UNIVERSITY OF NAIROBI

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES
SCHOOL OF COMPUTING AND INFORMATICS

PREDICTION OF STOCK PRICES USING PREDICTIVE DATA ANALYTICS.
(A CASE OF NAIROBI SECURITIES EXCHANGE)

BY ELIAKIM PETER OLUOCH OBUOLO
P53/66046/2013

SUPERVISOR
DR. ELISHA ODIRA ABADE.

Research Project Report Submitted in Partial fulfillment for the requirement for the award of the Degree of Master of Science in Distributed Computing Technology

AUGUST 2020

Declaration

I Peter Oluoch do hereby declare that this research project is entirely my own work and where there are work or contributions of other individuals, it has been referenced as acknowledgement.

To the best of my knowledge, similar research has not been carried out before or previously presented to any other University.

Sign: .... Date........08/09/2020...........

Eliakim Peter Oluoch Obuolo

This project has been submitted as part fulfillment of the requirements for the Master of Science in Distributed Computing Technology of the University of Nairobi with my approval as the University Supervisor.

Sign: .... Date........03/09/2020........

Dr. Elisha Odira Abade

School of Computing and Informatics, University of Nairobi

Dedication

I dedicate this research to Nairobi Securities Exchange, My Stocks Kenya, Investing.com among other numerous websites whose stock market analysis, insights, trends, opinion pieces challenged my knowledge on stocks trading. I want to appreciate the stock gurus who graced the trading floor before me and whose knowledge and insights I will forever cherish. The stock market gurus and stock traders include: Warrant Buffet, Alexander Elder, Jesse Livermore, William O'Neal among others.

Acknowledgement

First and foremost, I would like to give thanks to the Almighty God, for the gift of life, good health, knowledge, wisdom and understanding and for the strength to carry on, when I was almost giving up. The journey has been long and tiresome. I can attest that persistence pays. The battle is not for the strongest, but for the persistent.

I want to thank my family, the beautiful girls that decorate my world. The Girls: Molly, Abby and Kayla. My two beautiful clever girls make my world absolutely heavenly. They bring joy to my heart and gives me the strength to wake up every morning. When giving up was the best option left, they reminded me that I had to complete the task. I had to leave a legacy, set a good precedent and a good pace for them to emulate in the years to come.

My appreciation to my first Supervisor, Prof. William Okello-Odongo, the humble computing giant. A Wiseman. The computing guru that has nurtured thousands of Computer Science professionals over the last three decades. Death took you away too early when I was still doing my project proposal. I was dumbfounded. Shell-shocked. For months on end, I was lost for words as my world came tumbling down. To my second Supervisor, Dr. Elisha Abade, I appreciate your guidance, to explore the various aspects of predicting stock price movement and volatility. I was focused on one way of doing things and you opened my eyes to other ways of looking at the problem in question. When I was set on getting five years' data, I was back in the field to get 10 years data, to be able to get a good view of the prediction of the stock markets. Invaluable input, I must say.

My Parents Seth Obuolo Aduol (RIP), Mama Pilista Aoko Obuolo - Nyager, my siblings Jenifer Agola, Christine Akoth, Pamela Adhiambo, Lilian Akinyi, Scholastica Anyango and my siblings from the First Family, Selina Wanemba, Jenifer Owuor, Pamela Aketch, Joseph Makedi (RIP), Stephen Oluanda, Chrispine Omondi, your financial support and moral support over the years is forever cemented in my heart.

On the spiritual and marital front, my Best man and my Pastor: Pastor Geoffrey On'gondo and your wife Pastor Jane On'gondo, who kept pushing me to finish the Master's degree program, you are forever in my prayers.

I can't forget my colleagues at work, who were helpful in several ways in helping complete this massive milestone.

Abstract

The field of stock trading worldwide is not even, has never been even and will not be even in the years to come. The money bags, the big boys continue to benefit and make massive profits from the stock market and the small fish continue to lose millions and billions of money in the market.

Thanks to technology and the fourth industrial revolution, thanks to big data, predictive analytics and more so Holt Winters Exponential Smoothening, it is now very possible to predict the volatility at the stock market and to make significant gains like the big boys. With the predictive analytics, the massive stocks data from the stock market becomes a goldmine that generates leads for short term investors who are able to be warned in advance of the possible capital erosion due to their bad investment stakes at the market and be warned of the possibility of capital gains before investing in the stock market. The data used in this project is a ten years' data and hence the predictions is 90% accurate. While it is impossible to predict with 100% accuracy what will happen in the future since stock volatility is affected by many other economic, political, environmental factors, it is very much possible to use historical data from the Nairobi Securities Exchange, do stocks prediction and predict with accuracy the future stocks movement if all other factors that affect stock volatility are held constant. The constant, is one of the major assumptions in this project.

Unlike in the olden times where stock traders and stock market analysts would rely on fundamentals of a given company to decide whether to buy or sell, or rely on technical analysis by following graphs and trends, predicting the stock market is getting easier using predictive analytics. In the past, predicting the stock market volatility has been done by several technologies including Artificial Neural Networks, Decision Trees, Logic Regression, Ensemble Models, Bayesian Models among many other statistical models. This project looks specifically at Holt Winters Exponential Smoothing, a branch of supervised machine learning model.

TABLE OF CONTENTS

List of Figures

Glossary of Stocks Trading Terms

IPO— Initial Public Offering - This is when a private company offers its shares to the public.

ROI—Return on Investment. The amount of money an investor expects after investing a given sum of money into a given investment.

P/E ratio - Price to Earnings Ration

Closing Price: the last traded price.

High and Low Price : The maximum price and the minimum price a stock traded.

Market Capitalization : A company's currency valued based on the price of the stock

Market Value : the price of the stock at the stock exchange

Opening Value: the first price the stock traded for the trading day.

Publicly Listed Companies: These are companies whose shares are traded at the exchange.

Investing: Putting money in a given business venture expecting a return after a while.

Trading : Buying and selling stocks over a short period of time like a week or a month, with an aim of making profits.

Speculating : Buying shares expecting to resell them after a short while. This term is used interchangeably with trading.

 Betting : This is a debate between right or wrong. The right person wins some money and the wrong person losses the money. In betting, there is winner and a loser. The outcome is unpredictable.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

NAIROBI SECURITIES EXCHANGE, NSE GENERATES MASSIVE DATA

Stocks trading at the Nairobi Securities Exchange, NSE generate large amounts of data on a daily basis.

Big, unstructured and diverse data gets generated every passing second from the stock market as traders buy and sell shares on a daily basis in a bid to make profit and get a return on their investments and hence it requires a paradigm shift to handle both the stock trading complexities and also guide future traders in picking the right stocks and proper stock market analysis. Such predictions and analysis are essential in helping traders to avoid losses in the market or what we popularly call capital erosion. (Singh et al 2014)

For information technologist, the future looks promising especially with big data analytics as future forecasting will be possible as people handle the core challenges of big data.

As traders buy and sell shares on the stock market on a daily basis, the market becomes a place where data grows every second. The complexities of the stock market data hence necessitate big data architecture which helps in data organization, analysis and possible predicting stock price direction. Without such data architecture design and analysis of the market data, most traders will not benefit from the data analytics by making profits and hence loss making will be the norm.

Correctly predicting stock prices is not a walk in the park, it is a big problem as observed by Kaushik and Singhal (2009) and also other scholars in the stocks market prediction domain Kannan et al (2010). They agree on the fact that minting data from the stock market is a big challenge because of volatility and ever changing and increasing data. The India shares market which was the study area for Kaushil and Singhal (2009) is volatile as prices move up and down abruptly and could be challenging to new beginner traders who lose their investment capital.

MANY INVESTORS MAKE LOSSES

Stock market volatility requires expertise knowledge to succeed. Many Kenyan investors do not have that knowledge yet and hence making decisions based on mob psychology. Few people, if any are willing to share the success knowledge they have (Dunne, 2015).

 According to O'Neil (2009) many investors have made serious losses in the stocks market. He explains that it takes a lot of trial and error before you can nail down substantial gains in stocks.

Many authors acknowledge that trading in the stocks markets is a challenging task. According to Bonde and Khaled (2011), prediction of stocks prices has always been a challenging task. They explain further that the stock price of any given company does not necessarily depend on the economic situation of the country. It is neither linked with the economic development of any given country nor a given area. Indeed stocks price prediction has become more difficult, than before.

Back home, Kenyans have made many wrong financial decisions and many have lost their investments. Kenyans have lost money both in stocks market and in property market. Majority of these investors make haste decisions not backed by research. For example, many have lost homes, case in point Syokimau home demolitions. (Standard Media, July, 2013).

Investors who own such companies like Athi River Mining Company, ARM, have registered a considerable loss with the share price dropping from as high as 100/- per share to the current 5/- per share (Daily Nation, July 11, 2018). Investors holding Mumias Sugar Company stock bought in a second offer in 2006 at Sh49.50 have been the worst hit in the past few years by Mumias' share price erosion, counting capital losses running into billions of shillings (Daily Nation, July 11, 2018).

Other stocks that have experienced capital erosion and big losses include Britam, Home Africa among others (Daily Nation, April, 30, 2015). The inherent nature of the human greed and lack of research also led to many Kenyans losing money in the various ponzi schemes in the recent past. Such ponzi schemes include DECI, Clip, among others. It goes without saying that data analysis and research will be key for many investors to avoid losing their hard earned capital.

Our ability to process what we know aids our decision making process. Whether we learn by example or create new approaches, all these aids us in human reasoning and the decision-making process (Singh, Dimri and Rastogi, 2014).

In Kenya, thousands of people have suffered financial losses at the NSE in the name of "investing" money without taking into account data analysis. These financial losses are largely due to the fact that the investors made their investment decisions based on gut feelings that a certain counter would do well. Unfortunately, things didn't turn out as they expected and they burnt their fingers in the process.

With proper data analytics and stock prediction tools, wise financial decisions can be made that can help investors make good profit and good return on their investments. The financial losses we have witnessed before like Nation Media Group selling at Kshs. 320 during the 2010 period and now selling at Kshs. 58 some 9 years down line and hence investors have suffered a whopping 81% loss. Many other counters in the exchange have led investors to wonder what exactly could be the solution to their financial hemorrhage.

1.2 PROBLEM DEFINITION

Investors at the Nairobi Securities Exchange continue to make losses on their investment mainly due to four things; Wrong timing of the market says Metcalfe G (2018), Lack of big data analysis of the market prices observes Gutierrez, Anzelde, and Gobenceaux (2013), Lack of information on the company fundamentals and technical analysis notes Barber and Odean (2011), Mob psychology drives thousands of people to invest during IPOs says Pring (2018) , Hansen (2017) and Fenzl and Brudermann et al (2013).

It is the aim of this project to solve the problem of data analysis of market prices using Holt Winters Smoothening stock prediction technology. This technology has been studied by among others Bakar and Rosbi (2016), Awajan, Ismael and Wadi (2018), Rusu and Rusu (2003), Ortiz (2015) , Newberne (2006), Kannaiah and Chakrapani (2012) among others.

With this prediction technology, it is my hope that making the right financial decisions will be easier for equity investors at the stock market and that many will realize return on their investments, ROI.

Holt Winters Smoothening exponential stock prediction is a supervised machine learning model which has been studied by many scholars including Dzikevičius and Saranda, (2010), Bakar and Rosbi (2016) who studied reliability of exponential smoothening method for forecasting Islamic share price in Malaysian Stock Exchange. From the University of Jordan, Awajan, Ismael and Wadi, (2018), tried improving the accuracy of stock market data using empirical mode decomposition Holt Winter, EMD-HW bagging. Other stock prediction technologies that are on the radar of researchers for stock prediction purposes include: logistic regression, ensemble models, Bayesian models, decision trees and artificial neural networks. Many researchers are working on solving this problem of investors losing money at the stock exchange worldwide.

1.3 OBJECTIVES

1.3.0 BROAD / GENERAL OBJECTIVE

The general objective of this project is to explore the possibilities of using Holt Winters Exponential Smoothening stock prediction technology to help investors make accurate stocks prediction based on data analytics of Nairobi Securities Exchange historical data.

1.3.1 SPECIFIC OBJECTIVES

i.   Investigate how data analytics can be applied to historical prices and develop predictive patterns. (Perform big data analysis on the stocks data and find predictive patterns).

ii.   Develop stocks investment aid software to help investors in their investment endeavours.

heck accuracy of the of the tool developed in this project as regards stock price predictions and other tools developed in previous research studies on machine learning like Artificial Neural Networks and Bayesian Network, which are supervised machine learning models.

1.4 SCOPE

The past 5 years' stock market data of the companies listed in the Nairobi Securities Exchange will be used in big data analysis and the results will help advice investing public based on stock price movements and the best counters to invest in.

1.4.1 CHALLENGES

One of the major challenges encountered during the study is the acquisition of data (daily stock trading prices) from the Nairobi Securities exchange. The NSE is selling the daily stock prices at exorbitant rates. This setback led to the acquisition of data from other sources outside Kenya and more specifically from online sources in the United States of America.

1.5   RESEARCH QUESTIONS

a. Can data analytics of NSE stock market data be used to make stocks predictions?

b. Can Holt Winters smoothing prediction model help investors at the Nairobi Securities Exchange make good investment decisions and get a good return on their investments?

c. How accurate are the Holt Winters smoothening predictions of NSE historical stocks prices compared to daily stock price movements?

d. Based to previous research studies like ANNS, how effective is Holt Winters Smoothening prediction model in predicting daily stock prices at the stock market, the NSE?

1.6   LIMITATIONS OF THE STUDY

Trading in the stock exchange successfully requires lots of wits, experience and occasionally people make losses.

The system will use historical stock market data to predict stock prices based on data analytics. While this may work, the fact that a given company has done well in the past is not a guarantee that it will continue doing well in the future. Things happen. Disasters happen. Economic turmoil happens. Stock Markets crush. (Fenzl and Brudermann et al, 2013).

So the analysis may not be the gospel truth on investment. The system may not be the magic bullet on stocks market investments as unforeseen events can drastically change the economy and hence stock market prices like the global financial meltdown in 2008-2009.

The other limitation in this study is the cost of getting data from the Nairobi Securities Exchange, NSE. The stock price data is not free, the data is sold. Several years data could be obtained, but unfortunately, cost was a factor. Nairobi Securities Exchange has stored trading data for the last 20 years. I couldn't even manage to buy 5 years data for this study. With more finances, more data could have been bought and that could have perhaps improved the accuracy of the system under this project research study.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

In the stock market, investors operate in two spheres of human reasoning: hope or fear. Some investors hope that the stock they have bought will increase in price and they will sell and make profits, while others fear that the stock will drop in price and they'll incur losses.

According to O'Neil, (2009), only two emotions rule the stocks market, fear and hope. He noticed that the only problem that exists is that we hope when we should fear and we fear when we should hope.

For many people, making decisions is based on relying on the information at hand. On the other hand, reasoning is based on processing the information we already know. This new kind of reasoning is used by market players; stock traders, stock investors, fund managers to uncover the underlying data patterns and infer multiple meanings thereof. These meanings, it is expected to result in positive financial reward (Singh, Dimri and Rastogi, 2014).

As we enter the new millennium, and with the increasing developments in technology, it goes without saying that research and innovations will be a key ingredient for success in organizations. Success for many organizations will be achieved through innovation which will help them not only to gain a competitive advantage over their competition, but also to create products and services

which meet the client's needs and expectations. This will be significant if they need to succeed in the ever challenging business environment.

How to make money by investing in the equities market remains one of the most important decisions individuals make in their investment lives. Investors have to analyse the efficiency of the corporations under study. They as well have to analyse other factors like assets, debts, cash-flow of the company. Stock investments, forex investment, property market, business among other options are some of the investment tools for personal wealth. Thousands of people have made a fortune trading in the stock market. Some of the notable investors and kings of the stock market include: Warren Buffet, Alexander Elder, William O'Neil, among others. Those investing in the stock market are always keen in knowing the best stocks that will help them make profits and not looses. This is complex process which requires wits, experience and data analysis. Generally speaking, there are two methods which are commonly used in analysing stocks; they are ; fundamental analysis which looks at the company's financials like cash flow, management and business climate and technical analysis which simply observe the charts which monitor stock price movement.

Few investors however, adhere to fundamental analysis and technical analysis, majority of the investors depend on crowd psychology or broker recommendations.

2.2 BACKGROUND OF THE STUDY

While a few Kenyans have made significant fortunes in the Nairobi Stock Exchange based on wise financial decisions, many others haven't been lucky. It is on this background of unlucky investors making losses at the Kenyan stock market, the NSE, that the project intends to find a computing solution that will help investors make gains.

In the global arena, many stock investors have been unlucky. Only few experienced traders make profits. Some unlucky investors include those who invested in oil. The market price for West Texas

Intermediate Crude Oil   reached a record high of over US$ 140 in summer 2008 and by 2009, it was trading below $40 (Fenzl and Brudermann et al, 2013).

Some of the lucky investors include those who invested in companies like Equity Bank, Housing Finance, British American Investment (Britam), Centum Investment Company among others as at 2015.

While those who have suffered financial losses include those who invested in Kenya Airways, Mumias Sugar Company , Sameer Africa Limited, Uchumi Supermarkets, Trancentury Limited among others.

Based on historical records of how Kenyans have suffered financial looses in the name of investing in the Nairobi Stock Exchange, this project comes to offer a solution to the tough investment decisions being made on a daily basis. It will do big data analytics of the stocks market data to help predict stock prices and help investors make good investment decisions in the stocks market.

2.3   HOLT WINTERS EXPONENTIAL SMOOTHENING

Holt Winters exponential smoothening model of prediction has been used in various sectors in order to help in prediction and forecasting of events. In the software development and testing field for example, Yakovyna and Bachkai (2018) has been able to use Holt Winters exponential model in predicting software failures and reliability. They argue that Holt Winters model is simple, works well in practice and suitable for short term forecasts time series data.
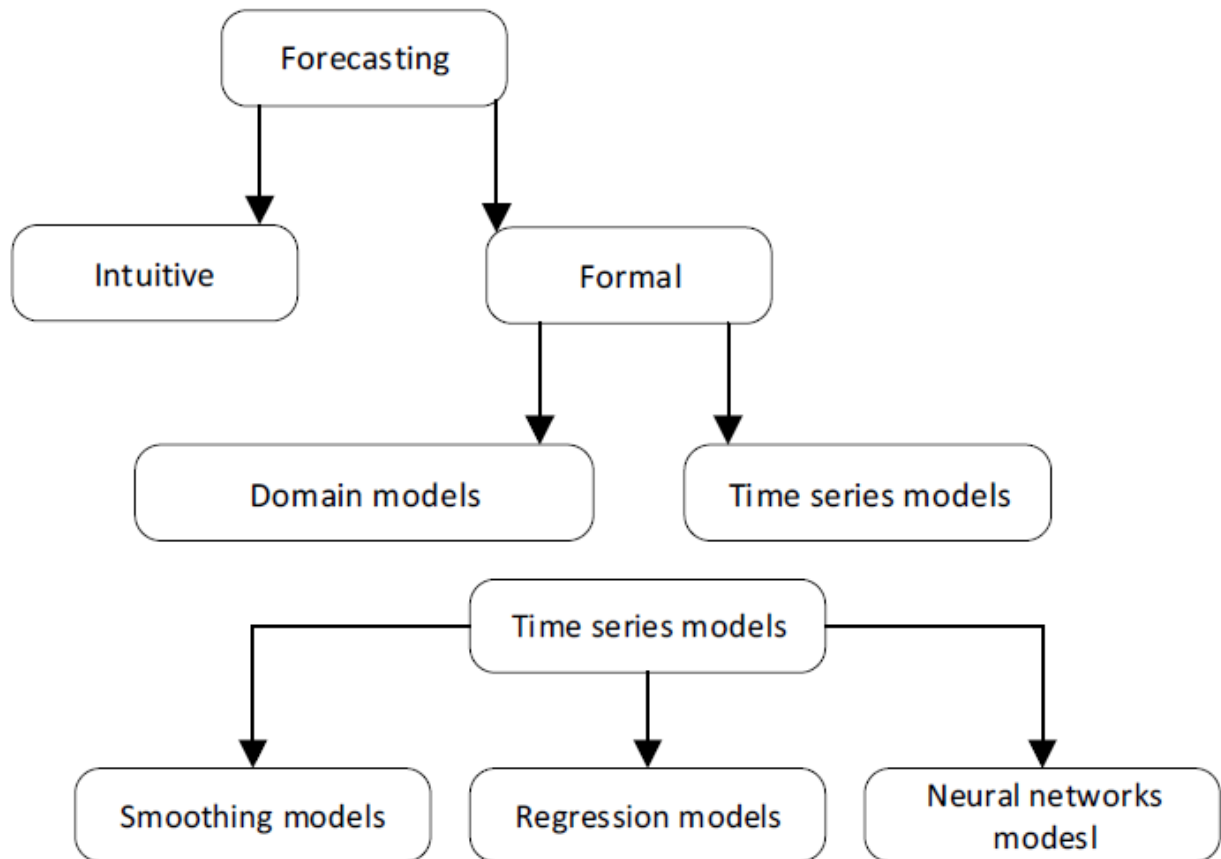
Fig. 1. Holt Winters forecasting is a model under smoothing models.

Holt Winters smoothening model for forecasting has been applied in healthcare sector. Newberne (2006) observes that health care managers often encounter the need for accurate and reliable forecast as regards decisions about staffing, purchases and healthcare delivery. Holt Winters model has been observed to save valuable time and resources. He further opines that Health Care Managers can use forecasting methods to estimate future health plan enrollment, outpatient visits, laboratory tests, radiology images and pharmaceutical stocks. With such data becoming available, health care managers can turn such data into information.

Investors in the forex market for example have been keen to use Holt Winters Smoothening model to forecast exchange rates. Ortiz (2015) observes that accurate forecasting of the imminent movement of the exchange rate is very important in investments, trade and economics. Holt Winters smoothing algorithm uses weighted historical trending .

In the stock market forecasting and prediction of share prices, several authors have explored Holt Winters exponential smoothening. Such researchers include Rusu and Rusu (2003), In trying to

predict the share price of companies dealing in gold and oil in the Malaysian Stock Exchange, MSE, Bakar and Rosbi (2016) used Holt Winters exponential smoothening tool.

On their part, Awajan and Wadi (2018) used empirical mode decomposition (EMD) Holt Winters bagging model to help in data forecast. In this model, the bagging method uses various ways of forecasting in order to generate a combined forecasting methodology.

Investment forecasting is the fodder for investment banks and insurance companies. This sector relies on credible forecasting process to make profits.   Unlike the general public which have inadequate risk assessment, desire to obtain abnormal profits, short term investment horizons. Investment banks and insurance companies relies on more efficient and accurate forecasting methods observes Dzikevičius and Šaranda (2010).

2.4 STOCK BROKERS RECOMMENDATIONS

Making investment decisions in the stock market has been a reserve of the stock brokers who work with investment banks. Investment forecast is the domain of investment banks and insurance companies says Dzikevičius and Šaranda (2010). In fact, the brokerage firms have research units that survey the market and come up with share / stock recommendations of various stock picks on whether to buy, sell or hold. Stock broker recommendations given out to investors are always packed as hot tips and are erroneous in many instance says (Pring, 2018). While this information from brokers may appear genuine to many investors, there is a catch! Whether a client buys or sells, the brokers make money. It profits whether you are in the loss making zone or in the profit zone. In fact the more times you follow the broker recommendations, the more commissions the broker earns. The broker earn about 4% for a combined buy and sell decisions.

Many other people rely on their gut feelings to make financial decisions. Gut feelings and banking on hope that a given stock will realize capital gains after one has bought it has not always been the case. In fact many losses have been made before. According to O'Neil, (2009), praying and hoping doesn't work in the stock market. An investor is better off buying, based on sure company fundamentals.

Both the print media and the broadcast media have made it worse for Kenyans. When the stock market is on a bull run (rapid increase in stock prices), it becomes news and many people join the

moving train hoping to cash in only to be slaughtered in the process. Making decisions on market noise have led many into making wrong financial / investment decisions.

Making decisions on stock market purchases / sales should be based on solid facts based on well analyzed data. Such major financial decisions should not be based on gut feelings, stock broker recommendations or on news headlines but from well analyzed data.

2.5 FUNDAMENTAL ANALYSIS

The first mode of analysis is the fundamental analysis. Fundamental analysis in most cases will focus on the company's operations and financial health to determine as to whether the company is doing well or not. This is important in predicting the future profits or losses. Case in point is Kenya Airways, Mumias Sugar Company and ARM Cement. These three companies have had huge debts which have crippled their operations. These have in effect negatively impacted on their share prices. Their shares have reduced in value tenfold in the recent past.   The method is more concerned with the company, rather than the stock. It is important to note that as the company gets affected negatively, the stock gets affected too. The Fundamental Analyst makes their investment decisions based on a company's previous performance to predict the future earnings (Shah, 2007).

Fundamental analysis may involve examination of company financial data, annual reports, the management board, business concept, business operations and the fierce competition a company faces (Suresh, 2013).

Fundamental analysis monitors profits and earnings and evaluates the risks that could spell doom to the company. It uses statistical, mathematical and financial algorithms applied to the financials of the company in order to determine the share price says Rusu and Rusu (2003).

The major focus of fundamental analysis as a mode of stock analysis is to predict future earnings that a stock will achieve and enable the investors to make profits by investing in the company at the present time (Baresa, Bogdan and Ivanovic, 2013).

2.6 TECHNICAL ANALYSIS

Technical analysis as a method of stock analysis focuses on historical stock price movement (as well all know, the stock prices move up, down or sideways), which might show a pattern which will hence be used by stock market players to forecast the future possibility of capital gains or losses which strongly determines the investor's choices of whether to invest or not. It is important to note that both technical and fundamental analysis work hand in hand for good stock analysis. Relying on one alone could be disastrous.

According to technical analysis, the price of a stock depends on supply and demand of that stock at the market place. Technical analysts have developed tools and techniques to study previous patterns and predict future prices. Such tools that are being used by technical analysts include Stochastic charts, Exponential Moving averages, Bollinger Bands among others. Technical analysis is basically, the study of markets only.

According to Rusu and Rusu (2003), technical analysis studies the changes in the share price and assume that the past behaviour has the potential to predict the future price direction of a given stock.

2.6.1 ASSUMPTION OF TECHNICAL ANALYSIS

One of the major assumptions of technical analysis and technical analysts in general is that ; The current or present trends or price patterns are influenced by the past trends and the projection of future trends is possible by an analysis of past price trends.

It is the analysis of the past price trends to determine the future price trends that this project is all about.

While some researchers have made significant profits using both technical analysis and fundamental analysis, others have tried other options of predicting the stocks markets. Such options include ; Deep Learning (Leemarkdeg, 2016).

2.7 DEEP LEARNING.

Deep Learning involves a computer system learning words from financial news and transforming them into vectors. These contextual vectors are fed into the deep learning computer system such as Recurrent Neural Networks (RNN) to predict stock prices says Pinheiro and Dras (2017). The two scholars observe further that the analysis of text data such as news announcement and commentary on events is one major source of market information and is widely watched by investors. A simple statistical technique is used to count the number of positive words and negative words that appear in the news and calculates the sentiment score. Then a regression model is used to investigate the relationship between the stock price and the sentimental score.

It's been observed that the deep learning approach can ride on the volatility and enable investors to achieve abnormal profits from market correction. From Singapore University of Technology and Design, Ding, Zhang, Liu et al (2015) developed a deep convolutional neural network that was able to model long term and short term influence of events on the stock market movements.

Data enthusiasts, Computer Gurus, Stock Market Traders and Researchers among other are exploring other methodologies and techniques for the determining the stock market price direction with an aim of making profits from the exchange. Nobody is interested in making losses. Some of these methodologies include mining data using decision trees.

2.8 DECISION TREES

Different methods have been used previously in the classification of data, one of those classification methods that will be of interest to us, is Decision Trees. It is important to note that Decision Trees is a pictorial illustration of all possible outcomes and the potential paths by which the possible outcomes may be achieved.

In the recent past, performing data prediction on huge data sets have been a preserve of data mining technologies and artificial intelligence technologies like decision trees, rough set approach and ANNs (Al-Radaideh, Assaf and Alnagi, 2013).

Data mining in brief can be defined as the extraction of knowledge from non traditional large data sets and databases. In general data mining is expected to achieve the following; discovery of concepts, data prediction and forecasting, trend analysis. Of importance to us will be data prediction and trend analysis as such could guide in the prediction of stock prices and could be of help to investors in the Nairobi Stocks Exchange.

Some schools of thought opine to the fact that data mining or the extraction of knowledge from large data sets is based on the fact that history can be used to foretell the future and thus historic data can be used to predict the future direction of stock market movement. Hidden patterns from historical data can be a goldmine to the stock investors who are eager to make a profit from their hard earned money says Kannan, et al (2010).

There are various learning algorithms that exist and an appropriate one can be used in training decision trees and artificial neural networks. In deed technical analysis have one assumption that price data leaves a trail. A pattern trail that is visible from far. Using artificial intelligence and data mining techniques, these patterns can be discovered automatically. It is the hope of stock investors and other market players that with the discovery of these patterns from the price data, price prediction becomes second nature. Once prices are predictable, investing then becomes easy, not a hard nut to crack as is the case now where wrong move can result in financial haemorrhage and losing one's capital.

One of the limitations of using a traditional relational database system is over reliance on rules based on raw data. These rules have traditionally been used to help users in the decision making processes. Manual analysis and manipulation of raw data has been used for decades to turn data into knowledge. Manual analysis of turning data into knowledge is becoming impractical as data volumes grow exponentially due to big data. The big data posses the challenge to stock price prediction due to the fact that volume of data being massive and hence the need for stocks data analytics that will be key in helping investors make good financial decisions on their capital investments and reaping a good returns thereafter.
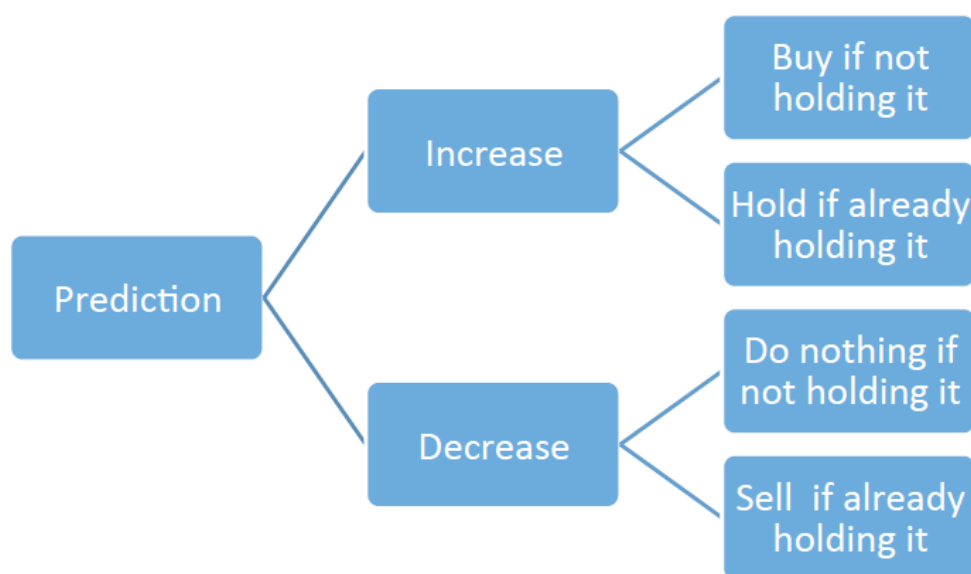
Fig 2 : Decision Tree on stocks prediction.

Source : Dai and Zang (2013) - Stock Price Trending Forecasting using Machine Learning

Decision trees and data mining are the recent ways of making decision to invest in the stocks market, it would be good to look into ways in which greatest investors have been doing stocks investment as well.

2.9 THE GREATEST INVESTORS OF ALL TIME.

In the last 100 years, how have people been making financial decisions especially in regards to stocks investing? According to Milosevic (2016), investors developed some analysis methods

based on the company's current and past financials data and other information. Financial balance sheet is key to the investors.

The father of stocks investing, Benjamin Graham invented 2 rules of investing. These rules are known by virtually every stocks investor in the world today.
Rule 1 - Don't lose money. Rule 2 - Don't forget rule 1. (Hagstrom, 2005)

Benjamin Graham, one of the greatest stock investors in the last 100 years and the man who mentored Warren Buffet in the stocks investment world, invented the two rules of investing. Rule 1, the "don't lose" was one of the guiding philosophy that guided Graham in selecting common stocks. This ensured that he was always having a margin of safety and reduced his chances of getting a loss. When buying a company stock, buy it for a third of its net asset value.

Once a given company stock was bought for a third of its net assets value, Graham was sure of profits thereafter. No consideration was given to a company's assets. All these company assets like machinery, plant equipment could be a product of loans, borrowing and corporate bonds which in turn lowered a company's net worth. When calculating a company's net-worth, all short-and long-term liabilities like bank loans, corporate bonds are deducted.

After all those deductions, Graham remained with net current assets. He would then purchase the stock should the price fall below this share value. The main reason being a margin of safety existed and the risks of capital loss reduced significantly. According to Benjamin Graham, this was considered to be a foolproof method of investing.
Unlike majority of investors who would prefer to buy stocks on market corrections, Graham argued that such a move would be unreasonable as stocks are unpredictable. The trend may not be in your favour.

The second approach to buying stocks according to the father of stocks investing involved buying stocks that had price depreciation and low P/E ratio. Secondly, the company must own more than it owes members of the public. A company that has too much debt would be avoided like a plague

as that would be considered a time bomb. Capital erosion would be eminent should the company fail to meet its debt payment obligations.

Benjamin Graham also advocated for investors to carry out 'technical analysis' on a given company before taking a position. Warren Buffet, one of the students of Benjamin Graham, urges investors to check out the business ethics, management team, financial strength and more importantly value system of a company before launching into the deep-end with any company (Hagstrom, 2005).

While it is recommended for investors to do their own research before investing, many investors are too lazy to do research and hence rely on stock brokers recommendations. These recommendations may be disastrous at times.

2.10   BUYING ON BROKER RECOMMENDATIONS

Many investors buy stocks based on broker recommendations. This can be disastrous to one's finances. Most brokers will recommend either a buy or a sell; neither will they tell you to hold a given stock for long term purposes. Their aim is to make profits on both buy and sell actions. A total of 4% commission goes to their bank accounts.

It is recommended that investors do their own research before investing money. This project will go a long way in helping investors make wise investment decisions having analysed stocks data in their fingertips.

2.11   BUY STOCKS YOU UNDERSTAND.

Perhaps one of the reasons Kenyans make losses instead of profits at the NSE is their lack of understanding of the domains or stock sectors they are dealing in. According to the world's billionaire investor, Mr Warren Buffet, before investing in a given counter, it is important to invest in an industry that you understand.

Buffet argues that it is a good practise to invest in a company you understand so that you can be able to follow its activities with ease. It is then possible to know whether a company is doing well or not. Without such powerful information, you are like a man walking in the dark, you can easily land on a ditch (Hagstrom, 2005).

The idea of buying stocks without understanding the company's operating functions is not healthy. It is a disaster waiting to happen. It is paramount that an investor should understands the company's operating function— the products and services that it sells to the public.

2.12 BUY STOCKS / COMPANIES WITH BIG MOAT.

Kenyans rushed to buy Eveready, Mumias, Home Africa, Deacons, Trans Century among others during the IPO, what they failed to know, is that the two companies didn't have a good moat. They didn't have a competitive advantage. They weren't the market leaders. Soon, Kenyans learnt albeit painfully, that such companies are best avoided.

Warren Buffet advises, investors to buy franchises, companies with a big moat.

In Kenya for example, Safaricom have a big Moat due to its money transfer service popularly known as Mpesa. The big moat have enabled Safaricom stock to grow tenfold, from 3/- to 30/- the current market price.

A moat gives a business organization a clear competitive advantage over others and makes it impossible to beat. Safaricom's mpesa gives it a clear advantage over the competitors like Airtel and Telkom in the telecommunications sector.

The bigger the moat, the more sustainable is the company and the happy the investors as products and services of the company are bound to deliver rewards over a long period of time. Coca cola's competitive advantage have helped it deliver rewards to investors over eight decades and counting. The stock will experience capital gains over a period of time, which is to the advantage of the investors. A dividend given out will be a plus to those who invested their capital in such firms.

A franchise and franchise businesses have been known to dominate the market and wade off competition by being the sole supplier of products people need. With that advantage in mind, franchise businesses have the advantage of increasing prices without losing clients. The ability of franchise to exercise this pricing flexibility is one of the defining characteristics of a franchise; it allows franchises to earn above average returns on invested capital and hence the reason most investors prefer starting a franchise business, that a brand new start up that has no known brand name. Franchise ride on the already created brand name and marketing from the parent company.

Look for the durability of the franchise that is Buffet advice to investors. The bigger the moat, the more durable the business ventures. A big business with a big moat is bound to last centuries. Examples of such businesses include Coca Cola, MacDonald, among others. (Buffet, 1994).

As an investor, your target is to make profit on your investments. It is the responsibility of the company management to work hard and create value for your money. Franchise corporate managers and company directors are charged with the responsibility of getting physical infrastructure like buildings, equipments and finances and using them to create value, which will give a positive return to the firm's investors. What the investors are interested in most is profits from their investments and not the business operations.

Directors from the Franchise Company formulate business strategies to achieve the goal of creating value and creating a return on investment for the company investors. This is done via business activities and daily business operations. In today's current world, a business is influenced by economic, social and political environment in which it operates. In Kenya, however, even the political environment has a huge potential to affect business operations. In the last three elections of 2007, 2013 and 2017,virtually all businesses were affected by the heightened political activities in the country. The better the competitive advantage, the better the chances of survival and the weaker the competitive advantage, the more the chances of failure as the competition is bound to set in.

2.13 FINANCIAL ANALYSIS

The main aim and objective of financial analysis is to use financial data records of a given company to assess, examine and evaluate both the current and past performance to ascertain its sustainability. Digging into the past financial records has a way of helping investors predict the future. It helps business investors in gauging the financial health of the company. A company with a good financial health record is sustainable for a long time. The analysis should be systematic, efficient and finally, make use financial data to explore business issues (Palepu, Healy and Peek, 2008).

Ratio analysis and cash flow analysis are assets for any serious investor worth his salt. He must use ratio analysis in evaluating business ventures. The performance of a given product in the market and financial policies are the main things that ratio analysis focuses on. A product that performs well in the market is bound to improve a company's sustainability. Cash flow analysis focuses on a firm's financial strength and well being. Financial muscle. A good company should be able to meet its financial obligations by paying off its expenses and its debt and remain with a good financial liquidity.

Financial forecasting is key in valuations and hence a very important tool for any investor who wants to invest in the given company. Company evaluation, credit evaluation, financial distress prediction, security analysis, mergers and acquisition analysis and corporate financial policy analysis are key metrics any serious investor must consider before putting their hard earned cash into any given venture.

2.14 RATIO ANALYSIS

The value of a firm is determined by its profitability and growth. A firm that is always registering consistent growth and increased profits will be very attractive in the eyes of investors. As shown in the figure below, the growth of any company is determined by its marketing strategies and financial management. A company determined to grow well must manage its finances, equipment

and assets, liabilities and dividend payouts. A serious growth company must ensure that it retains the bulk of the dividend income as opposed to giving it out to investors.

Ratio analysis as a tool of evaluating businesses has one objective; to evaluate the effectiveness of the firm's policies during business valuations. Business factors and business finances will be weighed and related together. In deed no analytical tool can act as silver bullet that can be used to give answers to all business queries and can be regarded as a good evaluation tool for business ventures.
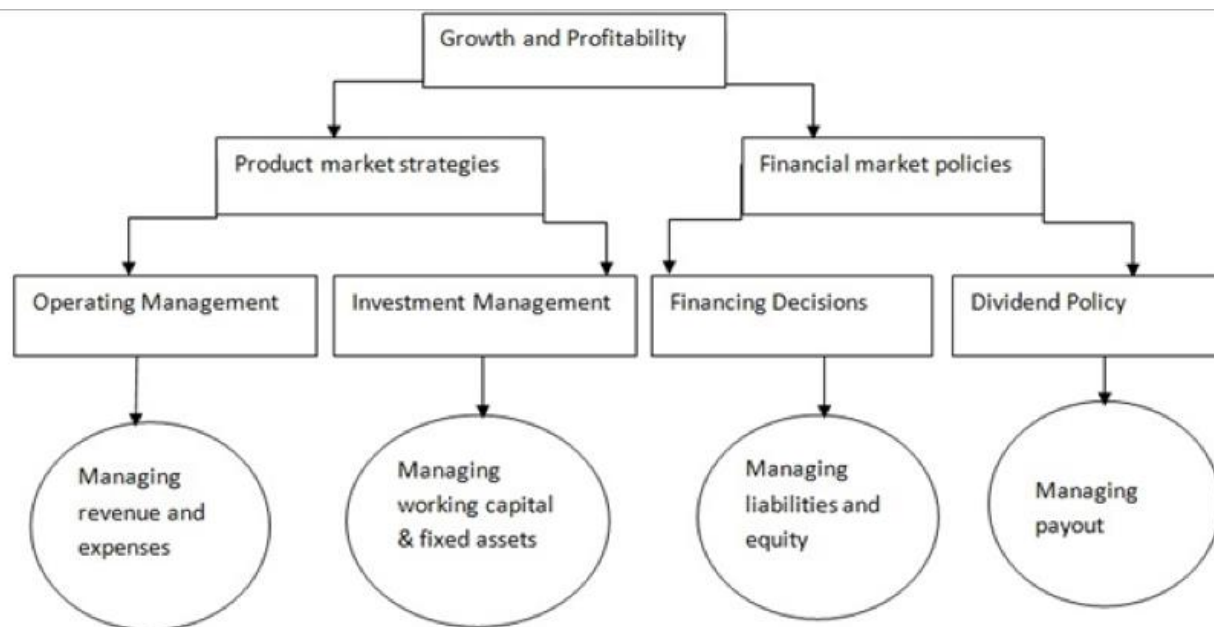
Fig 3. Drivers of a firm's profitability and growth.
Source: Business Analysis and Valuation IFRS Edition: Texts and Cases

2.15 FORECASTING

Data mining can also be combined with forecasting to get some good results. According to Hand, (2009), a combination of data mining and forecasting has a high probability of giving the right forecast. In deed having the right forecast every single time is the objective of any investor in the stocks market – an investor wants a right prediction and forecasting every single time. No investor is interested in wrong forecasts as this can be financially disastrous.

The aim is not to be correct sometimes, but to be correct in stock trading all the times.

Several scholars have found that stocks market prediction via forecasting have been beneficial. According to Wei et al (2008), several factors that affect the stock market notwithstanding, the challenges that affect the stocks market notwithstanding too, stock traders still benefit from stocks trading at the Hang Seng Index.

2.16 PROSPECTIVE ANALYSIS: FORECASTING

Taking cue from the forward looking nature of most investors, financial statement analysis task are also performed with an optimistic mindset. This is simply because investors are most often optimistic in their nature. They are looking forward to making a profit and hence they need an analysis with an explicit forecast. Financial forecasting and forward looking is second nature to a typical investor (Palepu, Healy and Peek, 2008).

Managers and company directors need forecasts to formulate the company's future business plans and provide yearly performance targets which must be met by the company employees; financial analysts on the other hand need forecasts to convince potential investors who are eager to invest and become the new shareholders of the firm; and bankers and those from the financial sector and concerned with loans and debts need forecasts to assess the chances of loan default or timely loan repayment. Those in the banking sector continue to suffer greatly due to the high number of non performing loans. Without assessing the likelihood of loan repayment, loan defaults may soar and this can result in massive losses to the bank and the shareholders of that given banking stock.
`
Moreover, in some cases forecast is summarised in the form of a company's estimated value. Such estimations and valuations are key to the shareholders of the company and the investing public.

Prospective analysis involves – company forecasting and company valuation - together they are bread and butter not only to market watchers, traders, investing public but also to the government

regulatory agencies that work day and night to ensure that the companies are well valued and the investing public are not swindled of their monies.

2.17 OVERRAL STRUCTURE OF THE FORECAST

Forecasting future performance of a company involves - producing a forecast of earnings, cash flows, balance-sheet, income statement. Such a detailed forecast will lay bare all the facts on the table to help investor in easy decision making. Potential risk areas shall have been identified and margin of safety created. Investing will hence be a well calculated risk. Unrealistic assumptions and over optimism by investors is guarded by a comprehensive forecasting approach.

For example, a forecast on growth in sales, earnings and profits require increases in working capital and plant machinery. It is almost obvious that growth in sales and resulting growth in profits is accompanied by heavy investment in working capital and plant machinery. It is impossible to realize increased profits without massive investments in assets and hence such a viewpoint will be regarded as unrealistic assumption.

A comprehensive approach involves many forecasts. Companies in the financial services sector, see the sales forecasts and profit margins as one of the key drivers. With such details being under the radar by the financial analysts, it is impossible to make wrong assumptions and to make a bad investment decisions that may result in looses to the invested capital.

In general, the manager or the financial analyst will be interested in a forecast of cash flows, earnings. Cash flows forecasting include sales, earnings, assets and liabilities. These are a must have for any serious investor who considers fundamental analysis before investing, but not for technical analysis investors who are more interested in market trends and not the underlying numbers.

2.18 70% PREDICTION OF THE STOCK MARKET USING DATA ANALYTICS

The approach taken here to bring together six methods of stock data analysis to help in stock price prediction. With the prediction, it will be easier to tell whether the stock price will go up or down. Stock prices move up or down on a daily basis depending on market sentiments. Bad news like reduced profit, the company making loses will most definitely result in stock prices going down. Meanwhile, good news like increased profits realized or increased earnings, new management in place or the company getting a new strategic investor will in most cases result in increases in stock prices. After these stock price predictions are released, they are compared to the next day's closing price. A 70% right prediction or 0.07 confidence level, would mean that the analysis is a useful aid in financial investing and hence should be considered by investors in the financial markets, those eager for profits (Kannan et al, 2010).

Furthermore, the 70% prediction or 0.07 confidence level results means that it is better than investing in the financial markets by a random selection of the stocks as is currently the case with many stock investors. Generally speaking, companies that trade in the money markets have invested heavily in information technology by buying the much needed financial and stock market prediction software to help them manage their financial portfolios with an aim of getting a good return for their clients. Over the years the financial markets have increasingly huge volumes of historical data stored electronically in various forms and in various databases worldwide both online and offline. This data volume is growing daily as the investing public continue to increase in number due to population increase. This massive data haven't been mined because it is not easy to discern, it requires a third eye. It requires some level of data analysis to make meaning out of it and to benefit greatly. Keeping the data about shares at the Nairobi Securities Exchange is not synonymous with carrying out explorative studies to detect the subtle buying patterns. Such explorative studies cannot be ignored in this day and age of information technology. Gone are the days when investors would take a guess by doing stock selection and investing in the financial markets. In this age, to benefit greatly, an investor needs data analysis of historical prices and explorative studies as suggested previously.

The developments and the improvements in computing technology over the last decade and the automated computing and predictive technologies help investor do automated trading in the markets. Unlike at the Nairobi Securities Exchange where most fund managers, investors and traders do trading manually by making trading decision individually, at the New York Stock Exchange, NYSE, for example most trades are executed automatically by built in computer programs. Data mining can be effective in predicting the future direction of any given stock and when used properly, it is bound to give a massive return on investment for the investing public. It goes without saying that every investor is keen on profits and would easily embrace such low lying fruits that comes with data mining, data analytics and predicting stocks movements. The technology unlocks concealed patterns from stocks data. These hidden patterns aid investors in making the right investment decisions. The resultant investment decisions made have a 0.7 confidence level or 70% accuracy. This is fairly good compared to 100% losses that investors get from the markets. Case in point investors in Mumias sugar company and ARM cement. The stocks have experienced more than 300% capital erosion in the last four years.

Data analytics can be used to maximise the predictions of financial stock markets.

Predictive patterns from quantitative time series analysis and data mining using quantitative analytical techniques have been one of the greatest gifts that technological advancement have given to investors in the financial markets. These technologies are helping stock market investors to discover patterns in the data that assist in knowing the buying and selling points of equities. Buying and selling equities have made millions of people rich worldwide with the best leading examples being Warren Buffet, Alexander Elder, and Charles Munger among others. With the discovery of these strategies that outperform the markets, investors have every reason to be happy smiling all the way to the bank.

2.19 OTHER WAYS OF STOCK PRICE PREDICTIONS:

2.20 ARTIFICIAL NEURAL NETWORKS

Stock market data gets generated every second and are highly time variant and usually non-linear and using such data to predict stock direction is a challenging not only to investors, but to the data analysts too (Patel and Yalamalle, 2014).

Prediction of stock price movement is not new. Investors have tried this in the past 100 years countless times with minimal results. Thanks to the developments in technology, prediction now assists traders in guessing stock price movements. Such knowledgeable information is important and can be utilized in decision making by customers and the investing public in making decisions on whether to buy or sell a particular equity or stock.

Buying a stock that is bound to experience capital gain is a big plus for any investor as they will make significant profits and buying a stock that is bound to experience a decrease in stock price will erode the paper wealth of any investor.

Many researchers are conducting studies to predict stock price movement either via data mining or data analysis methods. Other scientists opined to the fact that Artificial Neural Network (ANN) techniques can also be used to achieve the same results of foretelling stock movement. The past historical data stored in databases at the stocks market can be obtained and that historical data of a few selected stocks is usable in building neural network models which attempt to envisage future stock direction. The results from the artificial neural network model are then compared with the real stock market daily prices to determine the accuracy of the artificial neural network model. The

higher the accuracy and similarity of the model results and actual data, the more the confidence the model gains among investors.

A neural network, just like the biological neural network operates via a set of interconnected links with associated weights. With regards to biological neural networks, where neurotransmission is the main mode of communication between neurons and achieved by the movement of electrical signals across a synapse. Artificial Neural Network predictive capability is creating a new field of making reliable predictions able to optimize profits for the investors. At the moment, Artificial Neural Networks are a common occurrence in the financial, manufacturing among other sectors. ANNs aids scientists and computer technologists among other professionals in forecasting about probable outcomes of a given phenomenon under study. The probable outcomes enable investors for example to make concrete investment choices and get a good return on their capital investment.

Any ANN like earlier mentioned and compared to the human body neurons can be categorized into three layers; the input layer, the hidden layer and the output layer.

Have a look at the illustration of a neural network.

Input Layer   First Hidden Layer   Second Hidden Layer   Output Layer
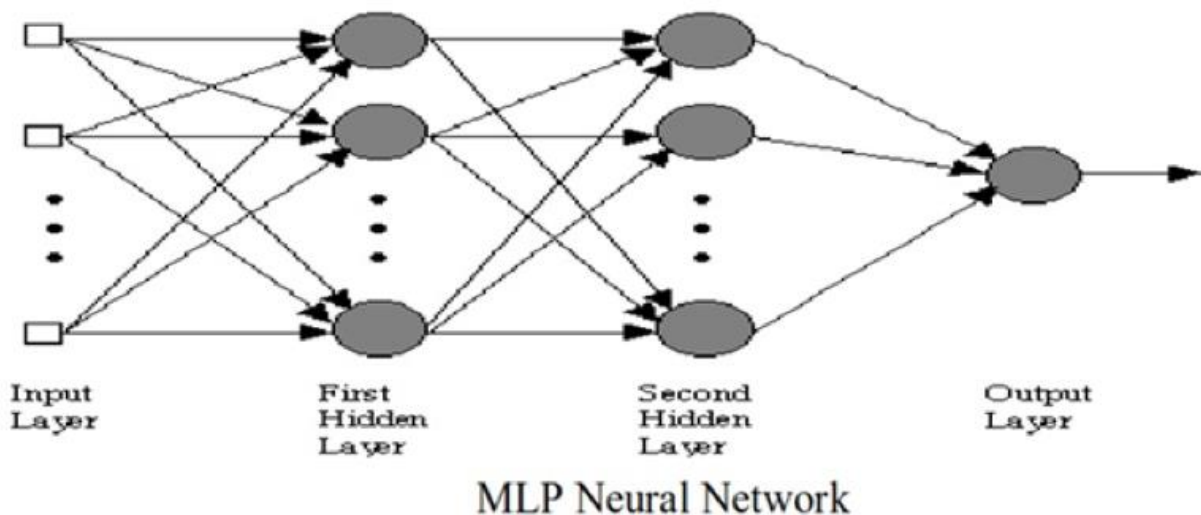
MLP Neural Network

Fig 4: Neural Network.
Source: Stock Price Prediction Using Artificial Neural Network. Patel and Yalamalle, 2014.

Dase R.K. et al. (2010) in their research study found that predicting stock markets and stock price movements using traditional time series analysis can be very challenging. It is not a walk in the park. However, they discovered that an Artificial Neural Network may be suitable for the task of predicting stock market prices with greater success. A Neural Network can extract useful information from large set of data. This useful information that is extracted from large datasets is very helpful to the fund managers who are eager to make a profit from the massive investment capital under their portfolio.

In their conclusions, Dase R.K. et al. (2010) observes that ANN is a must have tool for serious traders and investors as it aids in predicting and foresting the markets. This is something that we can borrow and adopt for predicting our very own stock market, the Nairobi Securities Exchange, NSE. ANN's ability to do correct prediction has enabled it to be used in other sectors like; financial sector, energy sector, manufacturing sector and sales predictions. It is interesting to note that ANN can be applied in virtually all fields. Over the years ANNs have found their way in mathematics, engineering, economics, meteorology among others (Kalogirou, 2006).

Akhter Mohiuddin Rather (2011) like many other researchers before have tried to predict stock market returns using ANNs. In his research study, he employed error metrics to evaluate the performance, this is important in the sense that errors would mean wrong analysis. Such wrong analysis when followed by the stock market investors would result in massive looses and hence should be avoided. Such models need repeated tests with real stock data to ascertain their accuracy. In this case, data from India Stock Exchange was used.

According to Wanjawa and Muchemi (2013) Stockbrokers, fund managers and financial advisors in general are better equipped with predictive systems to advice their clients better. The investing public need good investment advice and stock picks. A good investment advice that will result in investors making profits and having a good return on their capital investments is very important. He concluded that such a system can exploit the artificial intelligence of the ANNs.

A third of stock market prediction and certainly not the last method that will be discussed here is Time series method. A time series is a series of data points indexed in time order.

The other method of predicting stock prices that scientists, investors, fund managers and stock analysts have tied to use is using agents based artificial intelligence.

The learning process is very fundamental for the successful deployment of artificial intelligence systems. By gaining knowledge through the learning process, an agent increases its competency levels. Agents in artificial intelligent systems learn via supervised or unsupervised process. Stock movements form a pattern. These patterns are what the agent learns about. The good profitable moves by agents are rewarded because they result in profits. The bad moves on the other hand, result in losses and hence agents need punishment for such erroneous moves which result in losses for investors. The last mode of learning is the supervised learning.
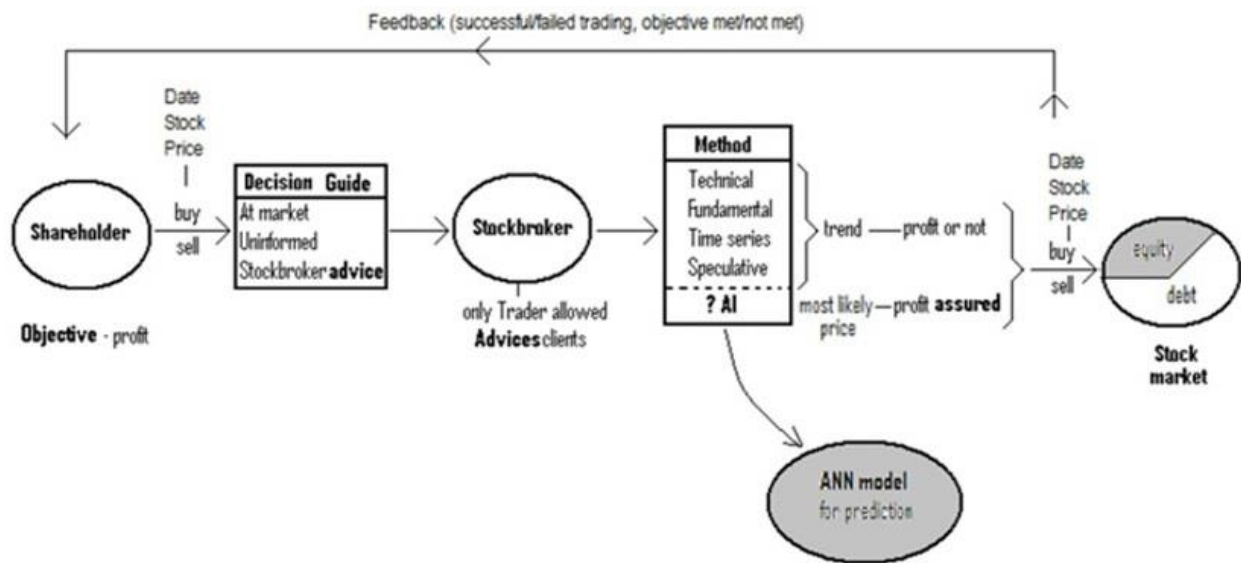
Fig 5 – *Illustration of the Stock Trade Cycle and Methods*
Source: Big Data Stream analytics for correlated stock price movement prediction-
Zhang, W., Lau, R., Li C., 2012.

2. The Big Data Stream Analytics Framework

 BIDSTA has 7 layers for effective forecast of stock price movement
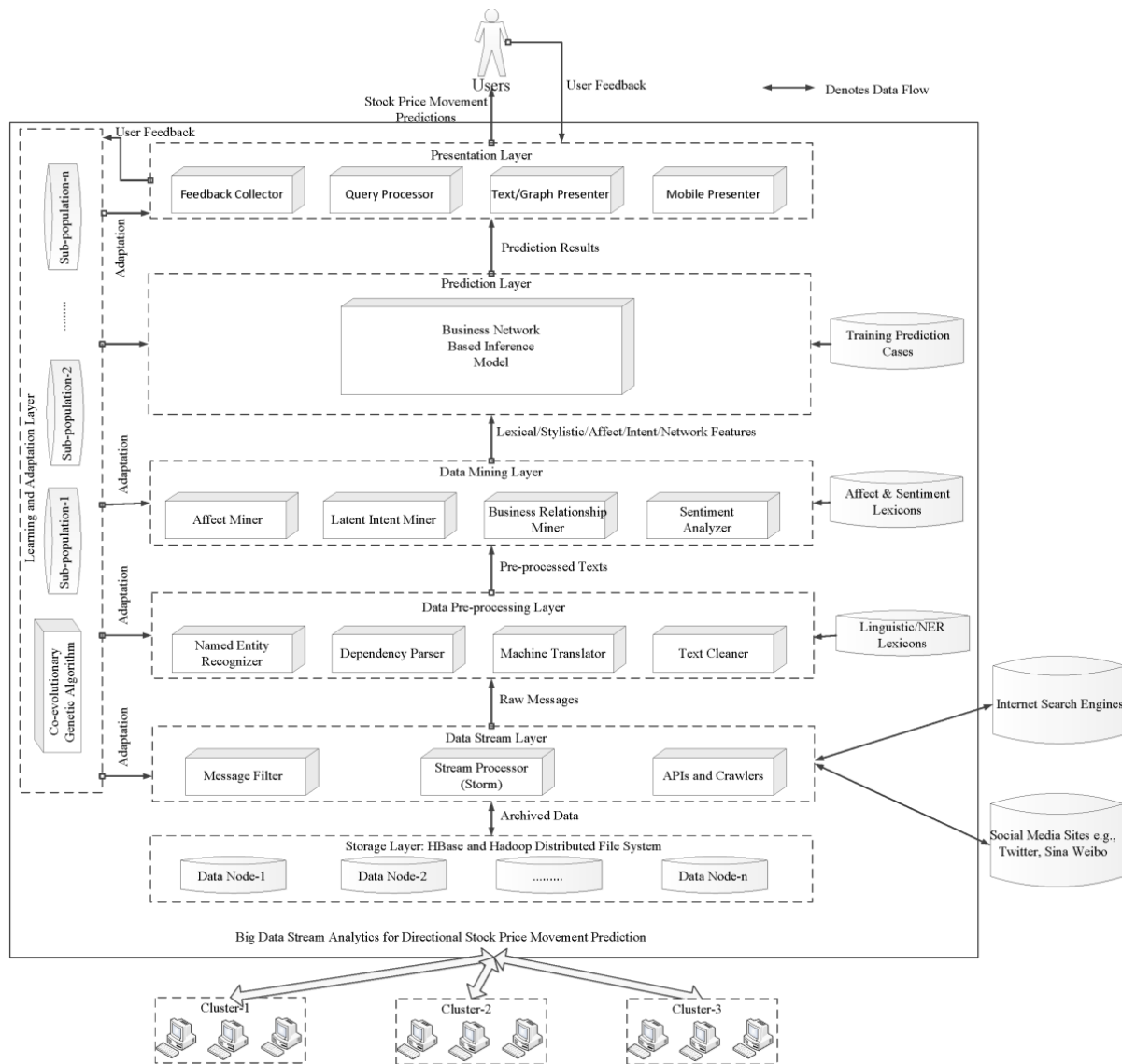
Fig 6. The Big Data Stream Analytics Framework

Source: Big Data Stream analytics for correlated stock price movement prediction-
Zhang, W., Lau, R., Li C., 2012.

The emergence of big data has provided scientists and data engineers with analytic opportunities and the possibilities of making meaning out of complex data. The other analytical opportunity provided by big data is the possibility of detecting deviations, anomalies from large data sets. (Singh, Dimri and Rastogi, 2014).

According to the traders and investors who religiously follow efficient market hypothesis, the equities and the financial markets in general reacts to news flow from all corners of the globe. With the improvements in technology and the emergence of social media, news and information comes in many formats ; text, picture, video, among others. The traditional database of data storage is overwhelmed by the different data formats we have today and it is not effective in handling data prediction and forecasting.

2.21 NATURE AND BEHAVIOR OF DATA IN STOCK MARKET

There is one serious concern with data storage among stock market players, intermediaries and stock brokers. Kenya has over 30 stock brokers, each with their own independent data storage system but all storing the same data. This is a global problem of data duplication. Take India for example, all the 83,000 stock brokers and intermediaries are developing own repositories which result in same data being stored in 83,000 different locations. Green computing practitioners avoid such wastage of resources and unnecessary costs. Information from around the globe could be unsettling like the financial meltdown in 2008-2009 which saw many investors lose their investment capital already in the markets.

2.22 BIG DATA ARCHITECTURE DESIGN FOR STOCK MARKET DATA

For successful stock analysis and prediction by data scientists, data is a mandatory requirement. This data is stored by intermediaries in different formats and databases at different places around the globe. (Singh, Dimri and Rastogi, 2014).
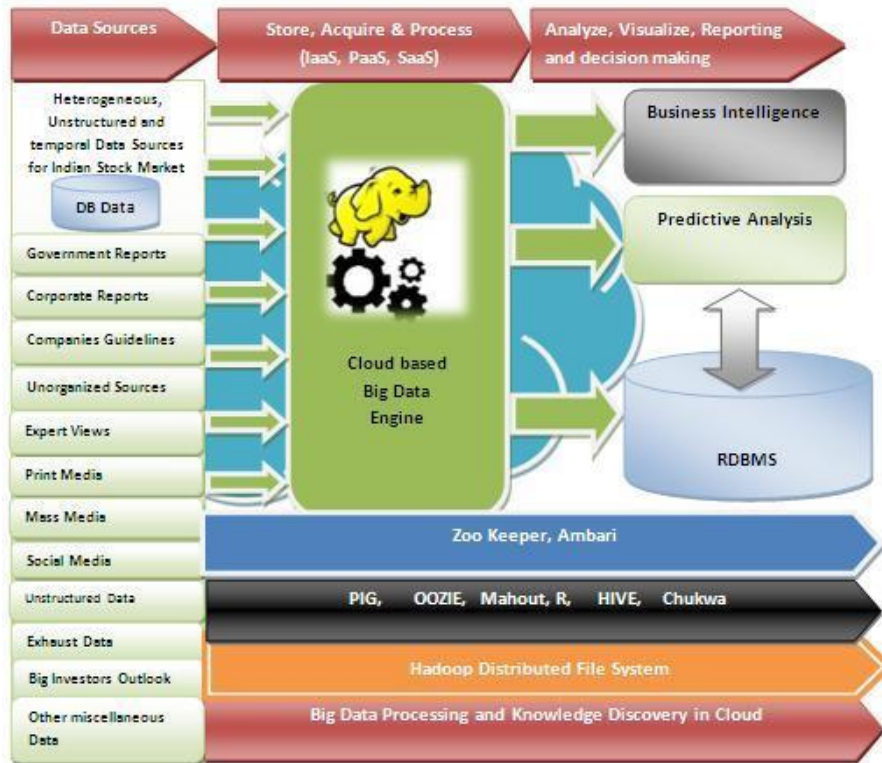
Fig.7 - Big data architecture design for stock market

Source: The Implications of Big Data in Indian Stock Market. Singh, K. K, Dimri, P. & Rastogi, K., N., (2014)

The financial market and stocks market in particular have a big data architecture that is divided into three major parts. These three major parts that make up the big data architecture are; data sources, storage and analysis. Data sources consist of news and information flows. News comes from all corners of the world and they can have serious ramifications on the financial markets which can lead to financial haemorrhage and loss of investment capital. To begin with, identify data sources that will be required for future prediction of the equities and the stock market. Secondly, after identifying the data sources, process unstructured data for heterogeneity.

 Finally, analyze the figures, visualize the numbers and make appropriate decisions that will help the investors make profits from their investments.

Financial resources are in limited supply to investors from third world countries compared to their counterparts in the first world countries. With that in mind, it is imperative to note that the easily affordable and technologically equipped big data architecture will help investors easily visualize the stock patterns and make good investment decisions that will help them make profits from their investments.

2.23 CAPITAL MARKET - DRIVERS FOR BIG DATA TECHNOLOGIES

Storage, management and data analysis is an uphill task for players in the financial markets. This has been the case for decades. The financial market players are always trying to build an easily usable and easy to visualize datasets. This endeavour has been expensive and billions spent in the process. Despite the fact that data sharing across the network would make life easy and data management cost effective for the capital market firms, each broker prefers working with his own dataset. This initiative of working independently with individual data sets has made enterprise wide data analysis extremely difficult.

Data analysis has been made difficult again by the number of sources of unstructured data. Advances in technology have made traders interested in a particular company acquire news analysis systems and tools that aid them in getting profits trading on news items. In the stock market, the most efficient traders and those who are quick to react to news benefits greatly. Those who wait for news over the mass media : radio , television, newspapers may miss the ride or other enjoy a short stint of it compared to those with tools that analyze news.

The emergence of social media has added a spanner into the works as stock market and the fund managers have additional channels to receive news in audio, video and through social media. Firms have to develop enhanced ways to analyse not only social media data but also data from many varied sources.

As firms continue to receive and analyse data from different sources, data storage continues as well to grow exponentially into terabytes and petabytes, making traditional relational database unable to store and handle the data. Advancement in computing allow storage, analysis of big data.

2.24 STOCK PREDICTION BASED ON TEXTUAL ANALYSIS OF FINANCIAL NEWS ARTICLES

Nowadays, a large number of online resources and websites offer valuable information about equities market, commodities market and financial market in general (Shah and Mohri, 2007). It is evident that a majority of these online resources comes from financial news articles about listed companies on various stock exchanges worldwide, company annual reports and expert recommendations. Experts in the stock market have not only authored good books, journals, magazines and various publications, they have online blogs where they give titbits of information on various stocks to buy or sell. Data from these online blogs are in text format and which makes it harder to use, compared to the numerical format. The main challenge now involves Mining of Text Documents to get the valuable stock market information in text format and Time Series Analysis to help in data analysis and stocks prediction.

Based on the news and information from the online resources and expert recommendations and analysis, the news impact on a particular equity can be either: Positive, Negative and Neutral. Positive implies that the given equity will experience capital gains after the news is released. Negative means the stock will decrease in price and neutral means that the stock price will not change.

Alternative studies relate to getting stock price movement patterns from the online resources and obtaining news alerts and pieces of information which correlate to a spike or decline in the stock price.

An automated crawler continuously crawls news articles looking for specific information from the documents. Words like "profits rise" or "share prices will go down" from news articles and blogs are index.

The general architecture of such a system is shown below:
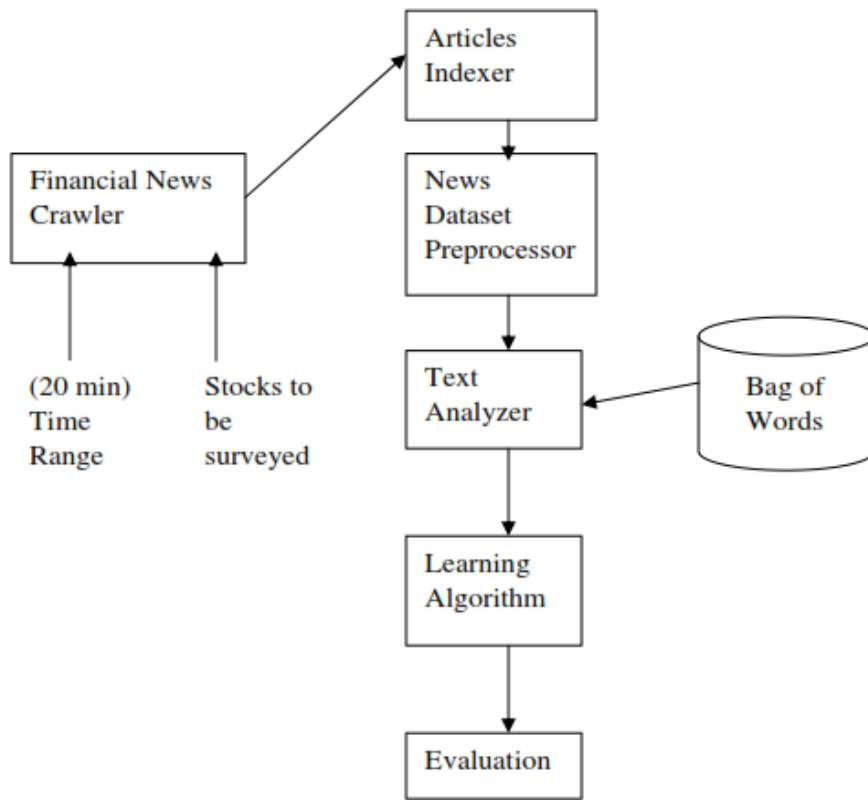
*Fig. 8. **Stocks Prediction based on textual analysis.***
Source: Foundations of Machine Learning. Shah, V., H., Mohri, M., (2007).

News items from credible sources can be given more weights as opposed to all articles being given the same weights. Secondly, news items with positive or negative predictive terms can be assigned more weights as news items with neutral term gets lower weights.

2.25 DATA MINING
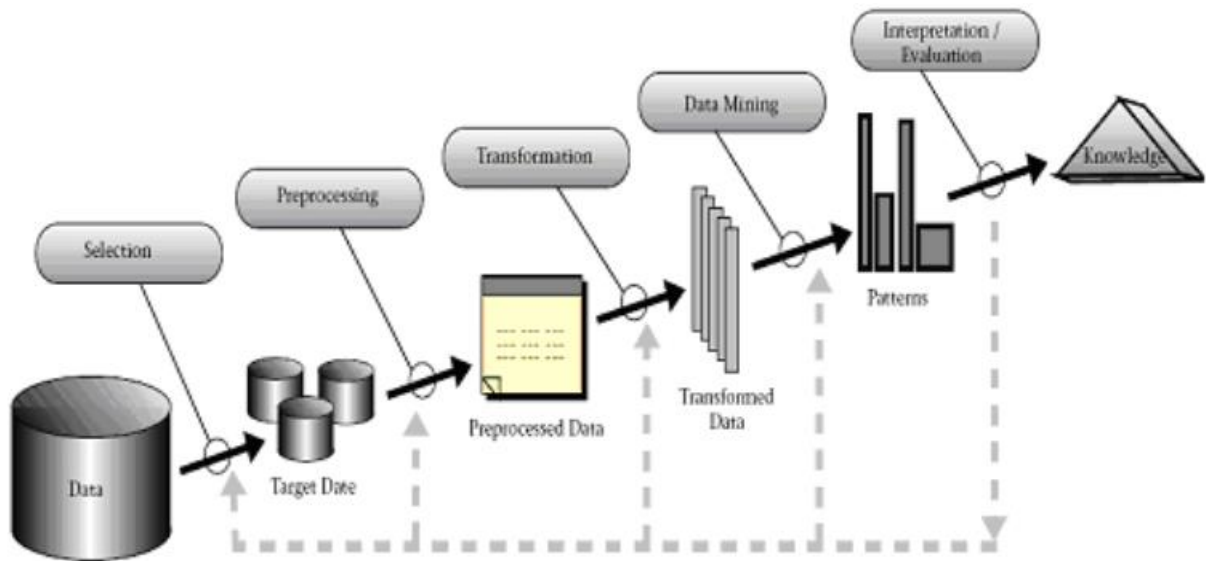
Fig 9. Data Mining

Source: Foundations of Machine Learning. Shah, V., H., Mohri, M., (2007).

2.26 SOCIAL MEDIA AND STOCKS PREDICTION.

The latest mode of news dissemination has been the social media sites such as Snapchat, Facebook, Instagram, Twitter, among others where almost everyone practises citizen journalism. Anyone can record and report on breaking news. While the majority of the social media users use it for entertainment purposes such as music, videos, photos, memes and news briefs, some others use the medium for predicting the stock price movement and stock market direction and making money (Bollen et al, 2010).

Bollen states further that, news as we know it is not easy to predict, but with continuous posting of the country's economic and commercial activities by users on social media early indicators may be obtained which also affect the stock market. It is a known fact that breaking financial news do affect the stock markets positively or negatively. (Schumaker and Chen, 2009).

Those who use the medium for making money to do so via, predicting the stocks market, using the words, tweets, news briefs emitted from social media.

According to several authors, social media content is key in tracking sentiments that show the public mood. These authors include Gilbert and Karahalios(2010), Mishne and Glance (2006), Liu et al (2007), Dodds(2009).

2.27 STOCK CHARTS

According to several other authors like Hughes(2014), stocks prediction can be based on charts. Chuck Hughes argues that predicting the stocks market can be very difficult and hence following the charts is an easy option, which can bring great rewards.

2.27.1 BUY SIGNAL

Hughes observes that a stock is best bought, when it has an uptrend signal. Uptrend is determined when 50-Day Exponential Moving Average line and the 100-Day Exponential Moving Average lines cross each other on the way up. In the figure below, a buy signal is shown for Apple stock.

A down trend or price reversal is indicated when the reverse happens, the lines (the 50Day) cross each other on the way down. If you are a short term investor, you should consider selling the stock when this occurs as shown in the next figure.

Fig 10. Stock Charts, Buy Signal

Source: Selecting Profitable Trades. Hughes, C.,(2014)

2.27.2 SELL SIGNAL

A sell signal is indicated when the 50-Day Exponential Moving Average cross below the 100-Day Exponential Moving Average. Short term investors owning Merck stock should consider selling the stock, since the price is headed south.

Technical analysts and stock chartists opine that the fifty days and one hundred days EMA line is one of the easiest ways to forecast stock market share price trend. The trend forecasts the buying or selling a stock.

Fig.11 - Stock Charts, Sell Signal

Source: Stock selection secrets, Hughes, C.,(2014)

2.28 CONCEPTUAL MODEL
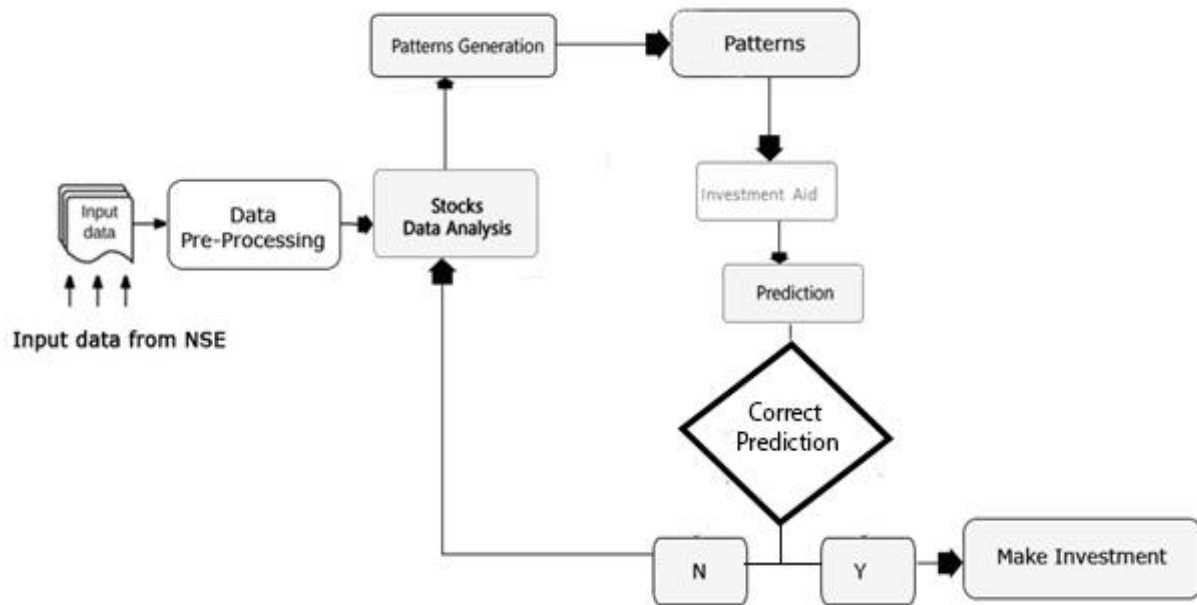
Fig. 12 - Stocks Price Prediction using Data Analysis.

2.29 CRITICISM OF PREVIOUS WORKS.

While Artificial Neural Networks have performed well in the prediction of the stock market prices as several scholars have shown. Having more neurons doesn't necessarily result in a better network and hence better prediction results. Having more neurons leads to over fitting hence a balance needs to be struck between having many neurons and few neurons. (Shah, Mirani et all 2016).

Since its discovery in 1940s by Warren McCulloch and Walter Pitts, Neural Networks, NN, has proved that it can be used to compute any arithmetic and logic functions. This being true, however, NN, has some shortcomings; complicated networks have reliability issues, statistical and discriminate analysis is a must to get training data sets, new IPOs have zero historical data and hence impossible to compute and predict future prices using this methodology. Integrity of results can be questionable since results depend on the data models used. Two data models could give 2 different results altogether. (Goel, Poovathingal, Kumari, 2016).

The econophysics studies have shown uncountable number of times that stock price distribution does not follow the normal distribution and hence some scholars have found it worthy to use Bayesian networks to predict the stock price.

It has been shown that Bayesian networks are probabilistic in nature. They are not just probabilistic; they are also graphical. They use the laws of probability for prediction and anomaly detection. They use Directed Acyclic Graph (DAG) to analyze data and their correlation. (Tabassum and Majumdar , 2018)

BAYESIAN NETWORK USES

It is used in anomaly detection, diagnostics and medicine prescribing. It can be used for the technology behind biomonitoring, document classification, information retrieval, semantic search, image processing, spam filter, turbo code, system biology etc. Reasoning and decision making under uncertainties are great use of Bayesian networks.

2.30 JUSTIFICATION AND SIGNIFICANCE OF THE STUDY

This project intends to come up with a viable solution to the problem of investors incurring massive losses at the NSE – the proposed solution is Holt Winters exponential smoothening forecasting model. Using historical data, from the past stocks trading prices, stock predictions will be done using data analytics to help investors make better financial decisions.

In the New York Stock Exchange (NYSE), many investors have lamented having missed Microsoft Corporation stock. Many others have lamented over missing Apple computers stock. These stocks have turned out to be some of the best investment picks and have resulted in a return of investment to the tune of thousands of money leaving the investors smiling all the way to the bank. Other good performing stocks include Cisco, Apple, Microsoft and Google among others.

In the words of Minervini, (2013), from March 1989 through May 1993, Cisco Systems averaged over 100 percent quarterly earnings growth. Cisco's stock price advanced more than 13-fold in that period.

It is evident that investors have made serious gains from investing in the stocks market. It is my belief that computing tools can help deliver such good performances via predictive data analytics.

According to O'Neil, (2009), some two American stocks, the likes of Home Depot and Microsoft both increased exponentially over a ten year period, from 1980s to 1990s. The stocks increased 20 fold. In fact Home Deport was one of the best performing stock immediately after the listing on the stock exchange.

In the digital world we live in today, large volumes of data from the stock market arises from online news sites like Yahoo Finance, social media sites like Facebook, Twitter, search queries from the search engines like Google, Bing among others. With these developments scientists need

research and more so development of new tools to process these large volumes of data. In most cases, big data is known by three dimensions; volume of data to be analysed, velocity or the speed with which the data is flowing in and variety of data types involved (Boden et al. 2013). Volume is the amount of data to be analysed or processed. The data is not only generated by employees, but by machines and network interactions too. Variety is the number of data types involved. Originally, data would be stored in spreadsheets but not anymore. As of today, data comes from several sources like emails, photos, videos, audio and files like pdfs. Velocity is the speed with which data flows in from the various sources for consumption by the management and subsequent decision making. The faster the flow of data from various sources the better for stock market players who benefits immensely from stocks analysis and predictions.

Study Significance?

Using data analysis of stocks market data to aid investors in making financial decisions at the Nairobi Securities Exchange will go a long way in helping investors realize a good return on their investment. With good investment returns, there are many benefits that can arise from it. One, the economy will experience growth, as investors will have more money that can be invested in other sectors of the economy.

After many investors have embraced the benefits of making financial decisions based on analyzed data and having realized good financial gains, the ripple effect is that many other investors will be interested in investing in Kenya. Kenya is bound to experience migration of tourist investors and increased Foreign Direct Investment, (FDI) who will be keen to invest in the stock market to realize the said capital gains and get a good return on their investment capital. With increased FDI, the country is bound to experience economic growth.

 Since Kenya is an economy driven by tourism and agriculture, the increase in tourist numbers is bound to lift the economic growth of Kenya as a country.

CHAPTER 3

RESEARCH DESIGN AND METHODOLOGY

Predictive analytics of stocks market data is based on quantitative research method. This is the method associated with measurements accuracy and repeatability of vital importance.

3.0 Constructive Research

Quantitative research method is based on constructive research. This is done by technical professionals to find a new solution to a specific persisting problem. The persistent problem in our case is predicting the future stocks volatility to enable investors to know with certainty whether they'll make a loss if they put their money on a given counter or they'll make losses.

The term "construct" refers to the new contribution being developed. Such new contribution could be a new theory, algorithm, model, software, or a framework observes Kumar (2011).

Trend studies

This project studies trends at the stock market hence, a trend study. It helps in mapping change over a period of time. By studying the past and the present, it becomes possible to predict the future with great accuracy. Trend studies help in forecasting.

Fund managers, Economists and players in the field of finance have explored the possibilities of using trend studies in forecasting by making use of the past and present trends.

As previously discussed, there are several ways of analysing stocks. Some researchers have tried combining five methods of analyzing stock market daily prices to forecast the possibility of the next day closing price being up or down (Kannan et al, 2010).

This project will take into cognisant Holt Winters exponential smoothening model of forecasting. It is a remote form of Moving Average Concept.

3.1 QUANTITATIVE RESEARCH DESIGN

OBJECTIVE ONE

Investigate how Holt Winters exponential smoothening forecasting approach to stock prices can be applied to Nairobi Securities Exchange historical stock market prices and develop predictive patterns.

OBJECTIVE TWO

The project aims at developing stocks investment aid software based on Holt Winters Exponential smoothening. It will help in predicting patterns of stocks prices.

OBJECTIVE THREE

Do a comparative analysis on the performance of the tool developed in this project, the Holt Winters Exponential Smoothening stocks forecasting and prediction tool as regards stock price predictions and other tools developed in previous research studies on machine learning like Artificial Neural Networks and Bayesian Network.

OBJECTIVE FOUR

Check accuracy of predicted results using Holt Winters exponential smoothening model against live Nairobi Stock Exchange stock prices.

3.2 METHODOLOGY OF SOFTWARE DEVELOPMENT

The SDLC is a process that ensures that programmers and system analysts are tasked with building good software. The end product is good functional software. The 5 phases ends with the implementation. (source www.airbrake.io)
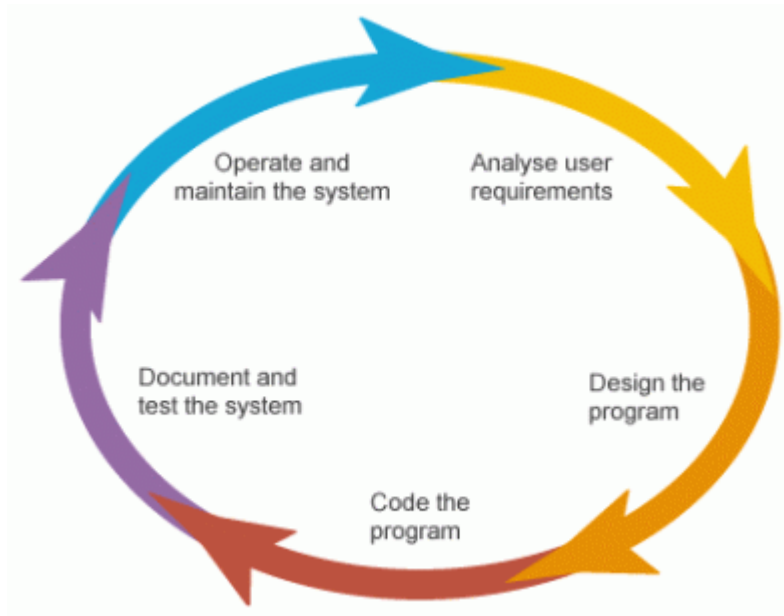
Fig.13 **The Software Development Life Cycle.**

Source.www.airbrake.io

Requirements Gathering/Analysis

This phase is critical to the success of the project. Expectations need to be fleshed out in great detail and documented. Once documented it is important that the parties sign. This is an iterative process with lots of back and forth communication taking place between stakeholders, end users and the project team. In most cases, users are bound to add some requirements they need implemented on the system.

Design

Technical design requirements are prepared in this phase by lead development staff that can include architects and lead developers.

Coding

The development team composed of system programmers, system analysts, user experience designers among others are deeply involved in the actual system coding and testing the process. Developers should be ready changes can be introduced.

This project will utilize waterfall software development life cycle model.

Testing / System Evaluation.

The system will be tested / evaluated against the current Nairobi Securities Exchange prices to test how well it has been able to predict the current day trades. Once the current trades can be predicted with accuracy, then we can be sure that the future trades will be predicted too.

3.3 WATERFALL LIFE CYCLE MODEL

The Waterfall model has a number of sequence of stages that must be accomplished one after the other without skipping any step. Each phase is completed and verified before development progresses to the next phase.
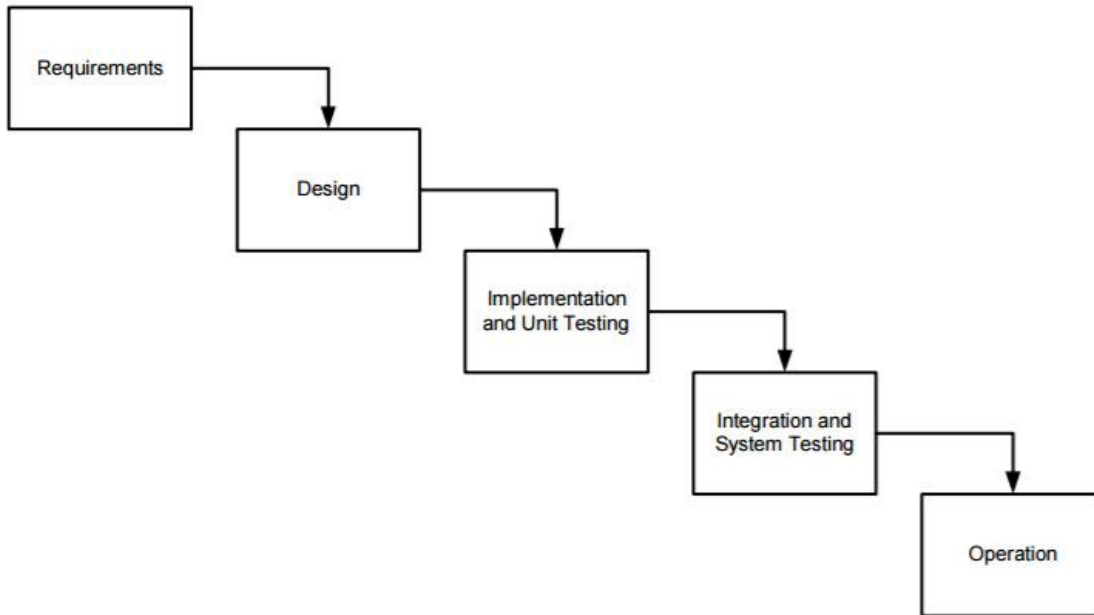
Fig 14. The Waterfall life cycle model. *Source: Kasetsart University.*

NSE DATA ANALYTICS TOOL EVALUATION

The Holt Winters Exponential smoothening forecasting tool will be evaluated against the current day stocks prices to determine the accuracy of the tool. Continuous evaluation will be done and compared to the real values.

An accuracy of 80% will be preferred. This will help many investors make good financial decisions.

An 80% prediction will be determined by how frequently the system does correct predictions. Out of 10 stock picks, 8 should have correct prediction s as to whether the stocks will move up or down.

3.4 PROJECT RESOURCES

The project will need a number of resources for successful project completion. These includes:

a. Hardware and Software - Laptop and python software, as the development environment.

b. Human Resource - 2 Research Assistants

c. Financial Resource - Salary to Research Assistants

3.5   PROJECT BUDGET

No.	Item	Cost
1	Laptop	50,000/-
2	Unlimited Wireless Internet Connection * 6 Months	20,000/-
3	2 Research Assistants	60,000/-
4.	Administration Cost (Telephone + Transport + Lunch)	30,000/-
5.	Insurance Cost	90,000/-
	Total	200,000/-

Table 2 : Project Budget.

3.6   PROJECT SCHEDULE
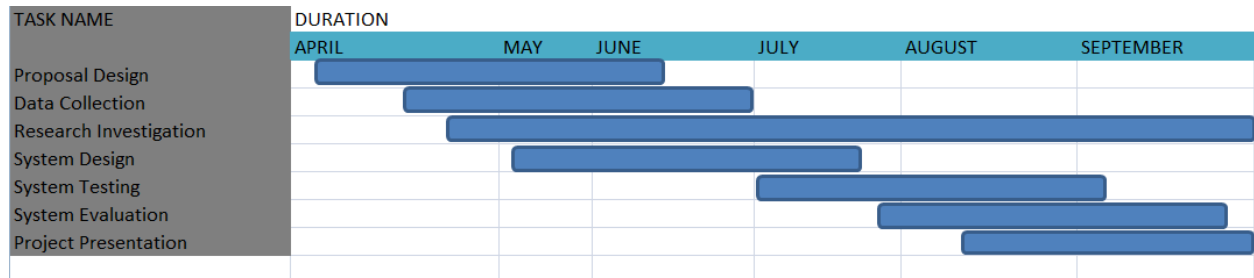
Figure 15 : Project Schedule

CHAPTER 4

RESULTS

In this Chapter, I will be showcasing the results as per the Holt Winters Method described in the previous chapters.

4.0 INTRODUCTION

Having analysed stock market data from the Nairobi Stock Exchange and more specifically, Nation Media Group, the leading media house in East and Central Africa, the results are in tandem with the movement of the stock over the last four months and the predictions are worrying. The stock is headed for a crash if it has not crashed already. The stock is on a freefall. The stock has been bleeding and the financial haemorrhage is catastrophic to the investors who bought the stock at a premium price of 320/- per share. The same share is now valued at 18/-, a 94.375% drop in paper wealth. Kshs. 1000,000 (One million shillings) investment in 2007 is now a paltry Kshs. 56,250 some 13 years later, debunking the myth that long term investment pays in the long run.

Long term investment has been beneficial for a select few companies at the Nairobi Securities Exchange and has been very traumatizing for a number of other companies at the Exchange. The bear run has been unforgiving and the predictive stocks analysis confirm the same.

NATION MEDIA GROUP

NMG - Plot of Historical Data.

NMG Historical Data - Stock Price Plot
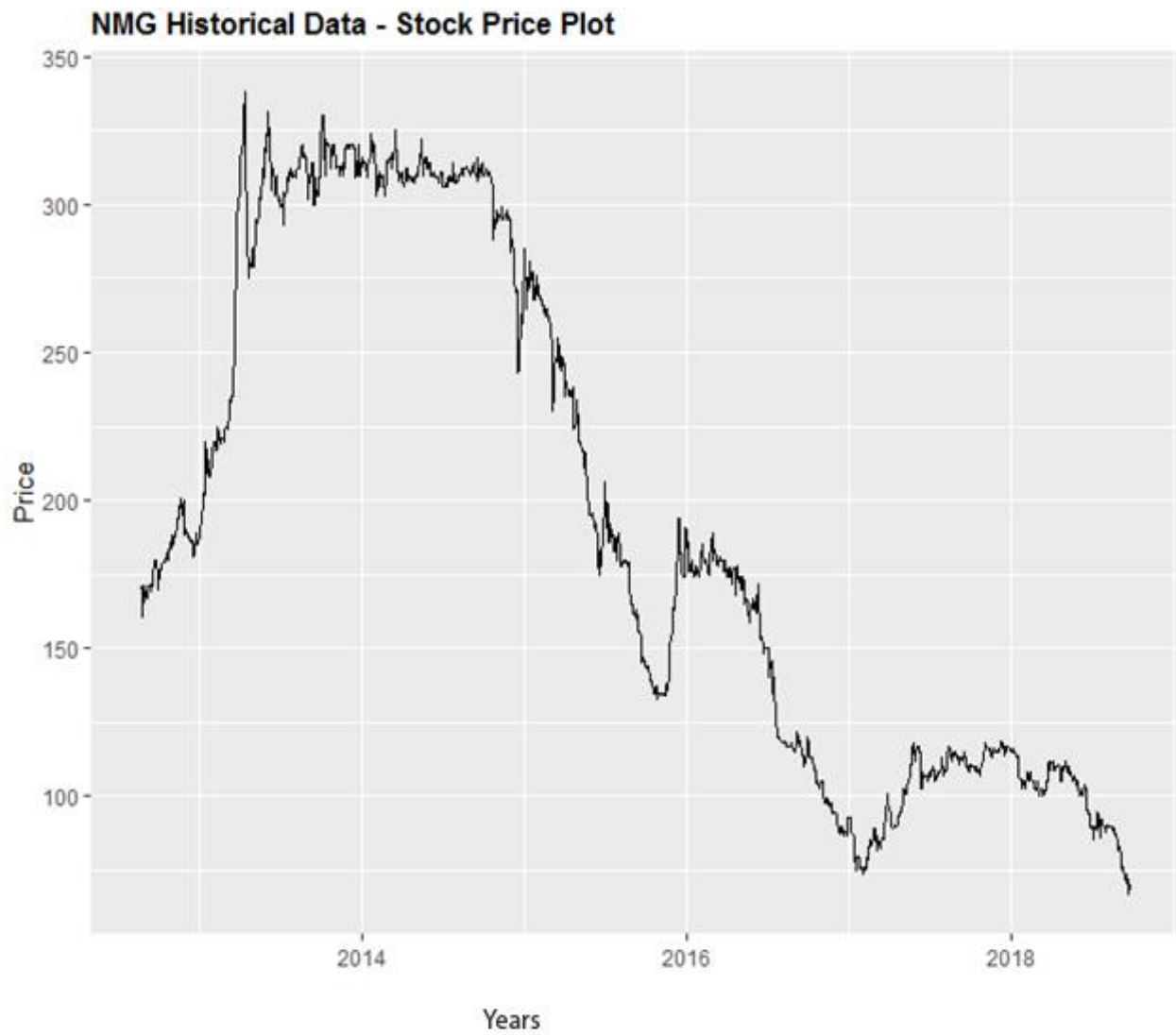
Fig 16. A Graph of Nation Media Group Historical Stock Price from 2013-2018

NMG Historical Data - Holt-Winters Filtering, Exponential Smoothing
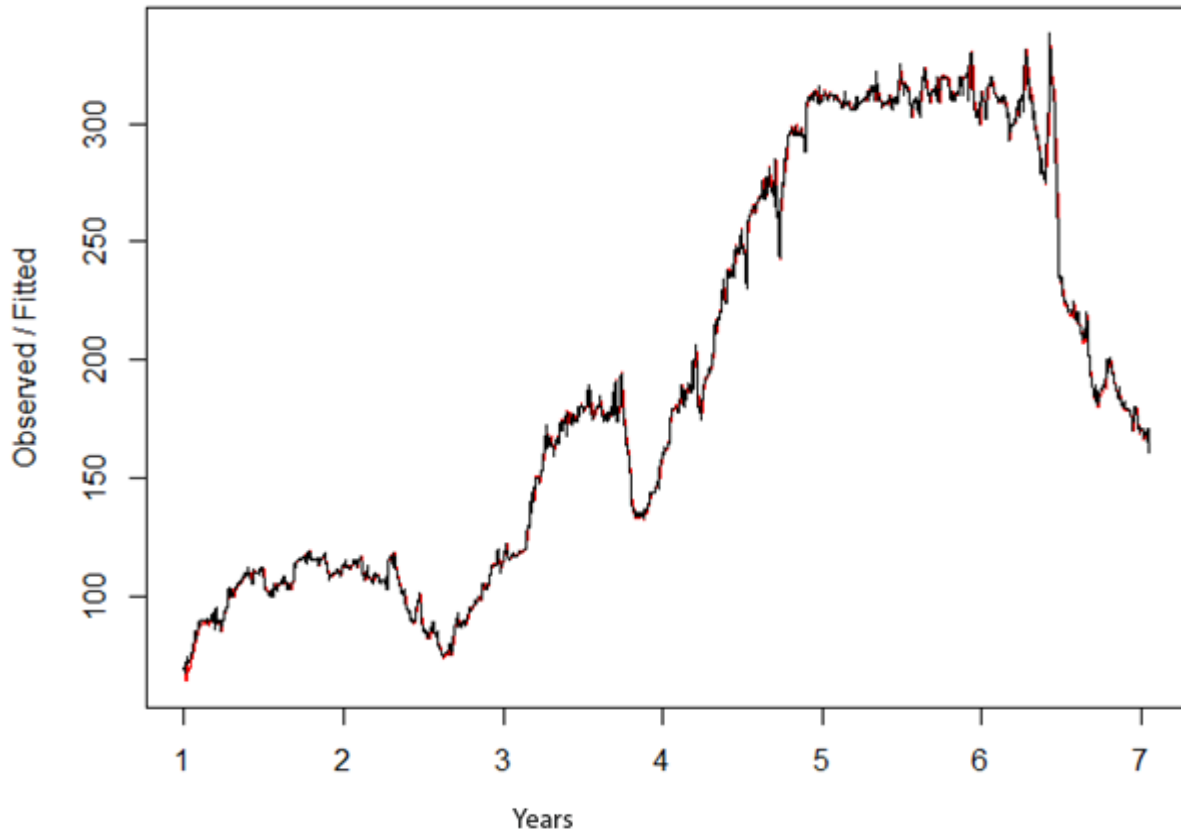
Fig 17. A Graph of Nation Media Group Stock Price with Holt Winter Smoothening, from 2014-2018

Exponential Smoothening, just like simple moving average mimics stock price movement of Nation Media Group over the 7-year period under review.

```
> fit_ES
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = ts(ticker.info[, 2], freq = 250), gamma = FALSE)

Smoothing parameters:
 alpha: 0.9209125
 beta : 0.02778579
 gamma: FALSE

Coefficients:
         [,1]
a 169.9775832
b  -0.3244122
> tickerForecast7
      Point Forecast     Lo 80     Hi 80     Lo 95     Hi 95
7.052         169.6532 165.2572 174.0492 162.9301 176.3762
7.056         169.3288 163.2759 175.3816 160.0717 178.5858
7.060         169.0043 161.5950 176.4137 157.6728 180.3359
7.064         168.6799 160.0699 177.2900 155.5120 181.8479
7.068         168.3555 158.6412 178.0698 153.4988 183.2122
7.072         168.0311 157.2781 178.7841 151.5859 184.4764
7.076         167.7067 155.9621 179.4513 149.7449 185.6685
> tickerForecast14
      Point Forecast     Lo 80     Hi 80     Lo 95     Hi 95
7.052         169.6532 165.2572 174.0492 162.9301 176.3762
7.056         169.3288 163.2759 175.3816 160.0717 178.5858
7.060         169.0043 161.5950 176.4137 157.6728 180.3359
7.064         168.6799 160.0699 177.2900 155.5120 181.8479
7.068         168.3555 158.6412 178.0698 153.4988 183.2122
7.072         168.0311 157.2781 178.7841 151.5859 184.4764
7.076         167.7067 155.9621 179.4513 149.7449 185.6685
7.080         167.3823 154.6811 180.0834 147.9575 186.8070
7.084         167.0579 153.4269 180.6889 146.2111 187.9047
7.088         166.7335 152.1934 181.2736 144.4963 188.9706
7.092         166.4090 150.9761 181.8420 142.8064 190.0117
7.096         166.0846 149.7716 182.3977 141.1360 191.0332
```

Fig 18. A Forecasting table for 7 Days forecasting and 14 days forecasting.

Holt Winters function, which is an inbuilt function in R programming language is instantiated upon getting the relevant seven year stocks data of Nation Media Group. The forecasting in done in two series, using a seven-day forecast, called **tickerForecast7** and a 14 days forecast called **tickerForecast14**

7 days and 14 days forecast works best for short term traders, speculators and trend followers who want quick gains on market turns. Trend followers specialize in making small profits on every market turn (Covel, 2007) and (Abraham, 2013).

NMG Historical Data - 7 Day Forecast, Exponential Smoothing
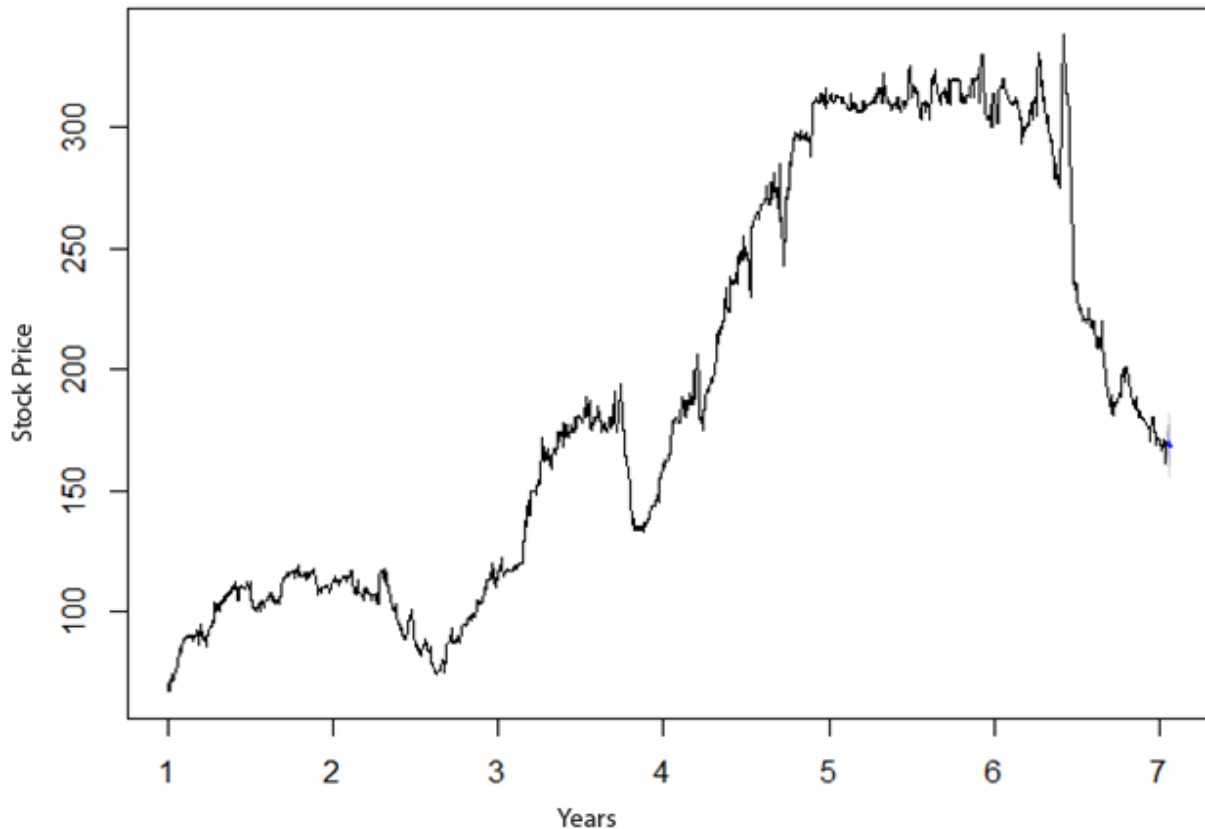
Fig 19. A Graph of Nation Media Group 7 Days Historical Stock Price Forecasting, Using Holt Winters Method.

- Based on the data fed on the Holt Winters Smoothening model, Nation Media Group 7 days stock price predicts a downward trend. As earlier mentioned by leading stock market trend followers Michael Coven and Andrew Abraham, short term traders capitalize on market fluctuations to make a profit. Markets are not steady, markets change on a daily basis and hence seven days is a long time to enable speculators make a move. Swing Traders, Day Traders and Short Term Investors have a field day making millions of money on a slight price change. Fundamental news from companies like product launches,

mergers and acquisitions, change in management cause a stock price change and hence swing traders benefit from change in stock price.

NMG Historical Data - 14 Day Forecast, Exponential Smoothing
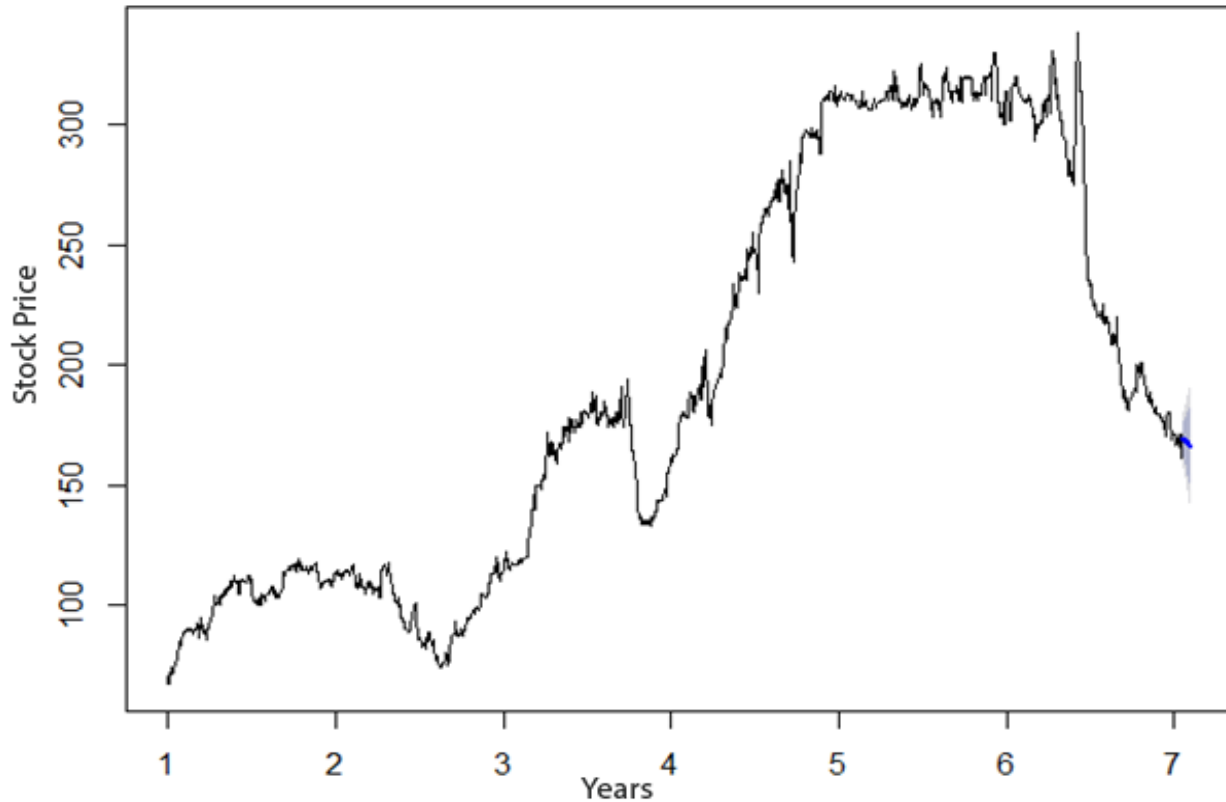
Fig 20. A Graph of Nation Media Group 14 Days Historical Stock Price Forecasting, Using Holt Winters Method.

- Based on the data fed on the Holt Winters Smoothening model, Nation Media Group 14 days stock price predicts a downward trend. Comparing the predicted values and the reality on the ground is that stock price had a downward momentum as per the Holt Winters smoothening stock predictions.

NMG Historical Data - 30 Day Forecast, Exponential Smoothing

Fig 21. A Graph of Nation Media Group showing 30 Days Historical Stock Price Forecasting, Using Holt Winters Method.

- The downward momentum on the Nation Media Group stock seems to continue based on the 30 days stock price prediction by Holt Winters Smoothening model. Comparing the system generated predictions and the actual values as per the trading data on the NSE, shows that the stock had a downward momentum as per the Holt Winters smoothening stock predictions.

NMG Historical Data - 100 Day Forecast, Exponential Smoothing

Fig 22. A Graph of Nation Media Group showing 100 Days Historical Stock Price Forecasting, Using Holt Winters Method.

- The downward momentum on the Nation Media Group stock seems to continue based on the 100 days stock price prediction by Holt Winters Smoothening model. Comparing the predicted values and the actual values as per the live trading data shows that the stock had a downward momentum as per the Holt Winters smoothening stock predictions.

4.1 WHO GAINS MOST FROM THESE PREDICTIONS?

The Stock Market have different kinds of investors: long term investors who invest for 20-50 years, medium short term investor who invest for 5 years and short term investors who invest from anywhere between one week and one two years. Holt Winters exponential smoothening model best fits short term investors as stock prices are highly volatile and highly unpredictable. They move in one direction, up trend and suddenly changes to downward trend based on company news

release or on the politics of the day or on the environment. Several factors affect the operations of a business venture. Such factors are given the acronym, PESTEL: - Politics, Economic, Social , Legal, Environmental, and Technological changes can seriously impact the operations of a company.

Nation Media Group for example has been seriously affected by technology. With the arrival of social media and citizen journalism, virtually everybody can report news incidences as it happens. This has reduced over reliance on mainstream media for breaking news.

Long Term Investors.

As earlier stated in the literature review, Long term investors are those who invest their hard earned money in a given stock for over 10 years. The short term prediction of the stock market using Holt Winter Exponential Smoothening model will not be very beneficial to them. This method works best for short term investors who are interested in stock price vitality and to gain from the price swings. Established stock investors like Warren Buffet call the daily stock price volatility as noise. Long term investors are not interested in the market noise. Their investment objectives are driven by a company's fundamentals like earnings, profits, dividends, products innovations, assets, management, acquisition and mergers and debts among others and less by daily stock market price swings.

At the Nairobi Securities Exchange, long term stock investors who put their monies in stocks like Safaricom, Equity, KCB, East African Breweries have gained significantly over the years. These companies have registered steady growth in their assets, profits, earnings, branches, production volumes and innovations.   The steady growth in share price over the last 10-20-year period of course is not without up and down price swings that affect virtually all stock market prices. All stock market share prices face price volatility (Abraham, 2013).

NMG Historical Data - 365 Day Forecast, Exponential Smoothing

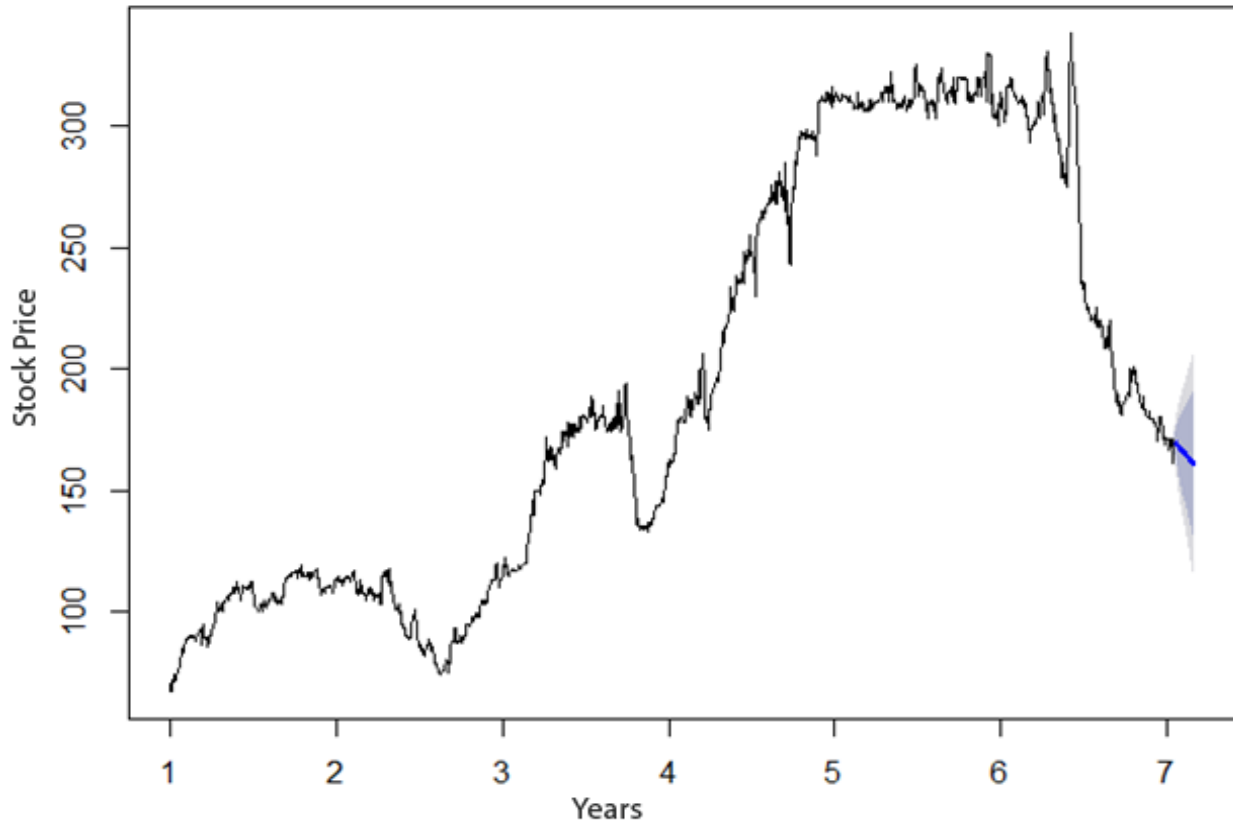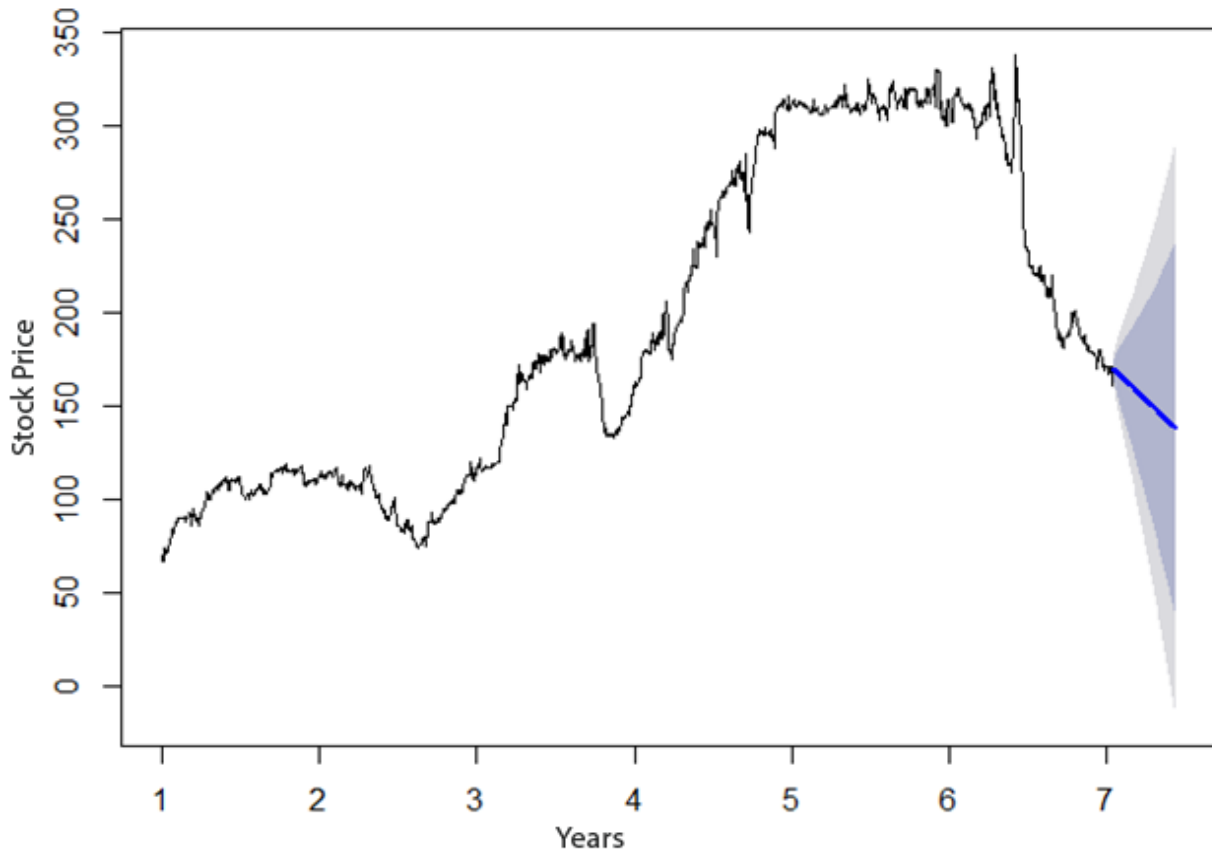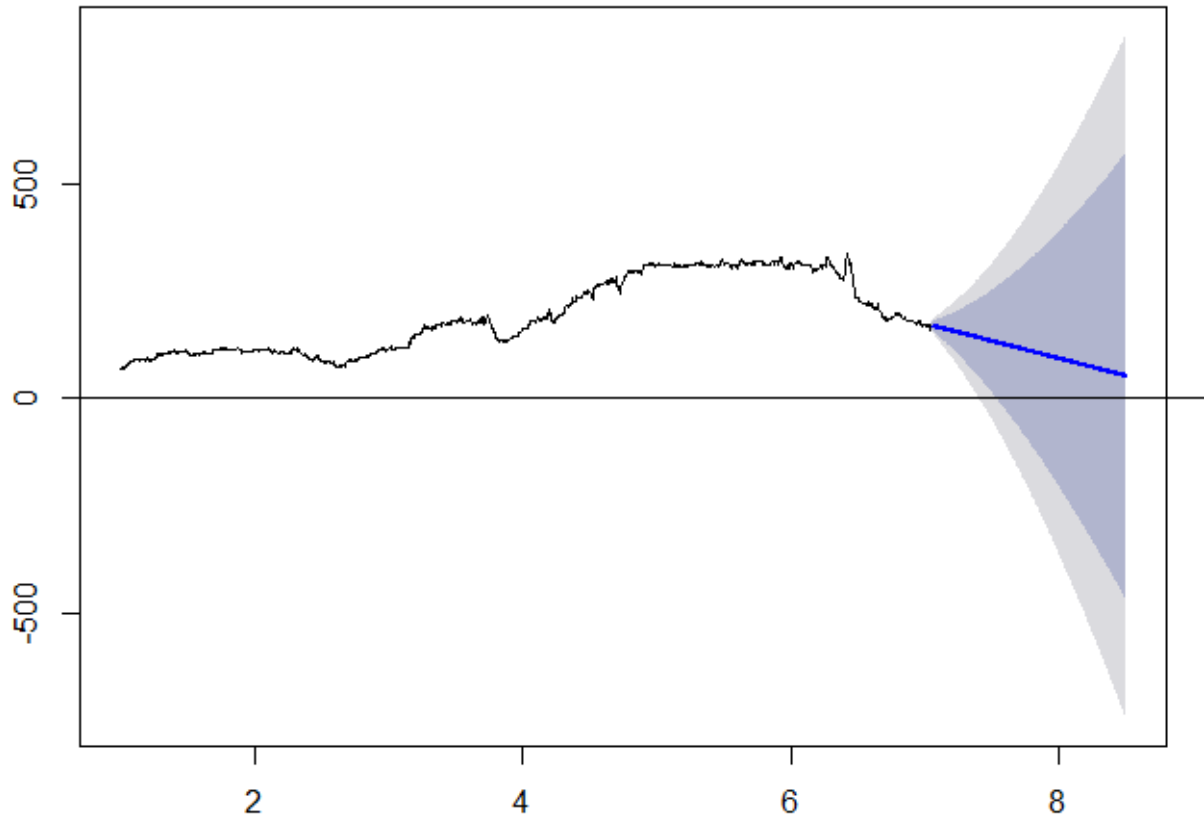*Fig 23. A Graph of Nation Media Group showing 365 Days Historical Stock Price Forecasting, Using Holt Winters Method.*

- Predicting Nation Media Group for 12 months, Holt Winters smoothening model revealed a bearish market that is bound to have a negative return on equity on those who invested their capital on this counter. Those who invested when the share price was 320/- per share will have the capital eroded by a considerable margin. Comparing the predicted values and the actual values as per the trading data on NSE, suggest that the stock had a downward momentum as per the Holt Winters smoothening stock predictions. The model seems to be 98% accurate as per the results shown.

NMG Historical Data 1825 Day Forecast, Exponential Smoothing
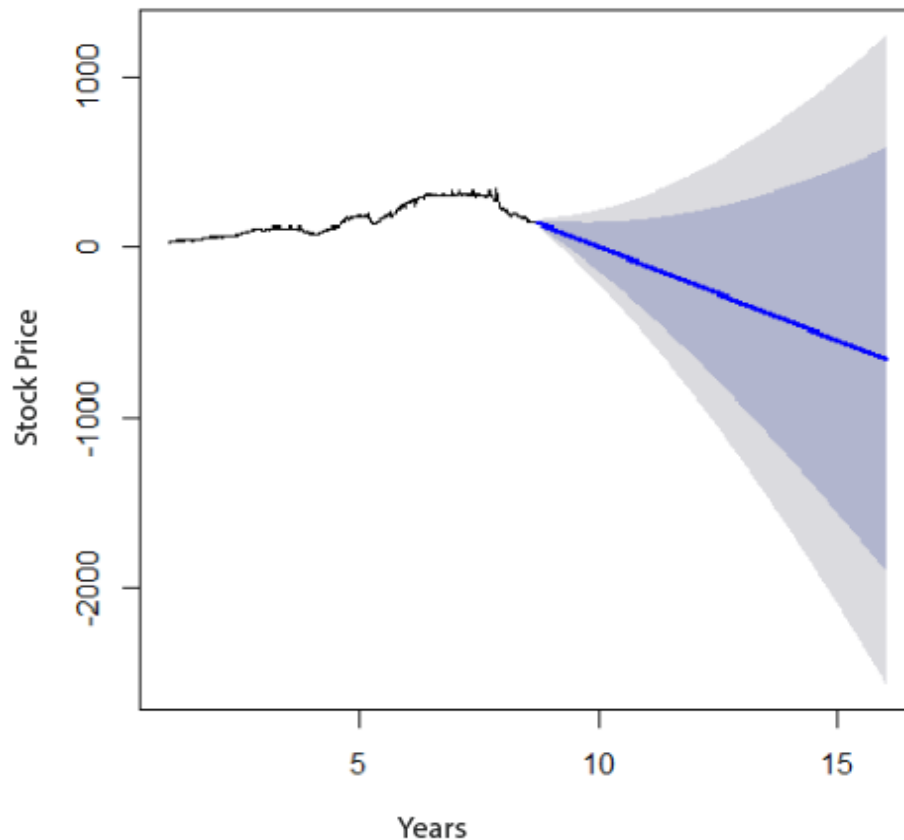
Fig 24. A Graph of Nation Media Group showing 1825 Days (5Years) Historical Stock Price Forecasting, Using Holt Winters Method.

Predicting Nation Media Group for 60 months, using Holt Winters smoothening model revealed an impossible bearish market with prices going towards the negative. This is impractical. The lowest stock prices could go to is a few cents, not even zero shillings. Case in point is Mumias sugar Company and Uchumi selling at 0.27/- and 0.35/- respectively. As the graph and the findings, the prices drops upto -600/- which is not possible. Such findings mean that the company would have filed for bankruptcy and wound up its operations long time ago.

CHAPTER 5

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.0 INTRODUCTION

This chapter presents a discussion on the summary of findings, challenges and limitations encountered during research, conclusions reached and finally recommendation on findings of the research.

5.1 DISCUSSION

This project focused on Holt Winters Exponential Smoothening model for predictive data analysis of the Nairobi Securities Exchange data over a ten year period. The bearish market that the Nairobi Securities Exchange, NSE, had a similar and corresponding results with the predictive model under the study.

Other scholars, doing a comparative study of Holt Winters Exponential Smoothing, compared with other models of stock predictions like ARIMA Model, Neural Networks, Time Series Model among others found that Holt Winters model outperformed the other tools.

As a comparative study by several authors including Ponnam and Rao et al (2016), analysed various stock predictive models such as Neural Nets, Holt winters and the Time Series linear model among other models, Holt Winters model outperformed Artificial Neural Network model and ARIMA model.

Why Holt Winters is preferred? Stock market data displays volatility, shift and seasonality, a trend that is best illustrated using Holt winters.

Why R? Inbuilt libraries and packages helps in solving time series data.

Evaluation Metrics

After the two models have been developed and tested, an accuracy or precision is checked. Mean Absolute Error is ideal to test the accuracy of the models generated. Two sets of data is provited Rtrain and Rtest. MAE measures the deviation of the predicted prices from the original prices. Small MAE are recommended, large values means a huge deviation and hence huge prediction errors.

$$MAE(f) = \frac{1}{|R_{test}|} \sum_{p_s \in R_{test}} (f(s) - p_s)$$
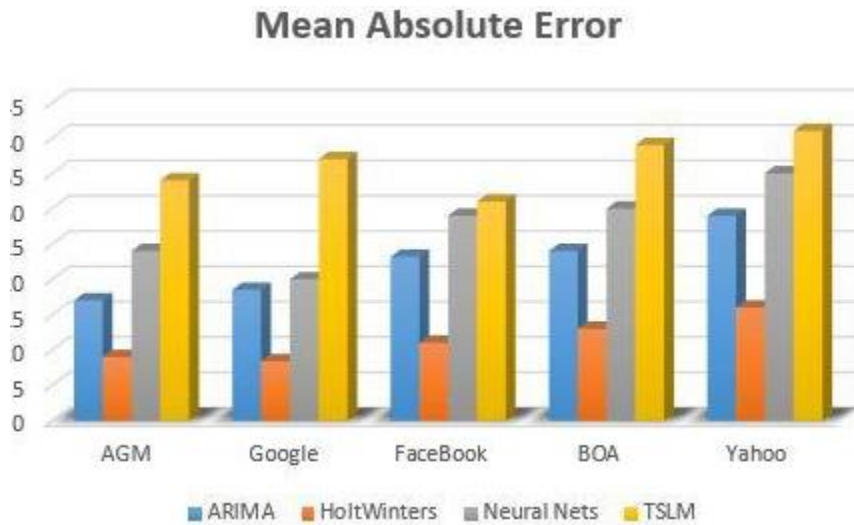
Fig 25: Histogram showing Mean Absolute Error of the various stock prediction models. Holt Winters Exponential Smoothening performs better than all the other indicators.

5.2 COMPARING HOLT WINTERS VS ARTIFICIAL NETWORKS

Holt-Winters Forecast - 365 Days
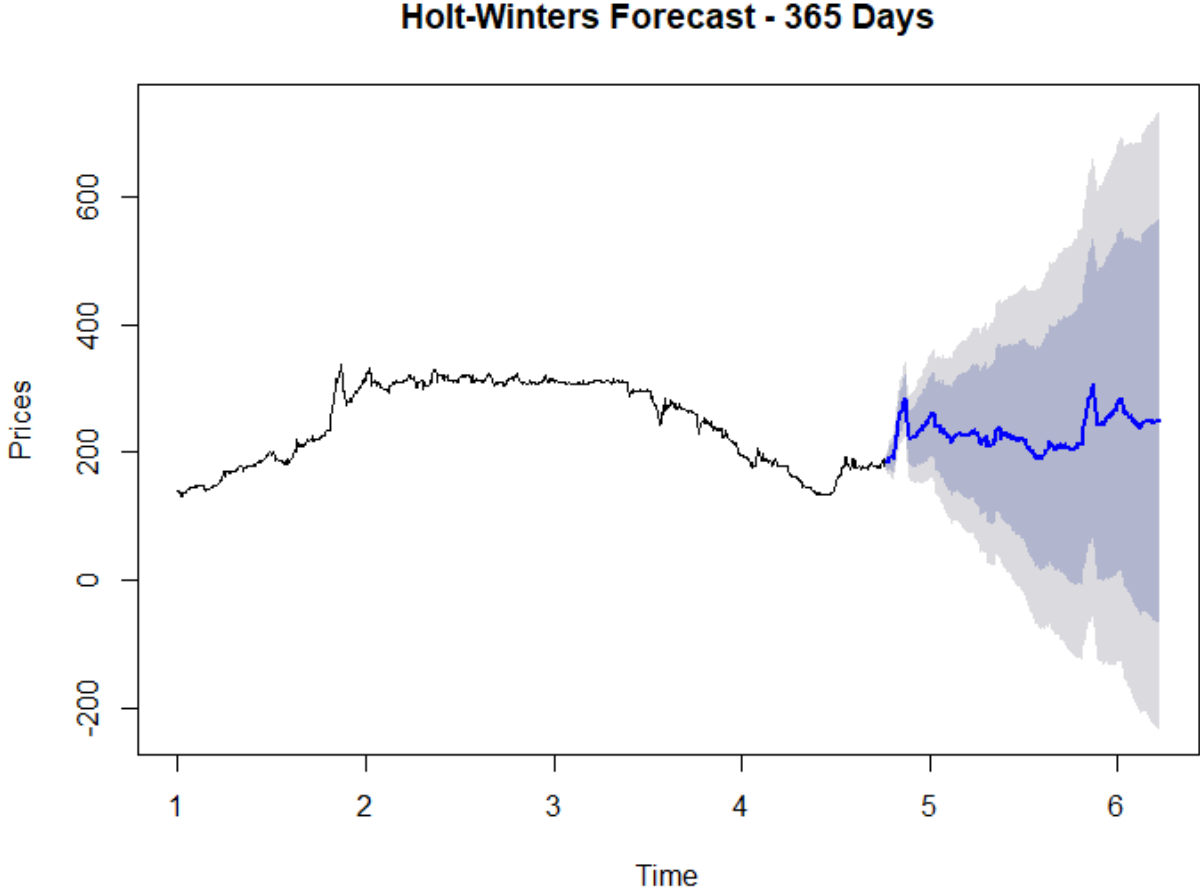
Fig 26. Graph of 365 days NMG Stock Prediction using Holt Winters Forecast.

Neural Network Forecast - NMG - 365 Days Forecast
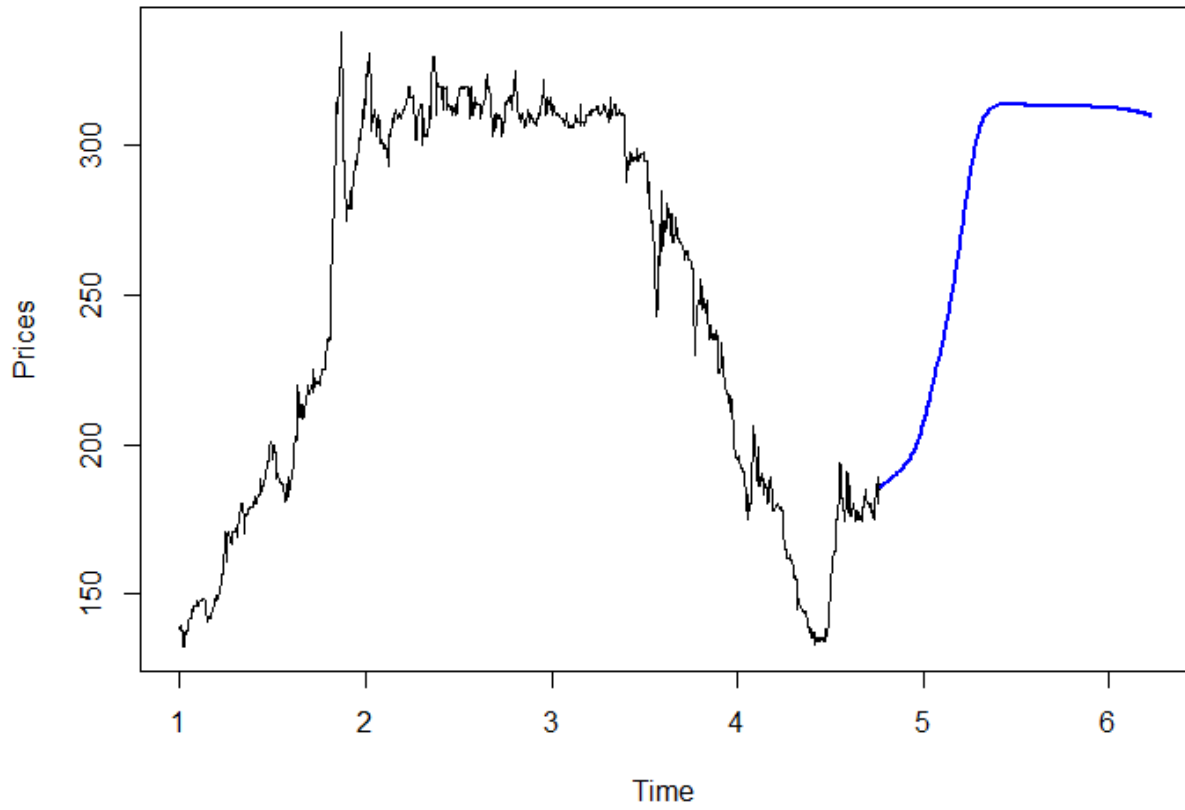
Fig 27. Graph of 365 days NMG Stock Prediction using Artificial Neural Networks.

Table: Forecasting Method Comparison Table

	Date	Train_Data	Test_Data	Holt_Winters	Neural_Network
943	2016-03-04	NA	183.0	183.1	185.4
944	2016-03-07	NA	181.0	185.8	185.8
945	2016-03-08	NA	180.0	186.3	186.0
946	2016-03-09	NA	179.0	184.9	186.3
947	2016-03-10	NA	179.0	187.1	186.5
948	2016-03-11	NA	178.0	191.8	186.7
949	2016-03-14	NA	180.0	192.9	186.9
950	2016-03-15	NA	180.0	191.5	187.1
951	2016-03-16	NA	180.0	192.1	187.3
952	2016-03-17	NA	180.0	193.8	187.5
953	2016-03-18	NA	181.0	191.1	187.6
954	2016-03-21	NA	180.0	192.1	187.8
955	2016-03-23	NA	180.0	205.1	188.0
956	2016-03-24	NA	176.0	218.4	188.2
957	2016-03-29	NA	176.0	227.5	188.4
958	2016-03-30	NA	180.0	230.3	188.6
959	2016-03-31	NA	174.0	241.7	188.8
960	2016-04-01	NA	174.0	248.2	188.9
961	2016-04-04	NA	175.0	255.8	189.1
962	2016-04-05	NA	176.0	263.6	189.4
963	2016-04-06	NA	174.0	263.7	189.6
964	2016-04-07	NA	177.0	266.4	189.9
965	2016-04-08	NA	174.0	267.6	190.1
966	2016-04-11	NA	177.0	271.6	190.4
967	2016-04-12	NA	172.0	279.8	190.6
968	2016-04-13	NA	174.0	279.0	190.9
969	2016-04-14	NA	173.0	285.6	191.1

Fig 28. Forecasting Method Comparison, Holt Winters vs Artificial Neural Networks – 2016 Data.

Fig 29. The reality. Nation Media Group Stock Prices. 2016 Prices Data

```
Table: Forecasting Method Comparison Table

|       |Date       | Train_Data| Test_Data| Holt_Winters| Neural_Network|
|1271 |2017-07-04 |        NA|    109.0|       250.6|          312.0|
|1272 |2017-07-05 |        NA|    109.0|       248.4|          312.0|
|1273 |2017-07-06 |        NA|    108.0|       248.5|          312.0|
|1274 |2017-07-07 |        NA|    109.0|       246.0|          311.9|
|1275 |2017-07-10 |        NA|    110.0|       245.3|          311.9|
|1276 |2017-07-11 |        NA|    108.0|       243.6|          311.8|
|1277 |2017-07-12 |        NA|    107.0|       243.2|          311.8|
|1278 |2017-07-13 |        NA|    109.0|       245.3|          311.8|
|1279 |2017-07-14 |        NA|    105.0|       241.2|          311.7|
|1280 |2017-07-17 |        NA|    107.0|       243.0|          311.7|
|1281 |2017-07-18 |        NA|    106.0|       236.7|          311.6|
|1282 |2017-07-19 |        NA|    107.0|       245.4|          311.6|
|1283 |2017-07-20 |        NA|    108.0|       245.3|          311.6|
|1284 |2017-07-21 |        NA|    108.0|       245.5|          311.5|
|1285 |2017-07-24 |        NA|    108.0|       246.1|          311.5|
|1286 |2017-07-25 |        NA|    108.0|       247.2|          311.4|
|1287 |2017-07-27 |        NA|    110.0|       249.0|          311.4|
|1288 |2017-07-28 |        NA|    113.0|       249.3|          311.3|
|1289 |2017-07-31 |        NA|    108.0|       250.9|          311.2|
|1290 |2017-08-01 |        NA|    108.0|       250.6|          311.2|
|1291 |2017-08-02 |        NA|    107.0|       249.5|          311.1|
|1292 |2017-08-03 |        NA|    109.0|       251.7|          311.1|
|1293 |2017-08-04 |        NA|    108.0|       248.3|          311.0|
|1294 |2017-08-07 |        NA|    109.0|       249.2|          310.9|
|1295 |2017-08-09 |        NA|    112.0|       249.4|          310.9|
|1296 |2017-08-10 |        NA|    114.0|       248.1|          310.8|
|1297 |2017-08-11 |        NA|    117.0|       247.6|          310.7|
|1298 |2017-08-14 |        NA|    117.0|       246.7|          310.7|
|1299 |2017-08-15 |        NA|    116.0|       246.6|          310.6|
|1300 |2017-08-16 |        NA|    115.0|       247.7|          310.5|
|1301 |2017-08-17 |        NA|    115.0|       247.7|          310.4|
|1302 |2017-08-18 |        NA|    114.0|       248.4|          310.4|
|1303 |2017-08-21 |        NA|    112.0|       248.3|          310.3|
|1304 |2017-08-22 |        NA|    114.0|       248.7|          310.2|
|1305 |2017-08-23 |        NA|    115.0|       249.3|          310.1|
|1306 |2017-08-24 |        NA|    114.0|       250.3|          310.0|
|1307 |2017-08-25 |        NA|    115.0|       251.6|          309.9|
```

Fig 30. Forecasting Method Comparison, Holt Winters vs Artificial Neural Networks – 2017 Data.

.

Fig 31. The reality. Nation Media Group Stock Prices. 2017 Prices Data

MEAN ABSOLUTE PERCENTAGE ERROR

The MAPE (Mean Absolute Percent Error) measures the size of the error in percentage terms. It is calculated as the average of the unsigned percentage error, as shown below:

$$\left(\frac{1}{n}\sum\frac{|Actual - Forecast|}{|Actual|}\right) * 100$$

Month	Actual	Forecast	Absolute Percent Error
1	112.3	124.7	11.0%
2	108.4	103.7	4.3%
3	148.9	116.6	21.7%
4	117.4	78.5	33.1%
MAPE			**17.6%**

Fig 32. Mean Absolute Percent Error

MAPE FOR HOLT WINTERS VS ARTIFICIAL NEURAL NETWORKS

A quick comparison between Holt Winters and Artificial Neural Networks reveal that Holt Winters does better at prediction with a lower, Mean Absolute Percent Error rate as shown below using Nation Media Group Stock Prices for the dates 2017-08-07 to 2017-08-25.

	A	B	C	D	E	F	G	H
1	NMG PRICES	HW	ANN		MAPE FOR HW			MAPE FOR ANN
2	109	249.2	310.9		128.6238532			185.2293578
3	112	249.4	310.9		122.6785714			177.5892857
4	114	248.1	310.8		117.6315789			172.6315789
5	117	247.6	310.7		111.6239316			165.5555556
6	117	246.7	310.7		110.8547009			165.5555556
7	116	246.6	310.6		112.5862069			167.7586207
8	115	247.7	310.5		115.3913043			170
9	115	247.7	310.4		115.3913043			169.9130435
10	114	248.4	310.4		117.8947368			172.2807018
11	112	248.3	310.3		121.6964286			177.0535714
12	114	248.7	310.2		118.1578947			172.1052632
13	115	249.3	310.1		116.7826087			169.6521739
14	114	250.3	310		119.5614035			171.9298246
15	115	251.6	309.9		118.7826087			169.4782609

Fig 33. Mean Absolute Percent Error outcomes for Holt Winters vs ANN. HW has lower error rate

5.3 CONCLUSION

Nairobi Securities Exchange, the Kenyan stock market is very volatile and prices can change abruptly. Prices can go up or down without warning, which to the untrained eye, is very unpredictable and almost confusing. Such confusion can lead to loss of investment money or trading capital if the prices moves against your amateur predictions.

This project has proven that to some extent, it is possible to predict the stock market movements. Stock market Gurus have been doing so for years, over the last 100 years and with technology, amateur traders can predict stock movement with great success and hence benefit like the stock gurus.

Stock Market Gurus like William O'Neil, Jesse Livermore, Michael Covel, Warren Buffet have made significant fortunes in the stock market in millions and billions of dollars. They have done it consistently over the years, decades in fact. They have made very good stock market predictions and benefited massively over the years, without incurring heavy losses (Abraham, 2013).

Holt Winters Exponential Smoothening, one of the predictive analytics employed in this project, has proven that it is indeed possible to predict the stock prices. When the market is going down, it is bound to keep doing down (especially in the bear market) hence it is good to be liquid. Be in cash. Cash is king. When the market starts going up, it is bound to keep going up for a while (the bull run). It is in the bull run when traders make money in the stock market. (Covel, 2007).

Stock market price prediction is possible using various models, apparently Holt Winters does better than all models, and more so ANNs as per our study. After all, we have successfully employed Holt Winters to help in the predictions of the stock market in this project. With such accurate prediction of the stock market, there is absolutely no reason why investors should make losses trading stocks.

5.4 FURTHER AREAS OF STUDY AND PROJECT LIMITATIONS

The Holt Winters Exponential Smoothening has one major limitation. It assumes a given direction in the movement of stocks, without taking into cognizance the fact that stock market has an upward and downward volatility. The prediction assumes that when the stocks is going up, it will continue going up and when it is going down, it will continue going down, which is not the reality. Stocks don't move up in perpetuity and they don't move down in perpetuity, till negative prices as has been the case in this project. Further research could be done to tweak the system to have the upward and downward swings that is the case in the stock market. The upward and downward volatility is driven by the speculators, who want to gain from the 10% stock capital gains and who want to dump a given stock to avoid further capital loss.

THE CODE

– Data Prediction of Nairobi Securities Exchange using Holt Winters Exponential Smoothening.

Table of Contents

##

 # 1) Working Directory

 # 2) Required Packages

 # 3) Data Import

 # 4) Forecasting

 # 5) Tables and Charts

##

1) Working Directory

##

 # Create a "Working Directory" folder

 # Place "Price_Prediction_Kenyan_Stock_Exchange.R" in the

 # Working Directory folder.

```r
# Place Kenya Stock Exchange historical data .csv files in the

# Working Directory folder.


# Go to "Session" - "Set Working Directory" - "Choose Directory..."


# Set the Working Directory to the folder holding .R & .csv files,

# for example:

setwd("C:/Msc/code")

###############################################

## 3) Required Packages

###############################################

# Uncomment the below "install.packages()" commands to

# install required R packages:


install.packages("timeSeries")

install.packages("tseries")

install.packages("xts")

install.packages("zoo")

install.packages("quantmod")

install.packages("PerformanceAnalytics")
```

```r
install.packages("forecast")

install.packages("ggplot2")

install.packages("knitr")


# Run the following code to load R packages into the library:

library(timeSeries)

library(tseries)

library(xts)

library(zoo)

library(quantmod)

library(PerformanceAnalytics)

library(forecast)

library(ggplot2)

library(knitr)


##################################################

## 4) Data Import

##################################################


# Find files in working directory

dir()
```

```r
# Input the name of the historical data file for price prediction,

# for example, "BAMB Historical Data"

# Therefore, FileName <- "BAMB Historical Data"


FileName <- "BAMB Historical Data"


# Run the following 2 lines to get ticker information

FileName2 <- paste(FileName, ".csv", sep ="")

ticker.info <- read.csv(FileName2)


# Re-formats dataset observation names

names(ticker.info) <- c("Date", "Price", "Open", "High","Low", "Vol", "Change")


ticker.info$Date <- as.Date(ticker.info$Date,format="%B %d, %Y")


##############################################

## 5) Forecasting

##############################################


# Fit Exponential Smoothing
```

```r
fit_ES <- HoltWinters(ts(ticker.info[,2], freq=250), gamma = FALSE)



# Forecast 7 Days

tickerForecast7 <- forecast(fit_ES, h=7)



# Forecast 14 Days

tickerForecast14 <- forecast(fit_ES, h=14)




# Forecast 365 Days

tickerForecast365 <- forecast(fit_ES, h=365)



# Forecast 1825 Days

tickerForecast1825 <- forecast(fit_ES, h=1825)




# Time series for Month plot, and Season plot

# ts() function converts stocks data into time series data

ticker_dataTS <- ts(ticker.info$Open, frequency=250)



#################################################
```

```
#################################################

##

## 6) Tables and Charts

##

#################################################

#################################################


# Creates a table of ticker information.

# See below RStudio console for print out

kable(ticker.info[1:10,1:7])


# Plot of Historical Data

ggplot(ticker.info, aes(Date, Price)) +

  geom_line(aes()) +

  ggtitle(paste(FileName, "- Stock Price Plot")) +

  theme(plot.title = element_text(lineheight=.7, face="bold"))


# Exponential Smoothing plot

plot(fit_ES,

    main = paste(FileName,

            "- Holt-Winters Filtering, Exponential Smoothing"))
```

```r
# Exponential Smoothing Table

fit_ES

# Predicted Stock Prices for the next 7 Days

tickerForecast7

# Predicted Stock Prices for the next 14 Days

tickerForecast14

# Predicted Stock Prices for the next 1825 Days

tickerForecast1825

# 7 Day Forecast Plot

plot(tickerForecast7,

    main = paste(FileName,

        "- 7 Day Forecast, Exponential Smoothing"))


# 14 Day Forecast Plot

plot(tickerForecast14,

    main = paste(FileName,

            "- 14 Day Forecast, Exponential Smoothing"))

# 365 Day Forecast Plot

plot(tickerForecast365,

    main = paste(FileName,
```

```
                    "- 365 Day Forecast, Exponential Smoothing"))

# 1825 Day Forecast Plot

plot(tickerForecast1825,

    main = paste(FileName,

                "- 1825 Days Forecast, Exponential Smoothing"))

# Seasonal Decomposition - Month plot

monthplot(ticker_dataTS, col="green",

        main="Month Plot, Seasonal Adjustment")



# Seasonal Decomposition - Season plot

seasonplot(ticker_dataTS,

        main="Season Plot, Seasonal Adjustment")
```

END.

BIBLIOGRAPHY

Abraham, A. (2013), The Trend Following Bible, How Professional Traders Compound Wealth and Manage Risk. Published by John Wiley & Sons.

Awajan AM, Ismail MT, AL Wadi S (2018), Improving forecasting accuracy for stock market data using EMD-HW bagging. https://doi.org/10.1371/journal.pone.0199582

Al-Radaideh Q., Assaf A.,A., Alnagi E., 2013, *Predicting Stock Prices using Data Mining Techniques*.

Bakar, N., B., A., and Rosbi, S., (2016), Reliability of Exponential Smoothening for forecasting Islamic share price to oil and gas sector in Malaysian Stock Exchange. International Academic Research Journal of Business and Technology 2(2) 2016 Page 38-44

Barber, B., M., Odean, T., 2011, The Behavior of Individual Investors.

Baresa, S., Bogdan, S., Ivanovic, Z., 2013, Strategy of stock valuation by fundamental analysis. UTMS, Journal of Economics.

Boden, C., Karnstedt, M., Fernandez, M., Markl, V. 2013. Large-scale social-media analytics on stratosphere. *Proceedings of the 22nd international conference on World Wide Web companion*, pp. 257-260.

Covel, W., M., (2007), Trend Following. How Great Traders Make Millions in Up or Down Market. Published by FT Press.

Daily Nation, June, 11, 2018, https://www.nation.co.ke/business/Mumias-Sugar-worth-all-time-low-of-Sh1bn/996-4607206-l6ep5m/index.html

Daily Nation, April, 30, 2015, https://www.nation.co.ke/lifestyle/money/Woes-aside-Britam-stock-is-a-long-term-buy/435440-2701892-q6dkdhz/index.html

Ding, X., Zhang, Y., Liu, T., Duan, J., 2015, Deep Learning for Event-Driven Stock Prediction. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

Dunne, M. , 2015, Stocks Market Prediction,

Dai, Y., and Zang, Y., 2013 Machine Learning in stock price trend forecasting.

Dodds, Peter. (2009) Journal of Happiness July, doi: 10.1007/s10902-009-9150-9

Schumaker, R. P & Chen, H. (2009) ACM Trans. Inf. Syst. 27, 12:1–12:19.

Dzikevičius, A., Šaranda, S., (2010). Smoothing techniques for market fluctuation signals. Journal, Business: Theory and Practice Issn 1648-0627 print / Issn 1822-4202 online

Fenzl, T., Brudermann, T., Malik, C., Pelzmann, L., 2013, *A mass psychological perspective on financial markets.* European Scientific Journal September 2013 edition vol.9, No.25 ISSN: 1857 – 7881

Fu, T., Chen, S., Wei, C., 2008, Hong Kong, Stock Index Forecasting.

Gilbert, E & Karahalios, K. (2010) Widespread worry and the stock market.

Goel, S., K., Poovathingal, B., Kumari, N., 2016, *Applications of Neural Networks in Stock Market Prediction -An Approach Based Analysis.* International Research Journal of Engineering and Technology (IRJET)   Volume: 03 Issue: 05 | May-2016

Gutierrez, J., Anzelde, T., Gobenceaux, G, 2013, Risk and Reward: The Effect of Big Data on Financial Services.

Hand, D J 2009, Mining the past to determine the future : Problems and possibilities. *International Journal of Forecasting,* vol 25 [2 July2015]

Hagstrom, G., R., 2005, *The Warren Buffet Way,* John Wiley & Sons, Inc., Hoboken, New Jersey

Hansen, K., B., 2017, *Crowds and Speculation*, A study of crowd phenomena in the U.S. financial markets 1890 to 1940

Kalogirou, S. A (2006), Artificial neural networks in energy applications in buildings, *International Journal of Low-Carbon Technologies*, Volume 1, Issue 3, 1 July 2006, Pages 201–216, https://doi.org/10.1093/ijlct/1.3.201

Kannan K., S., Sekar, P., S., Sathik, M., M. and Arumugam, P., 2010, Financial Stock Market Forecast using Data Mining Techniques, *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2010, Vol 1, IMECS 2010, March, 17-19, 2010,

Kombo, D. K. and Tromp, D. L. A.(2006) *Proposal and Thesis Writing : An Introduction.* Paulines Publications, Nairobi. Pg 83.

Kumar, R., 2011, Research Methodology, a step-by-step guide for beginners. 3rd Edition. Sage Publications Ltd

Leemakdej, A., 2016, Stock Price Prediction by Deep Learning.

Liu, Y, Huang, X, An, A, & Yu, X. (2007) ARSA: a sentiment-aware model for predicting sales performance using blogs. (ACM, New York,NY, USA), pp. 607–614.

Metcalfe G (2018) The mathematics of market timing.
PLoS ONE 13(7): e0200561. https://doi.org/10.1371/journal.pone.0200561

Minervini, M., 2013, Trade like a stock market Wizard. How to achieve super performance in stocks in any market. McGrow Hill Education, New York.

Milosevic, N., 2016 , Equity Forecast: Predicting long term stock price movement using machine learning. University of Manchester, UK.

Mishne, G & Glance, N. (2006). Predicting Movie Sales from Blogger Sentiment. AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs

Newbern, H., J., 2006, Holt Winters Forecasting; A study of practical application for health care managers.

O'Neil, J. W., 2009, *Making money in stocks. A winning system in good times or bad*. 4th ed, McGrawHill, New York.

Paul, J., and Moynihan, B., 2007, What I learned losing a million dollars. Infrared Press, Nashville, Tennessee.

Palepu G. K, Healy P., M, & Peek E, 2008, *Business Analysis and Valuation IFRS Edition : Texts and Cases*, 2nd edn, Cengage Learning, Hamphire, United Kingdom.

Patel, B. M., Yalamalle, S. R., 2014, *Stock Price Prediction Using Artificial Neural Network* , International Journal of Innovative Research in Science, Engineering and Technology [Vol. 3, Issue 6].

Pinheiro, L., D., S., and Dras, M.,. 2017. Stock Market Prediction with Deep Learning: A Character-based Neural Language Model for Event-based Trading. In *Proceedings of Australasian Language Technology Association Workshop*, pages 6_15.

Ponnam , L., T., Rao, V., S., Gujju, B., R., Nallagulla, S., N., (2016), *A Comparative Study on Techniques used for Prediction of Stock Market.* 6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India

Pring, M., J., 2018, *Investment Psychology Explained*, Classic Strategies to Beat the Markets. John Wiley and Sons inc.

Rusu, C., Rusu, V., (2003), Forecasting methods and stock market analysis.

Shah, V., H., Mohri, M., 2007 *Foundations of Machine Learning*, Spring.

Shah, V., V., Mirani, S., J., Nanavati, Y., V., Narayanan V, Pereira, S., I., 2016, *Stock Market Prediction using Neural Networks,* International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-6 Issue-1, March 2016.

Shah, V., H, 2007, *Machine Learning Techniques for Stock Prediction*

Suresh A., S, 2013, *A study of fundamental and technical analysis*, International Journal of Marketing, Financial Services and Management Research, Vol 2, No. 5 [May, 2013].

Singh, K K, Dimri, P & Rastogi, K N 2014, The Implications of Big Data in Indian Stock Market, *International Journal of Computer Applications*, vol. 99, no. 5. [2 July 2015].

Singh, K., K., Dimri, P., Rastogi K., N., 2014, *The Implications of Big Data in Indian Stock Market,* International Journal of Computer Applications, *[Volume 99– No.5]*

Tabassum, P., Majumdar, M., A., 2018, *Stock Price Forecasting using Bayesian Network*, Brac University.

Wanjawa, B., W., Muchemi, L., 2013, *ANN Model to Predict Stock Prices at Stock Exchange Markets,* School of Computing and Informatics, University of Nairobi.

Yakovyna, V., Bachkai, O., 2018, The Comparison of Holt – Winters and Box – Jenkins Methods for Software Failures Prediction

Zhang, W., Lau, R., Li C., 2012, Big Data Stream analytics for correlated stock price movement prediction, City University of Hong Kong.