ISEN 660- Quantitative Risk Analysis

Risk Analysis of Stock Market

Under the Guidance of

Dr. Nancy Currie-Gregg

Project Report by-

Shubahm Zope (930001594)

Vaishnavi Patki (430000405)

Vraj Thakkar (329006917)

TABLE OF CONTENTS

ABSTRACT

Facebook has been in talks because of several crises and this has impacted its stock prices. This study focuses on the development of ARIMA model based on the last five years of its data, for predicting its stock price. Results from the model has been successful in predicting stock price with optimal amount of risk involved. In addition, by visualizing data and interpreting curve fit model, it discusses major setbacks that company has faced during the period which has a lasting social impact.

INTRODUCTION

Financial planning plays a vital role in everyone's life. It helps to plan your future and take better control of your goals by setting realistic goals and by evaluating alternatives, taking effective measures. Financial planning includes managing earnings, investments, returns and taxes. Most investments include stocks, bonds, mutual funds, exchange traded funds and life secure or retirement plans. Directly or indirectly we have a linkage with stock market. The risks associated with the stock market has an active role to play in our daily life. The different risks involved in stock market are liquidity risk, headline risk, rating risk, obsolescence risk, detection risk, legislative risk, inflationary risk and interest rate and risk model risk. The liquidity risk and headline risk were chosen in the project because these two risks have a major impact on the price movement of the stock. The term liquidity in broad prospect means the amount of cash flow in the market. In stock market the term refers to how easily the stock can be sold and bought. The numbers of buyers and sellers in the market and the price difference at the top of the book indicates the liquidity of the stock. Headline risk involves any news story leading to price drop of the stock. Significant news has huge impact on the stock performance in the market. The primary of goal of any investor is to buy at low price and sell at high price. The price movement in the market are subjected to two types of analysis fundamental analysis and technical analysis. In this report technical analysis of stock is done using the major movement of price associated with volume. The Facebook stock was selected for analysis as this stock is popular and there is relevant data available online regarding news and events of the stock.

Financial forecasting or particularly forecasting the price of stock are one of the hottest fields of research, due to the high stake of investments in stock market and the kind of fascinating benefits it has to offer. The major challenge is to forecast the price movements in the market. The historical data was used in Bayesian approach to get posterior probabilities of price movement. These probabilities are indication of level of risk for the investor. Bayesian approach will provide probabilities for particular cases based on historical data, so time series forecasting was selected to predict daily price movement .

LITERATURE SURVEY

3.1. Bayesian Modeling

When there are number of events occurring, which are a result of interconnection between some common variables, Bayesian methods are useful in studying such systems. Such variables have probabilistic dependencies among them. Cooper et al.[1] explained such probabilistic dependencies of variables in detail in their research paper titled 'A Bayesian Method for the Induction of Probabilistic Networks from data. They explained the occurrence of 10 cases based on 3 variables x1, x2, x3 based on the inter-relationship between these variables. For example, conditional probabilities like P(x3=True|x1=True) should be calculated depending on the relationships.

Conditional probabilities are calculated using Bayes' Theorem as: Evidence (E) tells about occurrence of Hypothesis(H) and more about if that is True or false as per following:

Figure 3.1 : Casual relation between hypothesis and evidence[2]

Bayes' Theorem is given as:

$$P(H \mid E) = \frac{P(E \mid H) \times \bar{P}(H)}{P(E)}$$

[2]

When there are many different types of relationship among variables like causal relation, common cause relation or common effect relation, Bayesian networks are created. They contain variables in the form of nodes connected by acyclic arcs. Joint probabilities in such Bayesian networks is calculated as :

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid \pi_i),$$

[1]

Where $X_1, ..X_n$ are variables and π_i indicate their respective prior probability.

Bayesian Networks can also be used to access the performance of investment risks. There are many qualitative and quantitative factors which affect the financial investments like stocks, bonds, mutual funds ad options. Bayesian networks can be one of the best tools to accommodate such relationships. Chang et al. [] in their research article 'Market analysis and Trading Strategies with Bayesian Networks' have considered variables like Interest rate(LIBOR), Consumer Price Index(CPI), Unemployment rate(UR), Money Supply(MS), Housing Start(HS) and Implied Volatility Index (VIX) for equity index S&P500 [3]. This Bayesian model is shown below as an example.
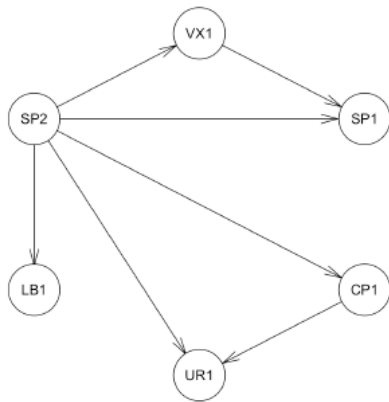
Similarly, for a stock, dividends, GDP, unemployment rate, economic or political events etc. can be the affecting variables.

Figure 3.2: S&P Bayesian Model [3]

3.2. Application of Statistical and Soft computing Techniques to Financial Forecasting [4]

As the fundamental theory behind investment in stock market is different, the major reliability for investment was heavily dependent on human expertise and justification. In this process there was a huge noise in transfer of information and as well miscommunication of data. Due to which there was huge volatility in price and volume in the market. To make stock market as a comparative safe investment method various models are developed for accurate prediction based on various statistical and soft computing techniques.

One of such statistical technique is Neural Network. Neural networks are a selection of algorithms modelled after the human brain to recognize patterns and interpret data based on trained patterns. Neural networks are mainly used for prediction of non-linear functions with better accuracy. Neural networks are less sensitive to error term assumptions and are capable to handling noise or chaos in the data. The better we can predict the risk the better we can try to prevent and pre-empt.
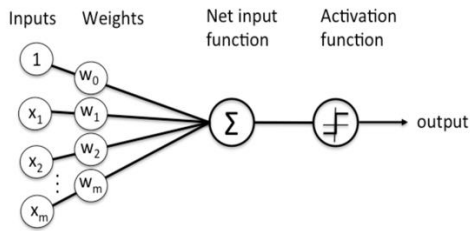
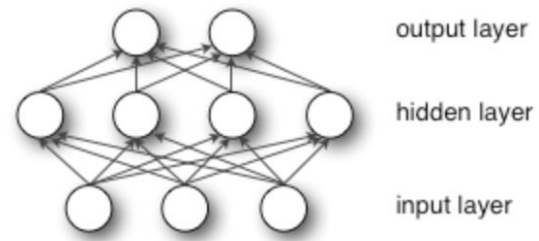Figure 3.3: Nodes of Neural Network [4]   Figure 3.4: Stages of Neural Networks [4]

Neural network model inputs different input values from the data and assigns weightage to each value. The input layer includes data obtained after multiplication of input values with their respective weightages. After that the values are passed to the activation function which reduces all values in range from 0-1in hidden layer and again weightage for each value is computed and multiplied with output from hidden layer. The process continues till the error in expected and trained sample obtained is the least.

There are two types of Neural Networks

1. Convolutional Neural Networks(CNN), 2. Artificial Neural Network(ANN)

The Artificial Neural Network was selected as the data from stock market is unsupervised. Models of ANN have gained popularity in past few years due to inherent capabilities to approximate any nonlinear functions with high prediction accuracy. In ANN one of the statistical techniques which was employed was Auto Regressive Integrated Moving Average (ARIMA). There are different types of time series forecasting models but there is no systematic approach or models available to accumulate variable frequency behaviour in time series forecasting. Time series forecasting involves all variables that are dependent on time are feed to a function. The data is trained with respect to time and processed to predict, classify, describe and transform to new data.

3.3. ARIMA

The ARIMA model is a combination of AR (Auto regressive),I (Integrating) and MA (Moving average models). Auto regressive model is that uses the dependent relationship between an observation and other lagged observations.

Integrating model use the concept of differencing to keep the variables of model stationary. Moving Average model uses the dependent relationship between an observation and residual error of lagged observations.

A standard notation for ARIMA model is ARIMA (p, d, q) where the parameters are substituted with integer values.
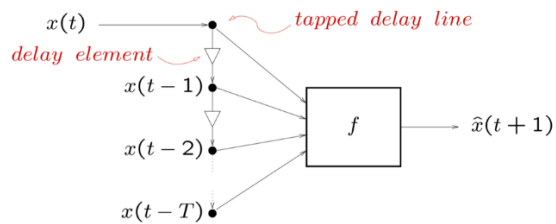
- p = Lag order of the observations

- d = No of differences in the raw model i.e. the degree of differencing of the model

- q = Moving average order

Figure 3.5:Trasnfer of data from nodes to function[5]

The model takes raw input and for a lag period calculates the relationship between the current observation and lagged observations and then passed it to differencing model to eliminate dependency on time further processing the data of observation to reduce the moving average residual error. Thus, the model takes into consideration both time dependency as well as the residual error of the observations.

EXPLORATION OF DATA SET

4.1 Exploration of Data set:

The data set was obtained from Investing.com for the Facebook stock. The data was available in daily basis. It had 4 prices Open, Close, High and Low respectively, date of trade and volume traded for the particular date. The Open price indicates the average trading price in the preopen session. The Close price indicates the average trading price during the post close session. The High price indicates the highest traded price for the particular day. The Low price indicates the lowest traded price for the particular day. The Volume indicates the total volume i.e the positions opened and positions settled in the whole day. The analysis would pe partial if it is performed on any one of the prices, so average price is calculated and the analysis is done using average price.

Average price=(Low + Open + High + Low)/4

Assumptions:

- Single stock is considered for the analysis to reduce the complexity of the project.

- Effect of events in other stock leading to price movement in the stock are not taken into consideration.

- Effect of sector on the stock price movement are not considered.

- Major events like very low liquidity, empty price book and resignation of Board member are not accounted.

- The data is reduced to 5 years to reduce the scope of the project.

4.2 Understanding of QQ plot

QQ plot is essentially a plot in-between quantile of theoretical Normal Distribution and fitted curve data. While plotting QQ plot we assume that our data theoretically follows Normal Distribution. To interpret the data distribution, we need to observe that how far does data deviates from the center line that shows the theoretical Normal Quantiles. If the data points coincide with centerline, we can interpret that it follows a perfect normal distribution. Otherwise, larger the deviation from line it shows how far is data from normal distribution.

Figure 4.1: Normal Q-Q plot of Facebook data set
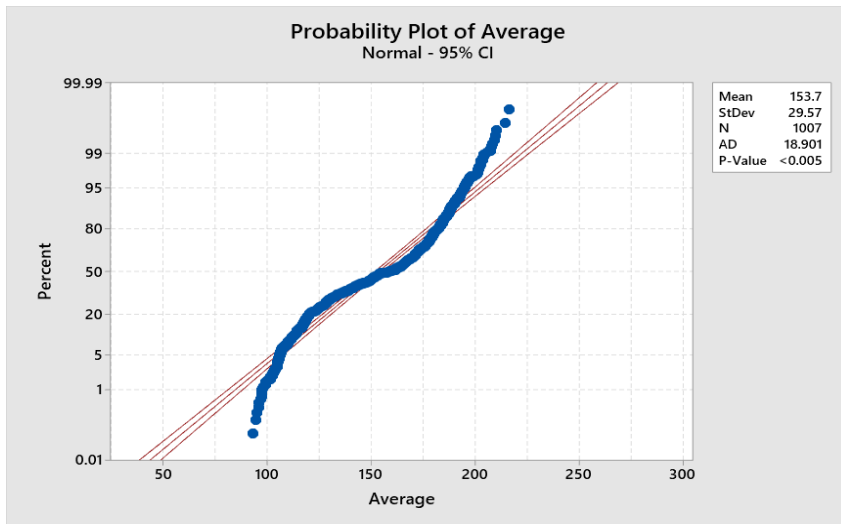
Here the normal QQ plot shows the p value is less than 0.005, which implies we can follow the normal distribution with certain tweaks, but the AD value is larger than that of Weibull Distribution. Hence, we will do our calculations based on Weibull Distribution with selected specific parameters.

Another way to interpret the distribution is by looking at Anderson-Darling goodness of fitness value.

4.3 Goodness of fit tests

Goodness of fit tests compares data of random sample with theoretical probability distribution function. There are several types of tests such as Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Chi-Squared (X^2). Among all the above described tests we have chosen to use AD test as its more sensitive than the other described tests. It essentially gives more weight to tails of the curve, which is required for comparing Stock Prices data.

It essentially measures the difference in area between the ideal curve and the fitted curve, that is plotted using data taken into consideration. Hence, lesser the AD value implies better fit of data. Essentially, AD curve is an extension of KS test where it gives more weight to tail of the curve. Hence it has more advantage for measuring more sensitive fluctuation in data as compared to KS.

10

Definition		The Anderson-Darling test is defined as:
	H_0:	The data follow a specified distribution.
	H_a:	The data do not follow the specified distribution
	Test Statistic:	The Anderson-Darling test statistic is defined as
		$A^2 = -N - S$

$$S = \sum_{i=1}^{N} \frac{(2i-1)}{N} \left[\ln F(Y_i) + \ln\left(1 - F(Y_{N+1-i})\right)\right]$$

F is the cumulative distribution function of the specified distribution. Note that the Y_i are the *ordered* data.

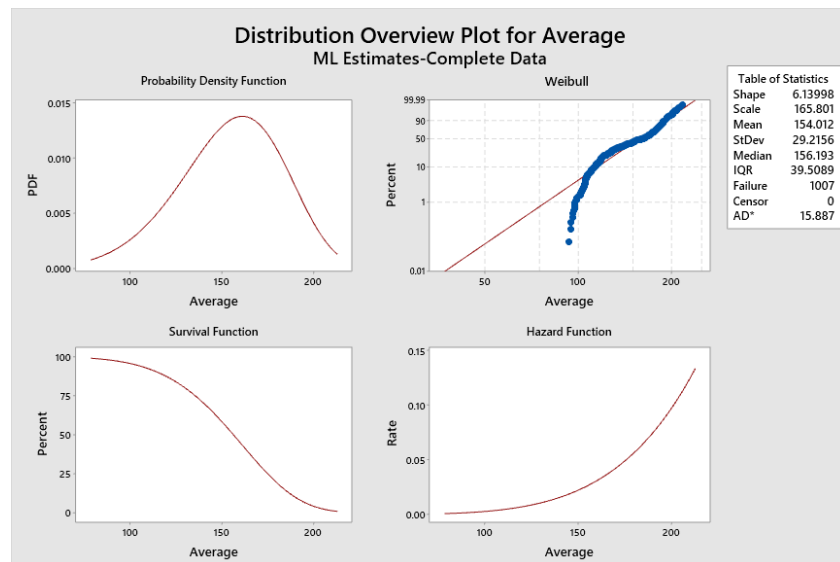Figure 4.2: Definition of Anderson-Darling test[6]

Figure 4.3: Weibull fit of Facebook data set

From here we can observe that AD value while fitting Weibull distribution is 15.867 which is 16% lesser than the value of Normal fit distribution.

While the Weibull Distribution parameters that are selected for curve fit is as follows:

Shape parameter – 6.139, Scale parameter – 165.801

4.4 Interpreting Major Events from the dataset

Moving Average Chart essentially measures the consecutive difference of Average of prices of each day. The consecutive difference allows us to measure the change in prices rather than actual prices which are essential from detecting the anomalies in behavior of trends.

Procedure of Plotting Moving Average Chart

	C1	C2
	Average	Moving Average
1	103.208	*
2	105.044	-1.8360
3	106.922	-1.8780
4	106.712	0.2100
5	107.194	-0.4820

Figure 4.4: Moving Average

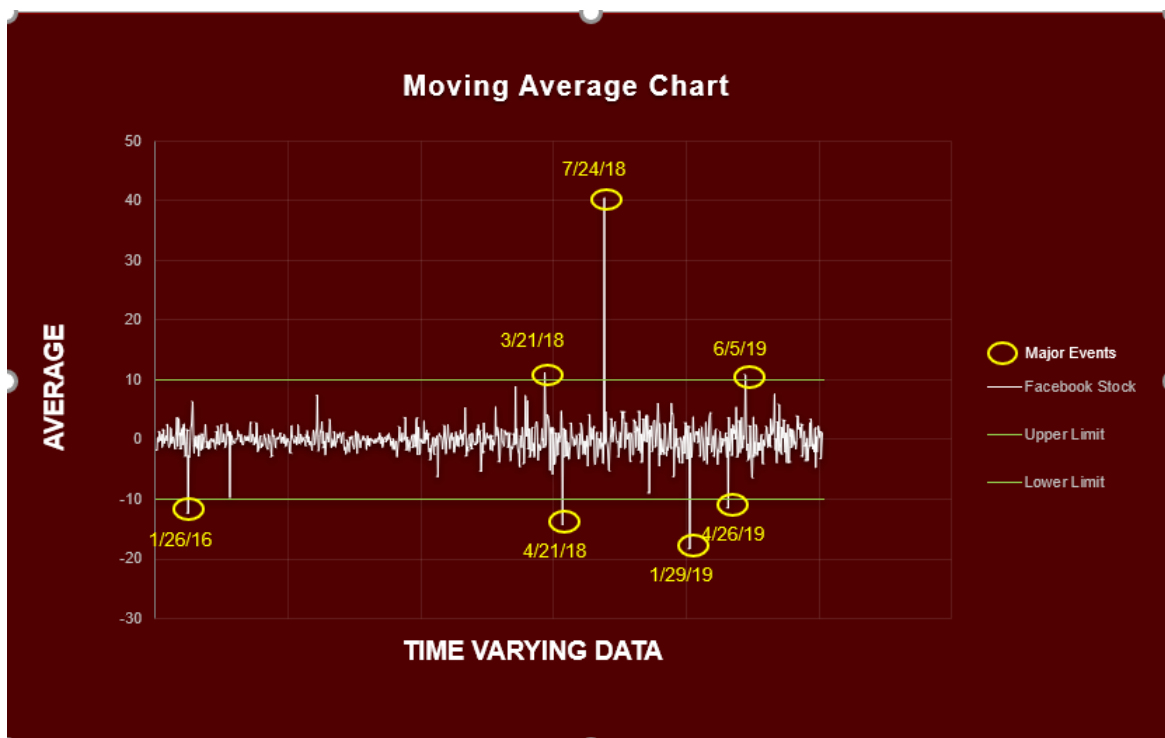$C_{22} = C_{12} - C_{11}$

$C_{23} = C_{13} - C_{12}$

And so on…

Figure 4.4: Moving Average Plot and Detection of Spikes (Major Events)

To describe events its essential to know the amount of stocks that are traded on that day and the trend that stock prices follows. This describes the kind of change i.e. the relative change in values is in positive direction or in negative direction. Hence, we plotted the trend with the amount of volume that is traded at that day, because the total value of stocks is determined by

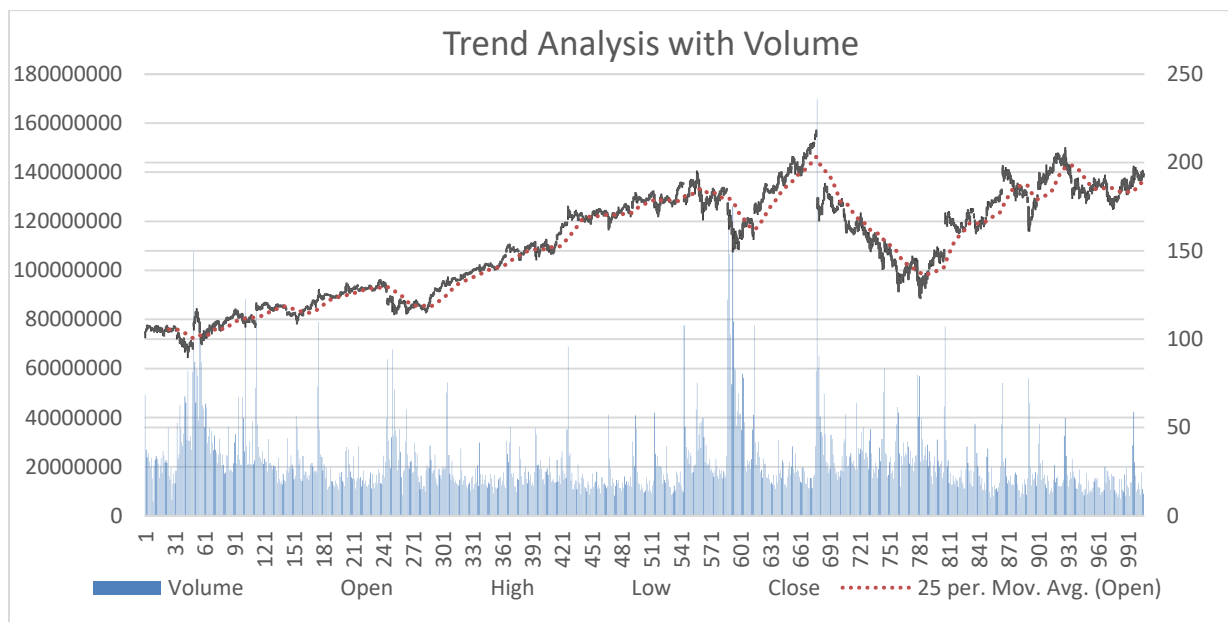Total Value of Stocks = Volume of Stocks * Average Stock Price

Figure 4.5: Trend Analysis of Facebook stock data with volume

Open, High, and Low prices are plotted on the same vertical line as that of closing pricing of the stock. The spikes in blue shows volume of stocks that are traded during that day.

Hence, we can interpret major events and happenings during the tenure of 5 years at Facebook i.e. from 2014 -2019 by looking to both the curves i.e. moving average plot and Trend Analysis with Volume. Here, we can observe total of five (5) major events happened during last five years of tenure of Facebook.

Date	Event Description
1/26/16	January Effect
3/21/18	Cambridge Analytica and Data Breach Scandal
4/21/18	CEO Mark Zuckerberg faces Congress
7/24/18	Major Event Stock Price Drop by 20% (Loss of $120 billion in Market Capitalization)
6/5/19	CEO was aware of Privacy Breach

Event 1: January Effect, 26th January 2016

January Effect is a kind of trend observed since last few decades in Bull Market where the prices start rising by the end of the month of January. In addition, its observed that, the kind of movement in first few days of the year determines the behavior of market for the whole year. For example, if the market performed very well and prices of the stocks are raised then this occurrence will be expected throughout the year and at the end of the year and people will expect to have higher returns. During the end of the fiscal year i.e. in December investors look at their portfolio and try to settle their losses if any in the last month of fiscal year. For instance, if an investor is facing loss having stock A in his portfolio then he tries to get rid of the stock so that he can file the losses in the billing of the fiscal year. Also, investors seek for the liquidity by getting rid of the stock with loses so that they can invest in the start of January in new stocks which are having low prices, since those stocks are expected to do well over the period of time.

Event 2: Cambridge Analytica and Data Breach Scandal

It came to media and to the people that Facebook was unable to protect its user's data and Cambridge Analytics, a political consulting firm has harvested 87 million of Facebook user's data, because of the loophole in Facebook's API. Its alleged that this has helped in Russia to spread fake news and also helped Trump administration to plan their political campaign based on the psychological merit of people, which he eventually won.

Event 3: CEO Mark Zuckerberg faces Congress

After this news became public people started losing faith in Facebook and its CEO Mark Zuckerberg has to face Congress to explain them what happened, and the measures company is taking to prevent this kind of blunders to happen with public data. This has reported in the decrease in stock prices of Facebook by around 18% by the end of week i.e. from 185$ to 152$. And, this has a prolonged effect of Facebook stocks as it required a lot of time to recover from the drop[8].

Event 4: Major Event in the history of Stock market, Facebook lost 20 Billion $ of the market capitalization in a day

After Facebook slightly regained its position in market, a major thing happened in the history of market and Facebook lost $120 billion of market valuation in a single day! The price of stock was around 20% lesser than the opening price of the stock. The reason for this kind of major setback can be termed as a cumulative effect of the ups and downs that company faced before few months. Even after company for generating profits, it was like company's progress was stalled and this thing made investors to lose their faith in the tech behemoth. This kind of downfall also resulted in the loss of faith in other tech companies in valley too.

Event 5: CEO was aware of Privacy Breach

During this period, it was rumored that CEO's mailed got leaked and he was aware about the data leaks and data breach that's happening at their company. And, Facebook is intentionally selling its customers data for political, psychological and influential marketing.

METHODS

5.1 Indicators of stock market

Bayesian analysis of stock market data of Facebook was performed. Firstly, major indicators of stock market were studied. There are 4 major types of stock market indicators:

1. Trend: It is the line which signifies the direction of the stock market. The concept of Bullish market and Bear Market can be visualized here. Uptrend line signifies Bullish market while downtrend line signifies Bear Market. Example: Moving Average

2. Volume : These are the units of the stocks that are being traded. Higher number indicates that there are more traders in the market. Example: On-Balance Volume

3. Volatility : It is the rate of change of price change. It is the measure of dispersion and signifies the stability of the market. Volatility brings risk of investing in sock markets.

4. Momentum : It shows how strong the trend is. It the difference between today's closing price and closing price N days ago. It also tells how many days the trend either positive or negative has lasted.

   prices. Examples: Rate of Change, relative Strength Index, Parabolic SAR

The Facebook stock data contained 4 prices viz., opening price, closing price, highest price and lowest price for that particular day. Mean price was considered for further analysis.

It is important to understand the importance considering price as well as volume while considering stock markets. High volume implies large number of participants in trading. Average trading volume increases when stock is in uptrend and vice versa. If price is increasing but volume is decreasing, this situation implies the lack of interest among traders. Decreasing price might be coupled with increasing volume. If there any negative new or fraud news about company, then stocks of that company might get affected in this way.

Facebook's shares are recently affected by major scandals and other incidents. This stock price data would be suitable for studying market behavior. One of the major events that hugely affected Facebook's stock prices is Facebook Cambridge Analytica Scandal.

5.2 Bayesian Model

The related importance two stock market variables i.e. average price and volume is studied using Bayesian tree. A is the event where average price increases and A' is the event where average price decreases. B is the Hypothesis of volume getting increased and B' is the hypothesis of volume getting decreased.

Prior probabilities of corresponding events are obtained from data by counting the days using FOR-loop where the price was higher than previous day and price was lower than previous day. Similarly, prior probability of hypothesis was calculated by counting days where volume traded was higher than previous day and it was lower than previous day. 4 posterior probabilities are calculated as per Bayesian tree method.

5.3 Artificial Neural network for time series forecasting of Facebook stock prices
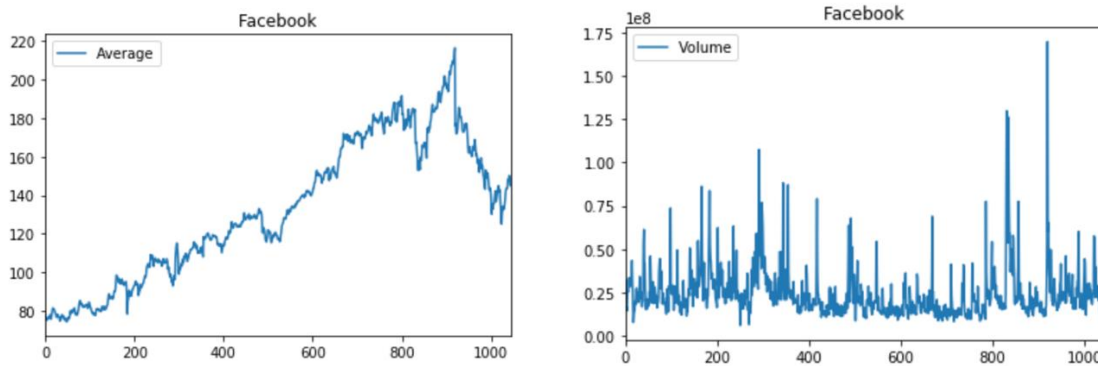

Figure: 5.1 indicates the average price variation of the Facebook stock from Nov-2014 to Jan 2019. The fig 2 indicates the volume variation of the Facebook stock from Nov-2014 to Jan 2019.

The span from Nov 2014 to May 2017 shows that there is significant increase in the stock price with constant variation in the volume. The points in graph where there is more than 10 percent of change in price with volume of approximately 25 million or greater indicates occurrence of events in news leading to drop in trust in the stock. The major drop in whole time zone is of 66% which occurs from May 2017 to Jan 2019.

17
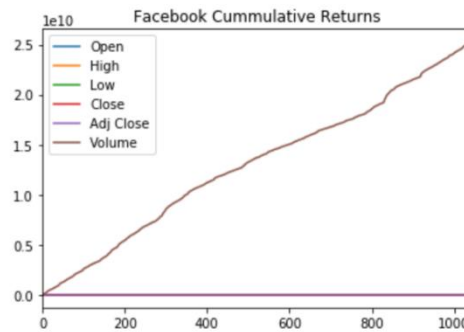
The Figure 5.2 indicates the cumulative returns of the Facebook stock from Nov-2017 to Jan 2019. The graph shows the performance of stock in the span of 5 years irrespective of events prior to Nov-2017. The positive slope indicates that the stock performance has been excellent fundamentally apart from major events happening in the span. Thus, we can specify that the stock can be selected for investment. The historical data indicates a return of 90% in the 5 years.
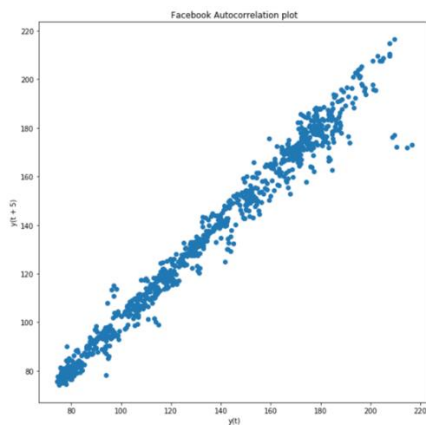
The Auto correlation plot indicates the relation of the lagged variables with the current observations. The straight line indicates that the lagged variables obtained from the Auto regressive method follow same relation followed by the average price.
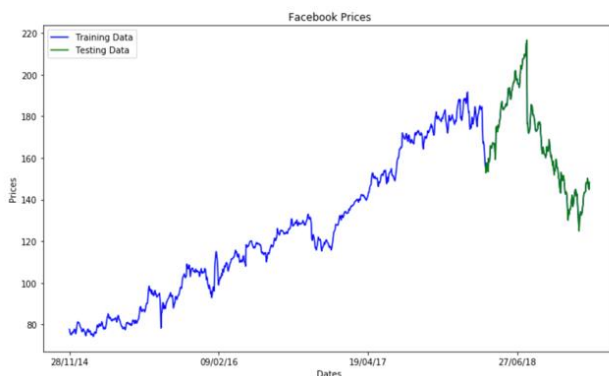
Figure 5.3:Auto correlation plot
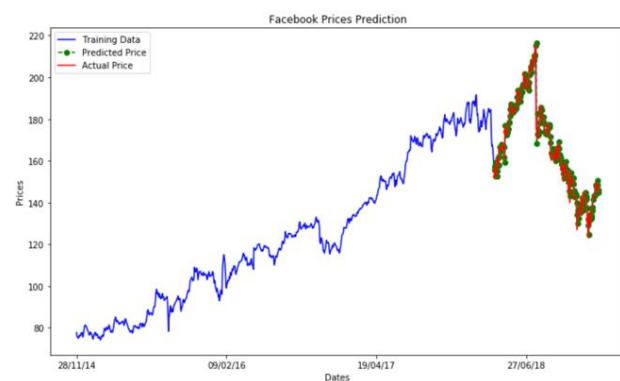

Figure 5.4: Split of average price

Figure 5.5. Prediction and comparison of Average price

To perform analysis of price movement we have plotted the average price vs time. The data is split such that the constant rise is in the training set of the ARIMA model and the volatile section is in the testing set. There is huge unpredictability in stock market, so the split makes sure that we are considering the worst possibility in the market that is we have invested money when the market was providing good returns and there is sudden volatility in the market due to news leak. After the data split, the data is trained with a lag of 5 days and predictions are made on the test data. The green points in figure 5.6 indicates the predicted price and the red points indicate the actual price.
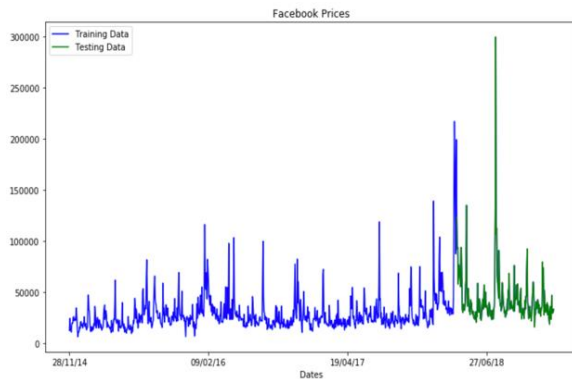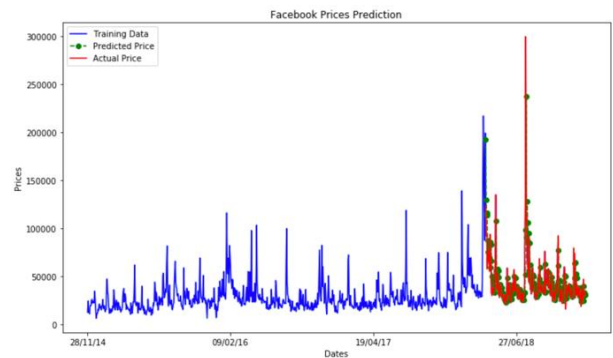

Figure 5.6: Split of Price x Volume   Figure 5.7 :Prediction and comparison of Price x Volume

To perform analysis of price movement we have plotted the PricexVolume vs time. The data is split such that the constant rise is in the training set of the ARIMA model and the volatile section is in the testing set. After the data split, the data is trained with a lag of 5 days and predictions are made on the test data. The green points in fig 5.6 indicates the predicted price and the red points indicate the actual price.The distance of green points and red points is far in the Price x Volume plot. Thus signifies that volume has a vital role in stock price prediction.

RESULTS AND DISCUSSION

6.1 Bayesian Network of Facebook-Cambridge Analytica Scandal
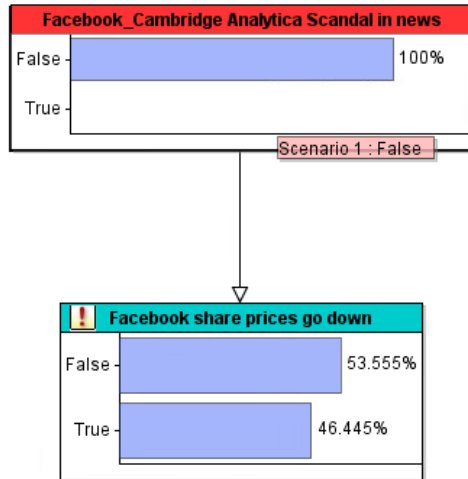
1.When this scandal was not in news   2. When Scandal was in news

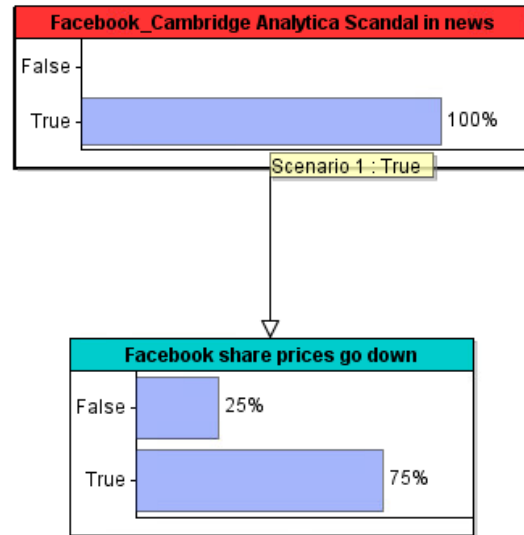Figure 6.1: Risk Map when the event is False   Figure 6.2: Risk Map when the event is False

The above two risk graphs show that the average probability of Facebook share price decreasing was 46.445% before this scandal was revealed. Then the major impact of this event was for 10 days which later gave rise to other events. For those 30 days, the probability of Facebook's share price decreasing was 75%. This indicates that such days where there is any negative news about the company, risk is much higher.

6.2 Bayesian Tree

Prior probabilities

Evidence: P(A)=0.5353, P(A')=0.4647, Hypothesis: P(B)= 0.4597, P(B')=0.5402

Conditional probabilities

	B	B'
A	P(A|B)=0.2482	P(A|B')=0.5312
A'	P(A'|B)=0.7518	P(A'|B')=0.4688

Bayesian Tree :
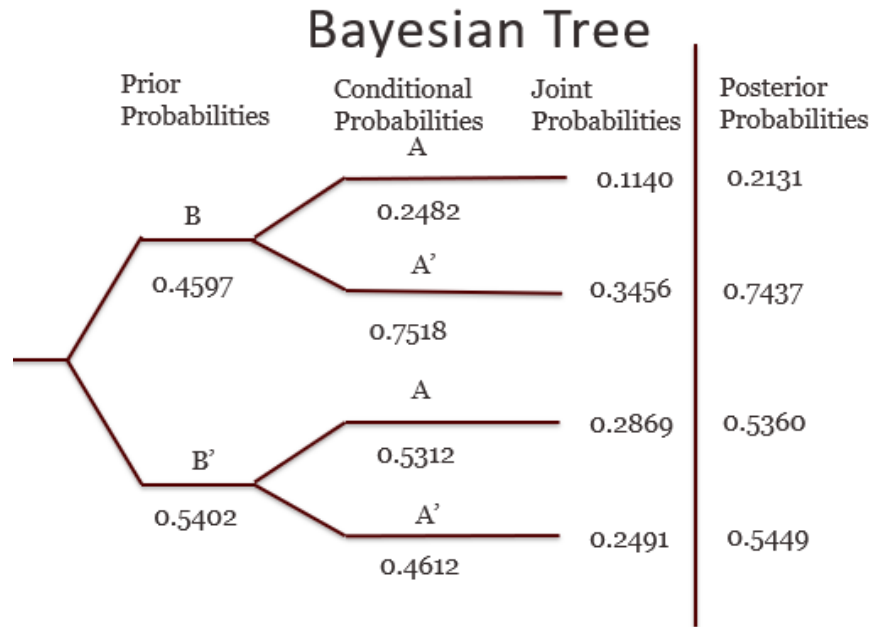

Figure 6.3: Bayesian Tree of volume movement given price

Posterior probability:

1. P(B|A) =P(B)*P(A|B)/P(A)=0.4597*0.2482/0.5353=0.2131

    This probability value explains the situation of the market where volume increases given price also increases.

    For the given dataset of Facebook Stock prices, overall it is observed that this kind of stable market situation

    has low probability of occurrence.

    This condition is suitable for Long Term Investors where they seek for the gain over long period.

2. P(B'|A) =P(B')*P(A|B')/P(A)=0.5402*0.5312/0.5353=0.5360

    This probability explains one type of volatility where price is increasing but volume is decreasing. Probability

    of such market situation for the given Facebook data is 53.60%. This is a concerning figure as this situation

    shows lack of interest among traders.

3. P(B|A') =P(B)*P(A'|B)/P(A')=0.4597*0.7518/0.4647=0.7437

    This probability explains the other type of volatility in the market where volume increases and price

    decreases. 74.37% of the times this situation has happened in past 4 years. Short Term investors want to sell

    their stock at such prices. However, there are a smaller number of buyers in the market.

21

4. P(B'|A') = P(B')*P(A'|B')/P(A')=0.5402*0.4688/0.4647=0.5449

This situation where stock volume decreases given price decreases, has happened 54.49% in last five years of Facebook stock prices. This situation is not suitable for investing. This is concerning figure as it exceeds 50%.

Event based predictions:

Above situations can help investor analyze the Facebook stock market situation based on the portfolio and term of investment. The increase in price with volume is a positive indicator of investment in the stock, and negative indicator of squaring off position in vice versa case. The opposite direction of price and volume indicates volatility of price in the market.

6.3 Time-Series forecasting using ARIMA

The term 'Long Term' and 'Short Term' are defined for holding period of 5years and holding period of 1 year respectively. The ARIMA model is fit for both the data of Long term and data of Short term.

1. **Accuracy of Model fitted for 5 years (Long Term)**

Symmetric mean absolute percentage accuracy(Price): 85.094%

Symmetric mean absolute percentage accuracy(Price x Volume): 61.249%

2. **Accuracy of Model fitted for 1 year (Short Term)**

Symmetric mean absolute percentage accuracy(Price): 93.586%

Symmetric mean absolute percentage accuracy(Price x Volume): 68.676%

The prediction accuracy for long term is less than the prediction accuracy of the short term.

This indicates that long term investments are subjected to more fundamental aspect than that of quantitative aspect. The probability of more events happening in long term is more than in short term. More the events happening higher the volatility of the stock so there is reduction in prediction accuracy.

The accuracy of Price x Volume fitted model is lower than the accuracy of Average Price fitted model. This indicates that volume has a significant role in price movement. The higher the variance of volume the lower the accuracy of prediction.

CONCLUSION

Risk Matrix:

Probability of correct price prediction	Minor	Moderate	Major
Likely	ARIMA Model(Short Term)		
Possible		ARIMA Model(Long Term)	
Unlikely			Bayesian Model

Type of Risk in Investment

The Bayesian model uses historical data to calculate the posterior probability to find safe zone for the investor to invest money. So, the probability of correct price prediction based on single iteration is very unlikely. The ARIMA long term model predicts prices based on 5 years of data so fundamental aspects in long term are considered in quantitative analysis. So ARIMA long term model has moderate risk. The ARIMA short term model predicts price for 1 year which has fewer fundamental aspects compared to 5 year. So ARIMA short model has minor risk.

REFERENCES

1. Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. Machine Learning, 9(1-3), 309–347.

2. Fenton, N. & Neil, M. (2019). Risk Assessment and Decision Analysis with Bayesian Networks, 2nd Ed., Boca Raton, FL: CRC Press.

3. Chang, C.& Tian, Z.(2015). Market Analysis and Trading strategies with Bayesian Networks. In Proceedings of the18[th] International Conference on information Fusion, pages 1922-1929.IEEE.

4. https://skymind.ai/wiki/neural-network

5. Neural Networks for Time Series Prediction 15-486/782: Artificial Neural Networks Fall 2006

6. Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical*

7. *Assohttps://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htmciation*, 69, pp. 730-737

8. https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram

9. https://www.investopedia.com/ask/answers/what-is-finance/

10.https://www.investopedia.com/articles/stocks/11/risks-every-stock-faces.asp

 11.https://pdfs.semanticscholar.org/bd56/830aee425f864658103e9b987a38763abce6.pdf

12. https://www.ersj.eu/dmdocuments/16_3_A_p13.pdf

13. https://www.sharetipsinfo.com/economy-stock-market.html

14.https://www0.gsb.columbia.edu/faculty/cjones/papers/2018.08.31%20US%20Equity%20Market%20Data%20Paper.pdf

15. https://www.econ.berkeley.edu/sites/default/files/Selene%20Yue%20Xu.pdf