements require information that is challenging to determine, however, for which there are adequate data and perceptions (Abhyankar 1997). It enables the ability to gain knowledge from experience and offers a workable, practical

solution to challenges encountered in the real world. Regardless of whether the information contains corrupt data with minimal measure of standard, ANNs can sum up and precisely deduce the inconspicuous information. Without knowledge of the relationship between the factors affecting information and results, ANNs can only do nonlinear displays to a certain extent. ANNs are a more versatile and inclusive estimating demonstration device as a result. ANNs can anticipate with precision marginally more noteworthy than half, however, since financial exchange information changes so incredibly with time and nonlinearly, the expectation is troublesome even with cutting-edge strategies like ANNs (B Patel, 2014).
To conclude, Artificial Neural Networks are the models used to forecast and classify different types of data in many different fields. Also, these models have a high capacity to handle non-linear patterns with a strong in handling missing data. Although they are sensitive to boundary determination - ANNs simply give anticipated target values for obscure information with no change data to evaluate the expectation.
According to the above studies mentioned, there were multiple metrics for checking the accuracy of the model. Mean Square Error (MSE) was used most frequently and to conclude MSE was calculated by squaring the difference between obtained output and actual output and the main limitation to it is that the value will increase with the increase in stock prices (B Patel, 2014). Subsequently, MSE is an extremely deceptive number. It is utilized by the framework for preparing the weights. It cannot be utilized for examination as it varies from the stock prices. Further, it is seen that the Artificial Neural Network strategy is extremely valuable in anticipating stock lists as well as stock prices of organizations. Also, different metrics have been used to check the accuracy, and it came to know that Multilayer Perception (MLP) can provide a satisfactory output if we're using the ANN model for the stock prediction (Ince, 2008). In addition, if we want extreme accuracy, it can be combined with another model like Naïve Bayes (NB) or Genetic Algorithm (GA) and can be further checked with the metrics to prove it.

Support Vector Machine (SVM):-

Support Vector Machine is an algorithm used for predictions. It falls under the best binary classifiers (madge, 2015). It is equipped for extricating the optimal solution with a little preparation set size. Nonetheless, despite the way that SVM has extraordinary execution, its execution and classifier's speculation capacity are in many cases affected by its aspect or number of component factors. Through SVM, it can give the ideal solutions and has astounding prescient accuracy ability but this algorithm is sensitive to the selection of parameters and outliners. All time, non-time and financial data can result in appropriate outputs. Support Vector Machine Algorithms typifies the Structural Risk Minimisation principle (SRM) which has demonstrated better than the Empirical Risk Minimisation principle (ERM) which is utilized by many neural networks (Lee, 2009). The benefit of this capability is that it can deal with different information sets, as there are not many circumstances in the math of the information. Prices that as of now reflect accessible data will change just given new data, so the upcoming value heading will be reliant just on new data that shows up tomorrow. If we want to predict the next day's prices, this algorithm won't be ready to foresee as all the data had been integrated into the cost. Therefore, SVM's trouble in anticipating the following day's stock cost upholds the Efficient Market Hypothesis. The model can anticipate cost bearing for certain stocks with more accuracy prominent than 80%, however, others can't

foresee with over 30% precision. The issue is that we don't yet be aware quite a bit early which stocks the model will want to foresee precisely and which it will not, so benefitting from the model is as yet troublesome without more trial and error (Madge, 2015).
However, determining is significant as it gives substantial information to venture decisions. Undertaking the mentioned studies and research, Support Vector Machine Algorithms have a high accuracy rate. SVM has excellent speculation for the paired characterization problems. The report addresses prevalence over the introduced approach SVM given RBF which conveys broadly upgraded esteem against different classifiers. The misclassification esteem is taken out from the last SVM. It has been seen in the analysis that the trained SVM gives 98% more exact outcomes when contrasted with the another (Mohan, 2020).

Genetic Algorithm (GA):-

The purpose of the Genetic Algorithm is to cluster, classify and forecast different time series data to obtain high-level analytical solutions by relying on natural evolutions. If in case, the problem is excessively computational and to get the solution near to ideal, GA can be utilised as the strategy (Chung, 2020). GAs are normally utilized as streamlining agents that change boundaries to limit or augment some criticism measure, which can then be utilized freely or in the development of an ANN which is mentioned before. While using financial time series, Genetic Algorithms are most frequently recognised to determine the optimal values of the parameters in a trading rule and can be incorporated into ANN models to identify stock and trade. Genetic Algorithms can also work with noisy data and so are the best fitted for times series. Reasonable for exceptionally difficult issues when next to zero information on the ideal capability is given and the pursuit space is extremely huge undertaking the parameter to be selected sensitively. From all the research studies (Ecer, 2020), (Inthachot, 2016), (Nair, 2010), (Sharma, 2021), (Sable, 2017), and (Gonzalez, 2015) we came across that the Genetic Algorithm is not analysed separately. It has been analysed hybrid either with ANN for the intelligence model or with Ensemble System (ES). One of the studies (Inthachot, 2016) shows that the hybrid intelligence of ANN and GA proved the stock prediction accuracy of almost 64%, which is not bad, which increased accuracy from just ANN which was 12.4% less. However, to archive higher accuracy, we can combine it with other AI models as well.

Long-Short Term Memory (LSTM):-

One of the most often used deep learning models at the moment is Long Short Term Memory (LSTM). Additionally, it is being used for time series prediction, a challenge that is made more difficult by the presence of long-term trends, seasonal and cyclical oscillations, and random noise (Yadav, 2020). LSTM are capable of learning long-term dependencies. Provided, its ability to recognize late and early examples by giving various loads for each while failing to remember memory it considers irrelevant to gauge the results, this kind of intermittent organization has shown brilliant execution on various difficulties. Because LLTM is a Recurrent Neural Network (RNN), it cannot be trained to connect the information in huge apps. Due to this, it does not have large gap problems. Subsequently, it can deal with expanded input successions better than other repetitive brain networks that can retain short groupings (Nelson, 2017). LSTM model is also capable of examining and taking advantage of the corporation and track tracking down the information. It makes the exact forecast since it looks

after the connections and hidden information within data and also holds the information for a long period. One of the main limitations of the model is it holds the memory cell to the size of repetitive weight which misses the mechanism of the component to list the memory while composing and perusing the information.
As discussed earlier with the other hybrid models, LSTM and GA combine and make a deep learning network for stock prediction (GA-LSTM). This model can help improve the forecast of the stock market (Jia, H., 2022). LSTM model provides an accuracy between 80-85% with the training and testing, also, they checked the accuracy with the help of MSE (mean square error) and RMSE (root).

Methodology:

The top financial cooperation like JPMorgan Chase, Bank of America, Citigroup and many more, are investigating stock market predictions for a long since it can be worth billions of dollars. It is a significant issue since there is no obvious solution, regardless of the way that approximations can be made utilizing various machine learning models and algorithms. The undertaking empowers strategies for viable ML applications, like collecting and analysing a significant measure of data and using a range of techniques to prepare the program and forecast potential results. Because of the increased investment or shareholders in the stock market, its impacts on monetary provokes and its ability to anticipate its various elements utilizing different methodologies like ANN, GA, etc, has been encouraged in different sectors. While foreseeing stock market movement, the nature of the stock is a significant part to consider. Since stock exchange information is changeable, applying an algorithm of machine learning for analysis is significant. In this view, AI strategies that depend on regression are utilized to analyse stock exchange data. Below is the overview of algorithms and models which will be used for the prediction of the stock market (Linea Regression, Polynomial Regression and ARIMA model).

LINEAR REGRESSION:

Linear Regression is one of the supervised machine learning algorithms which attempts to establish a relationship between two variables and develop a learning model by fitting a linear line to input data. Regression can make connections among indicator and target values that can be interpreted as a pattern. Other datasets with the unknown objective can likewise utilize this methodology. When there is just a single indicator variable, the forecasting approach is known as simple regression and its multi regression when there are different indicator factors. We are using the regression process to address the problem of forecasting stock exchange movement because forecasting is the basic premise of regression. The input and output variables must be significantly correlated. A scatter plot can be used to determine the degree of relationship between these two variables. The correlation coefficient is a crucial numerical indicator that expresses how closely two variables are related or correlated to one another. The correlation coefficient, which ranges from -1 to 1, expresses how strongly the observed data for the two variables are related to each other (Chapter 8 ARIMA Models | Forecasting: Principles and Practice (2nd Ed), n.d.). Linear regression includes searching for the straight line that fits the info information that focuses the best. A regression line is a line that fits the information the best. Information esteems that stray from the regression line after it has been displayed for a bunch of information are referred to as outliers. The wrong information is addressed as exceptions, which may likewise highlight an ineffectively fitted regression line. With such incorrect information values, exact expectations can be extraordinarily diminished. Outliners can fundamentally change the consequences of linear regression. The foundation of linear regression is frequently the assumption that there is a direct relationship between the output and input variables, which isn't always true. The best fit for nonlinear data is not provided by linear regression. This supports the use of polynomial regression to predict stock development.

To get the linear line, it follows the formula:

$$y = a0 + a1x + \varepsilon$$

where ε is a random error with a mean equaling zero.
To sum up, linear regression is used for stock prediction to find the relationship between variables for deciding to buy or sell stocks.

POLYNOMIAL REGRESSION:

Polynomial Regression is a type of linear regression in which both the depended (x) and independent (y) variables are demonstrated as an nth degree polynomial. In some cases, plotting may suggest the relationship between the dependent and independent is non-linear. Such graphs or relationships between variables can be analysed by polynomial regression. One of the advantages of polynomial regression is that it can have multiple independent variables from which we can predict the dependent variable if those variables are having interaction between them. In addition, sometimes it can also provide meaningless outcomes if some of the principles are not undertaken. Those are:

- If the model is provided with the correct size as data, results will be more reliable.
- Large values of independent variables should not be there causing a very high degree, i.e., input values should be downsized.
- Having a non-linear nature, avoid assuming values beyond observation. (Sakhare, 2019)

ARIMA MODEL:

Auto Regressive Integrated Moving Average model explains time series data given its previous data. ARIMA model opens another approach to deal with time series prediction. These are the two most broadly utilized ways to deal with the issue, another is Exponential Smoothing, which gives corresponding ways to deal with the issue. ARIMA model describes the autocorrelation of data (What Is ARIMA Modeling ? 2022). It may be utilized for any nonseasonal series of numbers that displays designs and isn't a progression of random events. One of the key qualities is the information is gathered over a constant and standard interval. A changed version can be made to show expectations over different seasons. The ARIMA model is turning into a famous ML algorithm for data researchers to utilize for estimating values. Stock prices, for instance, reflect the distinctions between the values and real values. To understand ARIMA, in short, “AR” stands for Autoregressive which means it predicts future values based on past values. “I” means the Integrated, difference between static data and previous data to get stable data that isn’t dependent on irregularity. “MA” stands for Moving Average, dependence between observed data and residual error from moving average which

is applied to past values (What Is ARIMA Modeling? 2022). This model is also used for clustering and is compatible with time series and financial time series data but does not work well with non-linear time series data. Also, the ARIMA model functions admirably for linear time series data. The best predicting strategies for social sciences. Processing large data takes much time for processing but in the short run, it gives a more productive analysis than relative models with complex structures. Further, this model requires some parameters to be set to predict future values according to previous values, which can be mismanaged sometimes, and may not result in accurate predictions.
In addition, after getting the prediction of stock prices, in the end, it's just the data and the model. Many big companies' decisions and strategies are based on it. So, for checking that prediction, many metrics have been already established. For the ARIMA model, if we check the accuracy with MAPE metrics, which is the most common and mostly used with this model metrics. Also, it can provide an accuracy of almost 97.5% according to (Dhaduk, 2021).

Test and trial:
For this research, mainly we will be implementing the ARIMA model which is mostly used by all the traders under the stock market for future updates. We will be testing this model after the data pre-processing to get the forecast. Data-pre-processing includes the checking of the full dataset file. It includes the functioning of checking for the missing values which provide interruption during prediction, so through this process, those rows will be eliminated. Further, it also drops off the columns which are not necessary for the prediction (in this case). For this project, we're considering the stock market prices of BARCLAYS BANK (BARC.L). The data which we will be required for the prediction is publicly available on Yahoo Finance. Our dataset includes the following:

1st column (Date): the date of the trading day
2nd column (Open): the opening price of the day
3rd column (High): the highest price of the day
4th column (Low): the lowest price of the day
5th column (Close): the closing value of the day
6th column (Adj. Close): the adjustment closing of the day
7th column (Volume): the volume of shares traded on the day

We are considering the data from 1st January 2018 to 1st January 2022 for forecasting. All the prices are in currency GBP(Pounds). Mentioned above are the attributes of our data in which the opening price of the day which was highest on 24th April 2018 opened the market with the rate of £216.75 whereas the highest close rate was £217 on 19th March 2018. In this 4 years, the highest trading price of the share was £220.10, on the other hand, the lowest trading price was £73.04 which was on 19th March 2020. The reason behind that is the COVID-19 pandemic. All the stock market prices were highly dropped during the lockdown period at different times in different countries. Moreover, the highest number of shares which was traded in these 4 years was 282344614 on 5th August 2019.

Using the dataset after processing, for analysing the ARIMA model, depends on the training data and testing data. From the dataset, we will be using the ratio of 80:20 (ideally) for Training and Testing data which will help in prediction as it is based on previous data. (46) Describes, the ratio of using 80% of training data for testing will result in better prediction.

(i) Training Data: The perceptions in the training set structure, and the experience that the calculation uses to learn. Every perception comprises an output variable and at least one observed input factor.
Considering the data from 1st January 2018 to 2nd August 2019 as training data. By the specified period data, data will be trained accordingly which will predict the results later.

(ii) Testing Data: The testing data is used to assess the performance of the model utilizing some metric. While importing data for testing data, make sure not to include any data from the training dataset.
Whereas, from 5th August 2019 to 31st December 2021, will be testing data based on training data.

Ethical Implications:

For this project, we are taking the data set which is secondary data and publicly available. The data which we obtained for our studies is adequate and relevant for our research. As we are taking our dataset from Yahoo Finance, it is completely anonymous and no identical participant can be identified. Further, if the data is publicly available it is implied to use and analyse.

Our Method:

For analysing the time series method or ARIMA model, also previously mentioned, data should be stationary. As the properties of statistical models do not change and one of the properties includes the stationary data. To check, if the dataset is in stationary status, there are certain tests which can be performed:

(a) Auto-correlation function (ACF)
(b) Partial Auto-correlation function (PACF)

These will also make sure the prediction through the model will be reliable or not. If data is not constant or according to the model requirement, it may not be a correct prediction.
Also, these methods are used for defining the relationship between the current data and the past data. It also helps with the lag, which represents the correlation with itself being it is value 0 and further resulting in the correlation of 1. Being more specific and functionality:

Determining ACF and PACF:

ACF is the model which determines the correlation between the present data and past data, whereas the PACF, determines the correlation coefficient between the stationary data and lagged values.

The general form of ACF is:

$$\boxed{\frac{Covariance\ (Xt, Xt-h)}{Std.dev(Xt).Std.dev(Xt-h)} = \frac{Covariance\ (Xt, Xt-h)}{Variance(Xt)}}$$

Where Xt and Xt-h represent the value of time series at a time ‘t’ for value ‘h’.

The general form of PACF is:

$$\frac{Covariance\ (y, X3|X1, X2)}{variance(y|X1|X2)variance(X3|X1, X2)}$$

Where, X1, X2, and X3 are predictors and y is the response variable.

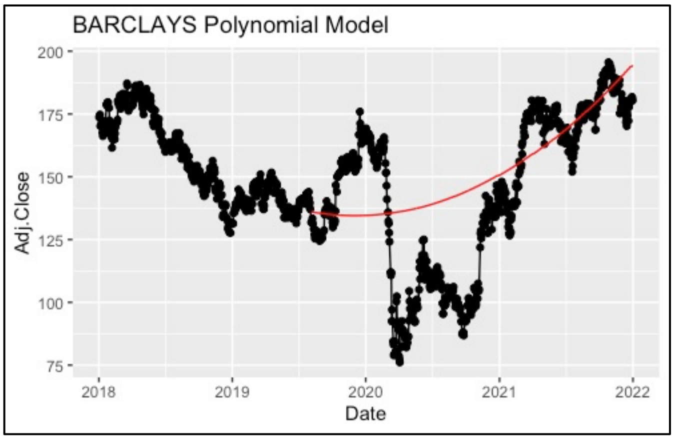
Linear Regression Analysis:

Stock exchanges and other financial business sectors consistently produce a lot of data required for analysis and forecast. They are therefore an interesting sector of how different strategies have been utilized to construct and improve these methods. Some of the commonly used methods for forecasting financial problems have advantages and disadvantages that might not be able to give favourable results (Gharehchopogh, 2013).

We used Barclays data for 4 years to predict future prices by using certain libraries installed in RStudio. We got the linear line through it but we don't the future circumstances, whether the prices will be better than in the past or not. We just made the standard assumption when applying the linear regression model, which is that it will be identical to the past. But in our scenario, the actual future data demonstrated that the near future will not be able to meet the past prices, but it would be not the case in future as we won't know which circumstances can come and hit the prices. And the so-called stock market is uncertain.

Polynomial Regression Analysis:

Polynomial Regression is also known as multiple Linear Regression. It predicts the estimated nth degree of the polynomial. When there is a non-linear relationship between the dependent and independent variable, some of the polynomial terms are added which is converted into Polynomial Regression. Similarly, we have identified the attributes from the dataset and analysed them through libraries already installed in RStudio. Below are the results of the Polynomial Regression of Barclays bank for 4 years:



We can analyse through the graph that; it makes a curve which is predicted by the training of the model.

ARIMA MODEL Analysis:

The analysis contains 3 basic methods: AR, I, and MA (p,d,q).
Auto Regression: the values of a time series are regressed on their own lagged values in auto-regression. (p) value.
Integrated: converting non-stationary data into stationary data with the difference in time series (d). If the value of d = 1, it's the difference between two times series data. If d = 2, it's the difference between the difference obtained by (d = 1) and keep going.
Moving Average: (q) values which are lagged numbers.

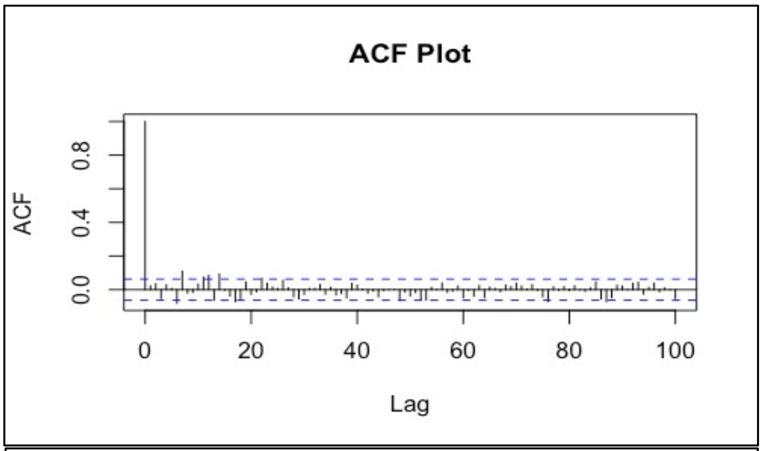
Following are the methods followed to archive our prediction:
Step 1: Ensuring data to be stationary: The Box-Jenkins method requires that a time series be stationary to be modelled. A stationary time series has no trend and a constant mean and variation throughout time, making value prediction simple. To check if the data is stationary or not we have implemented another test**: the Augmented Dickey-Fuller Test** (ADF) which means if our dataset provides us with a p-value of less than 0.05, the data is stationary. We use the differencing method to change a non-stationary process into a stationary process. Finding the differences between consecutive values in a time series is known as differencing it. A new time series dataset made up of the differenced values can be evaluated to find new correlations or other statistical features. After differencing once, if the data is still not stationary, we can try differencing method till we get our data to be stationary. It is important to have stationary data in time series analysation.

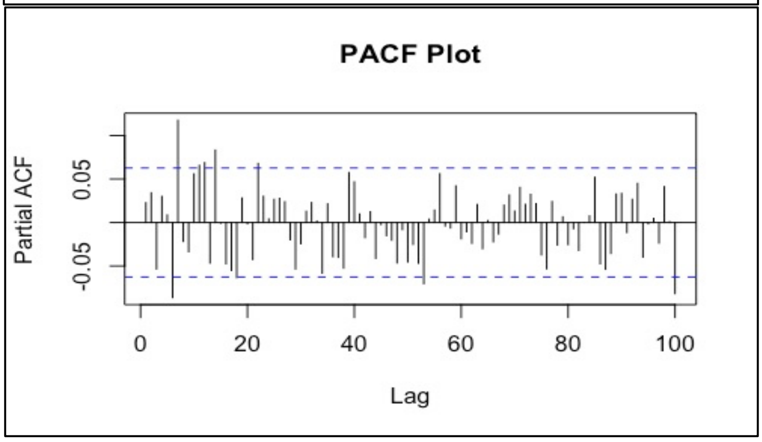
```
> # Conduct ADF test on log returns series
> print(adf.test(stock))

        Augmented Dickey-Fuller Test

data:  stock
Dickey-Fuller = -8.5003, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary
```

Step 2: Identification of variables p and q: The ACF will exponentially decay for AR models, and the PACF will be utilised to determine the order (p) of the AR model. If the PACF shows a single substantial spike at lag 1, we have an AR model of order 1. If the PACF shows substantial spikes at lags 1, 2, and 3, then the AR model is of order 3. For MA models, the PACF will exponentially dampen, and the MA process order can be determined using the ACF plot. An MA model of order 1 is present if the ACF shows a single substantial spike at lag 1. If the ACF shows substantial spikes at lags 1, 2, and 3, then the MA model is of order 3.



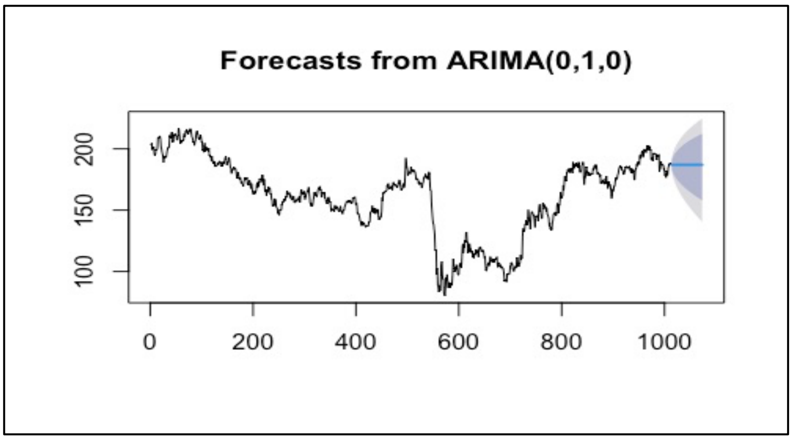


Step 3: Estimation and Forecasting:
Once the parameters (p, d, and q) have been established, we estimate the ARIMA model's accuracy using a training set of data, and then we use the fitted model to predict the values of the test set using a forecasting function. Finally, we compare our predicted values to the actual ones to see if they agree.

How closely a time series correlates with its prior values is known as autocorrelation. As is common knowledge, the ACF will exponentially dampen AR models. The correlation between the points, up to and including the lag units, is visualised using the ACF plot. Although we can see that the autocorrelation is strong for a large number of lags, the autocorrelation at later lags may be only a result of the initial lag's autocorrelation spreading.

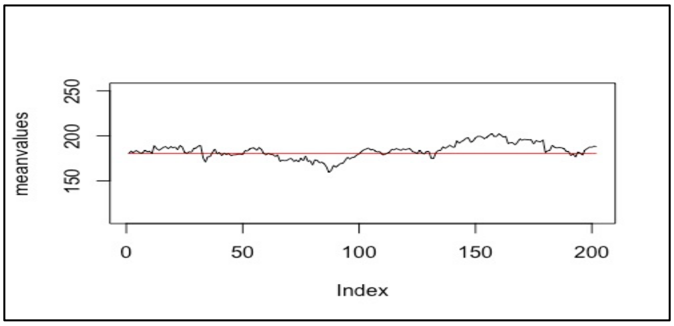
After plotting correlations, we can apply our dataset to the Arima model. Now after, the model has been fitted, we can move on to project the values of our daily close prices for the future. We concentrate on predicting the closing stock price for the upcoming 60 days or a typical month. Below are the results:

```
#Dataset forecasting  for the  next  60  days
price_forecast <- forecast(modelfit,h=60)
plot(price_forecast)
head(price_forecast$mean)
head(price_forecast$upper)
head(price_forecast$lower)
```



The blue represents the mean of the prediction. Also, it is provided with the dark and light blue coloured area which represents the confidence interval of respectively 80 and 95 per cent. As we can see from the forecast ARIMA came is (0,1,0) which is a special case in this model with a random walk.

Once our forecast has been applied to the train set, we plot the average tendency of our forecast over the recent price movement on the test set.



The red line predicts the mean forecasting over the closing price of the stock. It shows a linear approach to the future.

After forecasting, to check the accuracy of our model. For the above model considering the dataset mentioned. We got an accuracy of 58.82%. The formula we used to check the accuracy was:

```
> print(Accuracy_percentage)
[1] 58.82353
>
```

For improving the accuracy of the model, we can further increase the number of observations, try again to do the ADF test to check the stationary data and try to increase the number of independent variables. These are some tried methods to increase the accuracy.

Risk and limitations: -

Analysation and anticipation of the stock exchange stay fascinating and troublesome issues. To extract information from the data and analyse how stock costs are impacted, we deal with new issues as additional data has become accessible. Live testing issues, algorithmic trading issues, long-term forecasts, and sentiment analysis problems on organization filings are a part of the challenges. Most of the stock analysis and expectation affirms that the suggested methodologies might be applied progressively to generate significant profits in the stock exchange, which is relevant to the live test. It is a strong assertion to make since while algorithms might perform well in a controlled environment when back-tested, live testing may present various difficulties or be ineffective because of things including cost changes, immaterial news, and background noise.

Likewise, the limitation provided above, one of the real examples is the Knight Capital Tragedy1, wherein the business lost $440 million. Therefore, a valuable analysis is to grasp how a portion of the normal stock exchange approaches capability in the real world or virtual settings. (Schaefer, 2012). Not only a stock prediction but in every case of prediction future is very uncertain. Regardless of how intensive our investigation is, it must be all precise as the information we have as of now. What will happen tomorrow is unsure to us. Analysis of potential future development is led under the reason of "any remaining things being equivalent." This suggests that assuming things go on as they are, we guess that a stock will have an increment based on the trend. We cannot predict any of the contingencies. There are consistently days, weeks, months, or even years that challenge the chances, even though everything stays equivalent on certain days. Foreseeing at these times may be especially hazardous on the off chance that forecasts end up being misleading. For example, estimating an expansion in a cost while costs are declining could demolish a trader, particularly since we can never be sure of how the market will respond to new data or news that might open up. Indeed, even uplifting news probably won't be sufficient to altogether raise costs when costs are falling, and, surprisingly, awful news probably won't be sufficient to bring down costs when costs are expanding fundamentally (49). It is challenging to anticipate when a stock will rise, and it is extraordinary for a venture choice to incorporate a benefit or stop misfortune leave point. Unpractised dealers expect an expansion in their value positions, yet this isn't generally the situation, and accept they will want to exit at the top on the off chance that they are correct. Practically, an arrangement this undefined seldom succeeds. Hence, whether or not an exchange creates a benefit or a misfortune, all merchants should have an arrangement for how they will enter and leave the market.

Further, we used the ARIMA model in our report. Our analysis examines a time series with 1013 observations and considers Barclays Plc stock prices, and daily quotations for the period of 4 years from January 1, 2018, to January 1, 2022. The data was taken from Yahoo Finance.

Since the parameters (p, d, and q) should be manually settled, finding all that right fit can be a tedious experimentation methodology, even though ARIMA models can be exceptionally exact and trustworthy given the right conditions and information accessibility.

The model likewise vigorously depends on the precision of past information and information separation. For the model to deliver dependable outcomes and figures, it is vital to ensure that the information was accumulated unequivocally and over a drawn-out timeframe. An absence of automated updates. These models will generally be unstable, which makes numerous financial suspicion. In contrast to straightforward naive models, there is no programmed automated component as new information becomes accessible. All things considered, the whole methodology should be repeated, particularly the diagnostic checking stage as the model might have separated. The B-J models (Box-Jenkins) ordinarily have high costs because of their high information prerequisites, absence of effective updating processes, and the necessity to assess them utilizing nonlinear estimation. They are likewise somewhat a greater amount of art than science because of subjective input at the initial level. Both changes in perceptions and changes in model particular, the ARIMA model tends to be unsteady. Numerous specifications won't create any discoveries, and the results couldn't be completely exceptional similarly to most nonlinear assessment strategies (50). ARIMA model is suitable for a short run. If we want the result to be effective and profitable, we can use this model for the prediction of short-term analysis as it will be using high-frequency data.

After all these analyses, the alternative to the prediction can be; as we know prices do fluctuate in waves, so, all brokers should comprehend that prices do change in waves across all periods while taking a chart at any outline after getting a handle on the previously mentioned focuses. However, the trend is still up, dealers don't have to panic and leave their positions regardless of whether prices decline. If prices quit rising over the course of their period, they need to have an exit plan. To participate in one of these waves, short-term investors should be unattached to a specific course. The factual precept that costs move in waves is overlooked assuming that one makes an expectation that costs will just move in one manner. Also, the possibility that support or resistance will hold or that a break of these levels will bring about a huge breakout is a genuinely widespread misconception. Position traders now and again make expectations about what will occur. Exchanging members should comprehend that help and obstruction levels are simply huge cost zones. It is an attempt to foresee the market to make presumptions that a breakout will occur or that a level would prevent further development.

We can enter the market at crucial points in time by understanding that prices do move in waves and that we shouldn't expect that large differences will hold or break. We should do such in light of what the cost is doing, not what we expect it to do. Realizing that should make it simpler for traders to find themselves on the right side of the exchange than the wrong side.

Conclusion:

Many investors overall have communicated revenue in corporate shares. Simply deciding, is a complicated endeavour since there are countless factors influencing everything. Financial investors are anxious to forecast/predict the stock exchange's future condition to make profitable investments. Gains from even little expansions in estimate accuracy can be significant. By providing steady data like the guidance that stock prices will head down in the future, an effective prediction framework will help traders in making an investment that is more productive and efficient. Investors and dealers can execute on the financial business sectors utilizing any gadget that has a web association. Individuals have gotten more attracted to stock exchanges throughout the course of recent years. The presentation of innovation has changed the financial market, likewise, it has in every part of life. Individuals can build their investments now. The online stock exchange has just adjusted how individuals trade stocks. The monetary business sectors have formed rapidly and are associated with a worldwide market. These advancements open up additional opportunities.

Moreover, while doing the forecasting we came across some issues which are concluded further. It is important to have the correct size of the dataset for the prediction because some of the datasets may not be stationary and if use non-stationary data or the value of p is more than 0.05; the forecasting model violates some of the assumptions and the result may not be accurate. ACF and PACF are the tests of correlation between the dependent and independent variables. An important factor while using the ARIMA model is to predict the values of p, q, and d and test the model with a metric suitable for the model.

EXECUTIVE SUMMARY

Forecasting the movement of the stock prices and the market becomes more crucial to avoid significant losses and make informed decisions because the stock market is inextricably linked to a nation's economic growth, attracts significant investments from investors, and issues equities to the public interest. Investors and dealers can transact in the financial markets using any device that has access to an internet connection. People have gotten more drawn to stock trading over the past few years. The introduction of technology has altered the stock market, just like it has in every other aspect of life. People can increase their investments now. Online stock trading has only altered how people buy and sell stocks. The budgetary markets have developed quickly and are connected to a global market. These developments open up new possibilities. As time passes by, more and more people are getting into stock trading and making it their business which the prediction has become important to most people. People who don't want to trade, then to they are interested in gaining what's coming next. To get the right predictions, researchers have introduced some machine learning models which will help to predict the stock market prices. Some of the models are ARIMA which we used in this project, Support Vector Machine, Artificial Neural Network, and Long Short Term Memory, which are mostly used to predict the trend based on their historical data. These models only predict from the way they are trained. In this report, we were concerned about the prediction and the process of prediction. As not only the past data but other factors like company performance, industry trends, and unexpected events can also change the movement of price and can further prove the prediction wrong. In this case, accuracy will also not help due to very uncertain factors. In this report, we focused mainly on the ARIMA Model by identifying their main factor p, d, and q which represents the autoregressive terms, nonseasonal differences and stationary of data respectively by which the model is trained for prediction. After getting the prediction we also checked the accuracy which came approx. 60% considering the data for 4 years.

We identified during our research, our ARIMA model came (0,1,0) which is a special case in which Y axis series might not be stationary after analysis. To fix this, we might have to take more data, do an ADF test, and check again for another ARIMA model.

This prediction and the process can be used in different sectors that want to make decisions and strategies by predicting the future. It can be useful to get at least get an idea about their pricing and likewise they can improve their methods to overcome it. The most organization may be concerned about the accuracy of the model, but it depends on the attribute we choose and how we train our model. If this model can give an accuracy of 58.82%, then after careful training and testing it can also give results into a good percentage. Also, if we improve the length of the training data, it will result in better accuracy it will improve the prediction and can give efficient results which are necessary for every sector. Future studies can concentrate on merging data from sentiment analysis of stocks with numerical values relating to historical stock values to forecast stock prices. By utilizing both information, more efficient stock recommendation algorithms may also be created. Further use of deep learning-based methodologies can lead to more effective feature extraction methods. Based on the models, visualization data is also important which has been shown throughout the project and can be used as a predictor in short-term data. The machines learning models or algorithms have so far not been very helpful in predicting the course of a company's worth over a longer period. In this research, we describe a machine learning-aided method for assessing the long-term future price of a share. There are some limitations to the models which are used in the project.

Some of the major limitations to it are the data should be stationary. Most of the time data provided for the analysis may have some blank spaces or NA values by which, predictions will not be accurate. For solving or using any of the models, data has to be stationary which will be done by the ADF test. It can be done for multiple periods till the data comes in a stationary position. Further, for this report we have considered the prediction of 60 days, we can easily change the value of (h) in the code and can run through it again to predict the mentioned period.
To conclude, our report can be used as a way to prove the prediction of stock market prices and the process to get through with the different methods. Also, alternative methods can be used which have been explained with their properties, limitations, and accuracy. This report can help small traders, investors, businesses, and other sector companies to predict future prices and make their strategies to improve their near future.

References

PROPOSAL:

1) AS, S. (2013). A study of fundamental and technical analysis. International Journal of Marketing, Financial Services & Marketing Research, 2(5), 44-59.
2) Clarke, J., Chen, H., Du, D., & Hu, Y. J. (2020). Fake news, investor attention, and market reaction. *Information Systems Research*, *32*(1), 35-52
3) Khot, V. (2022, September 26). Get a glimpse of the future using time series forecasting using Auto ARIMA and Artificial Intelligence. Medium. Retrieved 29 September 2022
4) Dikov, D. F. (2021, December 30). Forecasts with the Polynomial Regression Model in Excel. Medium. Retrieved 29 September 2022

FINAL REPORT REFERENCES

1) Bustos, O. and Pomares-Quimbaya, A., 2020. Stock market movement forecast: A Systematic Review. Expert Systems with Applications, 156, p.113464.
2) G. González'Rivera and T. H. Lee, "Nonlinear time series in financial forecasting" in Encyclopedia of Complexity and Systems Science, New York, NY, USA:Springer, 2009.Fama, Eugene F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance 25: 383–417.
3) Fama, Eugene F. 1995. Random walks in stock market prices. Financial Analysts Journal 51: 75–80.
4) Brockwell, P. J., & Davis, R. A. (Eds.). (2002). Introduction to time series and forecasting. New York, NY: Springer New York
5) Lee, CL., Kuo, SC. & Lin, CJ. An efficient forecasting model based on an improved fuzzy time series and a modified group search optimizer. Appl Intell 46, 641–651 (2017).
6) Jiake Li, "Research on Market Stock Index Prediction Based on Network Security and Deep Learning", Security and Communication Networks, vol. 2021, Article ID 5522375, 8 pages, 2021
7) The Importance of Stock Markets. (2020, July 27). Retrieved 30 September 2022.
8) Bosco, J. (2018, April 6). GRIN – Stock Market Prediction and Efficiency Analysis using Recurrent Neural Network. Retrieved 30 September 2022
9) What Does Prediction Mean in Machine Learning? (n.d.). Retrieved 30 September 2022
10) Devadoss, A. V., & Ligori, T. A. A. (2013). Stock prediction using artificial neural networks. *International Journal of Data Mining Techniques and Applications*, *2*(1), 283-291.
11) Yumulu, S. Gurgen, F, Okay N (2004). "A comparison of global, recurrent and smoothedpiecewise neural models for Istanbul stock exchange prediction", Pattern recognition letters 26(13), 2093-2103.
12) Abhyankar, A., Copeland, L.S., and Wong, W. (1997). Uncovering nonlinear structure in real-time stock market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. Journal of Business & Economic Statistics, no 15, p.1-14.
13) S. R. Y. Mayankkumar B Patel, "Stock prediction using artificial neural network," International Journal of Innovative Research in Science, Engineering, and Technology, 2014.
14) Ince, H., & Trafalis, T. B. (2008). Short-term forecasting with support vector machines and application to stock price prediction. *International Journal of General Systems*, *37*(6), 677-687.
15) Madge, S., & Bhatt, S. (2015). Predicting stock price direction using support vector machines. *Independent work report spring*, *45*.
16) Lee, M. C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications*, *36*(8), 10896-10904.

17) Mohan, L., Pant, J., Suyal, P., & Kumar, A. (2020, September). Support vector machine accuracy improvement with classification. In *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 477-481). IEEE.
18) Chung, H., & Shin, K. S. (2020). Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. *Neural Computing and Applications*, *32*(12), 7897-7914.
19) Ecer, F., Ardabili, S., Band, S. S., & Mosavi, A. (2020). Training multilayer perceptron with genetic algorithms and particle swarm optimization for modeling stock price index prediction. *Entropy*, *22*(11), 1239.
20) Inthachot, M., Boonjing, V., & Intakosum, S. (2016). Artificial neural network and genetic algorithm hybrid intelligence for predicting Thai stock price index trend. *Computational intelligence and neuroscience*, *2016*.
21) Nair, B. B., Mohandas, V. P., & Sakthivel, N. R. (2010). A genetic algorithm optimized decision tree-SVM based stock market trend prediction system. *International journal on computer science and engineering*, *2*(9), 2981-2988.
22) Sharma, D. K., Hota, H. S., Brown, K., & Handa, R. (2021). Integration of genetic algorithm with artificial neural network for stock market forecasting. *International Journal of System Assurance Engineering and Management*, 1-14.
23) Sable, S., Porwal, A., & Singh, U. (2017, April). Stock price prediction using genetic algorithms and evolution strategies. In *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)* (Vol. 2, pp. 549-553). IEEE.
24) Gonzalez, R. T., Padilha, C. A., & Barone, D. A. C. (2015, May). Ensemble system based on genetic algorithm for stock market forecasting. In *2015 IEEE congress on evolutionary computation (CEC)* (pp. 3102-3108). IEEE.
25) Yadav, A., Jha, C. K., & Sharan, A. (2020). Optimizing LSTM for time series prediction in Indian stock market. *Procedia Computer Science*, *167*, 2091-2100.
26) Nelson, D. M., Pereira, A. C., & De Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In *2017 International joint conference on neural networks (IJCNN)* (pp. 1419-1426). Ieee.
27) Jia, H. (2022). Based on the LSTM-GA Stock Price Ups and Downs Forecast Model. In: Sugumaran, V., Sreedevi, A.G., Xu, Z. (eds) Application of Intelligent Systems in Multi-modal Information Analytics. ICMMIA 2022. Lecture Notes on Data Engineering and Communications Technologies, vol 136. Springer, Cham.
28) Sakhare, N. N., & Imambi, S. S. (2019). Performance analysis of regression based machine learning techniques for prediction of stock market movement. *International Journal of Recent Technology and Engineering*, *7*(6), 655-662.
29) Chapter 8 ARIMA models | Forecasting: Principles and Practice (2nd ed). (n.d.). Retrieved 30 September 2022, from https://otexts.com/fpp2/arima.html
30) *What Is ARIMA Modeling?* (2022, August 4). CORP-MIDS1 (MDS). Retrieved 30 September 2022, from https://www.mastersindatascience.org/learning/statistics-data-science/what-is-arima-modeling/#:%7E:text=ARIMA%20is%20an%20acronym%20for,future%20data%20in%20a%20series
31) Dhaduk, H. (2021, July 18). *Stock market forecasting using Time Series analysis With ARIMA model*. Analytics Vidhya. Retrieved 30 September 2022, from https://www.analyticsvidhya.com/blog/2021/07/stock-market-forecasting-using-time-series-analysis-with-arima-model/

32) Gharehchopogh, F. S., Bonab, T. H., & Khaze, S. R. (2013). A linear regression approach to prediction of stock market trading volume: a case study. *International Journal of Managing Value and Supply Chains*, *4*(3), 25.

33) Schaefer, S. (2012, August 2). *Knight Capital Trading Disaster Carries $440 Million Price Tag*. Forbes. Retrieved 30 September 2022, from https://www.forbes.com/sites/steveschaefer/2012/08/02/knight-capital-trading-disaster-carries-440-million-price-tag/?sh=6ed546553f0a