



Stock Market Prediction Using Machine Learning(ML)Algorithms

M Umer Ghani^a, M Awais^a and Muhammad Muzammul^a

^aDepartment of Software Engineering, Government College University Faisalabad
umerghani92@gmail.com, muhammadawais@gcuf.edu.pk, m.muzammul275@gcuf.edu.pk

KEYWORD

Stock Market Prediction;
Machine Learning(ML);
Algorithms;
Linear Regression;
Exponential Smoothing;
Time Series Forecasting

ABSTRACT

Stocks are possibly the most popular financial instrument invented for building wealth and are the centerpiece of any investment portfolio. The advances in trading technology has opened up the markets so that nowadays nearly anybody can own stocks. From last few decades, there seen explosive increase in the average person's interest for stock market. In a financially explosive market, as the stock market, it is important to have a very accurate prediction of a future trend. Because of the financial crisis and recording profits, it is compulsory to have a secure prediction of the values of the stocks. Predicting a non-linear signal requires progressive algorithms of machine learning with help of Artificial Intelligence (AI).

In our research, we are going to use Machine Learning Algorithm specially focus on Linear Regression (LR), Three month Moving Average(3MMA), Exponential Smoothing (ES) and Time Series Forecasting using MS Excel as best statistical tool for graph and tabular representation of prediction results. We obtained data from Yahoo Finance for Amazon (AMZN) stock, AAPL stock and GOOGLE stock after implementation LR we successfully predicted stock market trend for next month and also measured accuracy according to measurements.

1. Introduction

Stock market is trading platform where different investors sale and purchase shares according to stock availability. Stock market ups and downs effects the profit of stakeholders. If market prices going up with available stock then stakeholders get profit with their purchased stocks. In other case, if market going down with available stock prices then stakeholders have to face losses. Buyers buy stocks with low prices and sell stocks at high prices and try to get huge profit. Similarly, sellers sell their products at high prices for profit purpose (Tae Kyun Lee *et al*, 2019). Stock market (SM) work as trusty platform among sellers and buyers. Advances in Artificial Intelligence (AI) supporting a lot in each field of life with its intelligent features. Several algorithms present in AI that performing their role in future predictions (Eunsuk Chong *et al*, 2017). Machine learning (ML) is a field of artificial intelligence (AI) that can be considered as we train machines with data and analysis future with test data. Machines can be trained on the basis of some standard that are called algorithms. Stock market predictions can be great beneficial to businessman. SMP provide future trend of stock prices on the basis of previous history (Bruno *et al*,2019). If stakeholders get future predictions then investment can lead him toward profit. Predictions can be 50% correct and 50% wrong as it is risk of business. Risks facing capability in business filed can



lead toward success. In any field of life, we take risks for success. Similarly, we rely on ML predictions about future prices of stock. In this chapter we would like to explain these ML algorithms with the help of their working methodologies and examples (Thomas Fischer *et al*, 2018). Before working on actual problem SMP, complete understanding of ML algorithms role in prediction is also necessary. That's why in this chapter we explained complete working scenario and problem definition (Bruno *et al*, 2019). Several Machine learning algorithms can be used for stock market prediction but in this research we used few algorithms like Linear regression (LR), Three Months Moving Average (3MMA) and Exponential Smoothing and if we further consider many other algorithms can also be used for Stock Market Prediction(SMP) (Bruno *et al*, 2019).

1.1. Data Analysis (DA) and Stock Prediction (SP)

Data analysis (DA) in machine learning (ML) is a process of applying technical skills (ML Algorithms) on historical data to obtain statistical as well as tabular results about predictions. It also considered as technical process of data illustration and evaluation. Two authors (Shamoo *et al*,2012; Resnik *et al* ,2013) explained about DA, according to their theory DA is process of distinguishing signals for decision making with statistical fluctuation of results. DA also included collection as well as analyzing process, it can be iterative according to problem statement. Several statistical techniques implemented in DA. Data scientists find patterns of entire data with special observations. Several types of quantitative as well as qualitative approaches as content analysis, history analysis, sentimental analysis and bibliographic analysis involved in DA (Tae Kyun Lee *et al*, 2019). DA study formulate predictions on the basis of historical data that can be present in form of notes, files, documents, tables, audio or video tapes. Accurate analysis of different research findings can lead to valid knowledge discovery. Inaccurate statistical presentation of data destroys the research findings of any scientist and guide wrong destinations to readers (Shepard, 2002) and public perceptions about research also influenced negatively (Bruno *et al*,2019). Integrity and accuracy are two main parts of statistical data analysis. In this research we used several data resources in form of datasets, research theories and financial resources of data presentation. Yahoo Finance, Quandl, Kaggle and several other similar platforms provided data that is used in stock market predictions. We obtained data from these platforms for different stock exchange companies and after applying ML algorithms we presented stocks predictions results statistically (Ben Moews *et al*, 2019).

1.2. Introduction to stock market

Stock market can be defined as combined platform of several markets and exchangers with regular process of buying and selling goods that shares issued publically (Comparison analysis performed at public platform). At this platform several situational financial performed for formal exchange process under defined rules and regulations (Hiransha *et al*, 2018). Multiple stocks trading places can be available at different place in a country where transections on stocks can be performed. There are two different terms used in stock market definition as stock exchange and stock market with consideration of formal trade assets. We can define stock market as, if someone is part of stock market will be considered as stock market trader, as he buy or sell his shares on one or different stocks (Ben Moews *et al*, 2019). Several stock exchange markets available at national as well as international level including Pakistan and world level biggest stock market i.e. New York Stock Exchange (NYSE), Chicago (CBOE) both of these are considered as national stock markets of USA (Tae Kyun Lee *et al*, 2019).

1.3. Introduction to stock market trend analysis

Trend is considered as direction of stock movement that is totally based on stock market ups and downs. Continues movement of stock in any direction upward or downward for specified duration or time period can be considered as trend. In stock market prediction trend analysis at current stage support a lot in future trend prediction (Thomas Fischer *et al*, 2018). Trend growing analysis for continues interval of time can be considered as future grow or continues down in trending market share prices can be supportive for future predictions as down. Stock market prediction always based on big amount of historical data analysis. Similarly trends also based on big data analysis results. Prediction about future trends in any stock market cannot be considered as 100% accurate. Trends presence in share market place provide predictions about trends in stock market. Gaining

in profit always based on trends, if investors move according to trend directions, they can be succeeded in their trade marketing (Bruno *et al*, 2019).

1.4. Problem Statement

Everyone want to be rich in his life with low efforts and great advantages. Similarly, we want to look in our future with inner most desire as we do not want to take risks or we want to decrease risk factor. Stock market is a place where selling and purchasing can provide future aims of life (Kang Zhang *et al*, 2019). Now the question is that how we can get advantages from stock market? Or what are the steps that can give us stocks market predictions before taking yourself in risk zoon (Yue-gang Song *et al*, 2018). How Artificial Intelligence with Machine learning algorithms can be supportive for future market trend predictions?

1.5. Stock Market Prediction (SMP)

If stock market trend predicted then we can avoid wastage of money. SMP is a process of predicting future on the base of past data. Prediction decreases the risk level to investors and increases the confidence level for investment. If they predicted goals before reach then they can avoid loss of money. All these consideration work as SMP. On the basis of historical data trends, we guess future trend that is called SMP(Changju Lee *et al*, 2019).

2. Material and Methods

AI is an intelligent field of latest research which is providing great help in solution of real time existing problems. AI supporting in each field of life as we use it in data processing in ATM machines, Bank accounts, Airways services, Reservation, X-RAYS, Auto door opening, recognition-based devices and weather forecasting. In other word we can say AI made our life easier and we can predict future. Earning money is major issue to face society and rick people considered as role models everywhere (Ben Moews *et al*, 2019). Stock market investments polices supported a lot in this process(Yue-gang Song *et al*, 2018). Stock market provide great benefits to intelligent investors but foolish investor's do not get much advantages form stock market. Why I said foolish? There is a big reason behind this, they invest in stock market without predicting future trend analysis of stocks ups/downs, that repel them in darky night. Some investors follow stock market predictions rules and enjoy great profits. Several tools support in stock market predictions as python, Java, Rstudio with R programming and Python programming as well (Wang Qili *et al*, 2018). Mostly predictions based on ML algorithms that are also used in these tools. Some famous prediction algorithms that are supported by several tools as well as programming languages are Regression (Linear, Logistic), K-Nearest Neighbor (KNN), Decision Tree (DT), Artificial Neural Network (ANN) and Simple Moving Average (SMA) with the help of Time Series Forecasting (TSF). In this research, we applied these ML algorithms on different stock datasets obtained for GOOGLE, FB, AMAZON and some other companies (Thomas Fischer *et al*, 2018). We predicted results by Rstudio and Excel sheets by using statistical as well as graphical presentation of results. After obtaining results by different predictors and ML techniques we performed comparison analysis to argue about best considerations. Several Machine learning algorithms can be used for stock market prediction but in this research we used few algorithms like Linear regression(LR), Three Months Moving Average(3MMA) and Exponential Smoothing and if we further consider many other algorithms can also be used for Stock Market Predic-tion(SMP).In Whole research we found Exponential Smoothing predictions results are best rather than Linear Regression(LR) and Three Months Moving Average(3MMA) (Bruno *et al*, 2019) .

2.1. Machine Learning (ML) algorithms Implementation for Stocks Prediction

We will implement machine learning algorithms on above explained datasets and we will also analyses the trends of data manipulation as combined analysis of GOOGLE, FB,AMAZON,AAPPLE data. Mostly data is obtained from yahoo finance (FengmeiYang *et al*, 2019).

Table 1.: Amazon Data representation from Jan 2019 to July 2019

Date	Open	High	Low	Close	Volume
1/2/2019	1465.2	1553.36	1460.93	1539.13	7983100
1/3/2019	1520.01	1538	1497.11	1500.28	6975600
1/4/2019	1530	1594	1518.31	1575.39	9182600
1/7/2019	1602.31	1634.56	1589.19	1629.51	7993200
1/8/2019	1664.69	1676.61	1616.61	1656.58	8881400
1/9/2019	1652.98	1667.8	1641.4	1659.42	6348800
1/10/2019	1641.01	1663.25	1621.62	1656.22	6507700
1/11/2019	1640.55	1660.29	1636.22	1640.56	4686200
-	-	-	-	-	-
7/25/2019	2001	2001.2	1972.72	1973.82	4136500

2.1.1. AMAZON(AMZN) data analysis and Stock market predictions with Machine Learning (ML)

AMZN considered as huge stock market that attracts investors to buy/sell its shares and its becoming a trendy business market in the world.

2.1.2. AMZN Data Representation and data ranges

In this research we considered data from start of year as January 2019 to 25 July 2019 then we evaluated our approach with different prediction methodologies.

Amazon One Month Stock Data and Normalized Value

Table 2: July 2019 data for normalization (Amazon Yahoo Finance)

Date	Open	High	Low	Close	Adj Close	Volume	Normalized
7/26/2019	1942	1950.9	1924.51	1943.05	1943.05	4921900	0.96
7/25/2019	2001	2001.2	1972.72	1973.82	1973.82	4136500	0.98
7/24/2019	1969.3	2001.3	1965.87	2000.81	2000.81	2631300	0.99
7/23/2019	1995.99	1997.79	1973.13	1994.49	1994.49	2703500	0.99
7/22/2019	1971.14	1989	1958.26	1985.63	1985.63	2900000	0.98
7/19/2019	1991.21	1996	1962.23	1964.52	1964.52	3185600	0.97
7/18/2019	1980.01	1987.5	1951.55	1977.9	1977.9	3504300	0.98
7/16/2019	2010.58	2026.32	2001.22	2009.9	2009.9	2618200	0.99
7/12/2019	2008.27	2017	2003.87	2011	2011	2509300	1.00
7/11/2019	2025.62	2035.8	1995.3	2001.07	2001.07	4317800	0.99
7/10/2019	1996.51	2024.94	1995.4	2017.41	2017.41	4931900	1.00

Date	Open	High	Low	Close	Adj Close	Volume	Normalized
7/9/2019	1947.8	1990.01	1943.48	1988.3	1988.3	4345700	0.98
7/8/2019	1934.12	1956	1928.25	1952.32	1952.32	2883400	0.97
7/5/2019	1928.6	1945.9	1925.3	1942.91	1942.91	2628400	0.96
7/3/2019	1935.89	1941.59	1930.5	1939	1939	1690300	0.96
7/2/2019	1919.38	1934.79	1906.63	1934.31	1934.31	2645900	0.96
7/1/2019	1922.98	1929.82	1914.66	1922.19	1922.19	3203300	0.95

We obtained data from yahoo finance for AMZN stocks prediction. We applied normalization process on data and obtained values that are present in right corner of Table 2. Normalization is considered as ratio between close prices and maximum close prices. If we see the maximum close prices in Table 2 is 2020.99. We divided all close prices by maximum close price and obtained normalized column.

2.1.3. Graphical representation of Amazon trend analysis

Trends according to date can be checked about any product that is part of stock market. For this analysis we drawn graph between open/high/low/close prices. We can check how prices are moving day by day. Instead of complete study of data graph can give fully analytical view of market trends. Market always remain in changing process and statistical analysis give entire data outlook without any complexities.

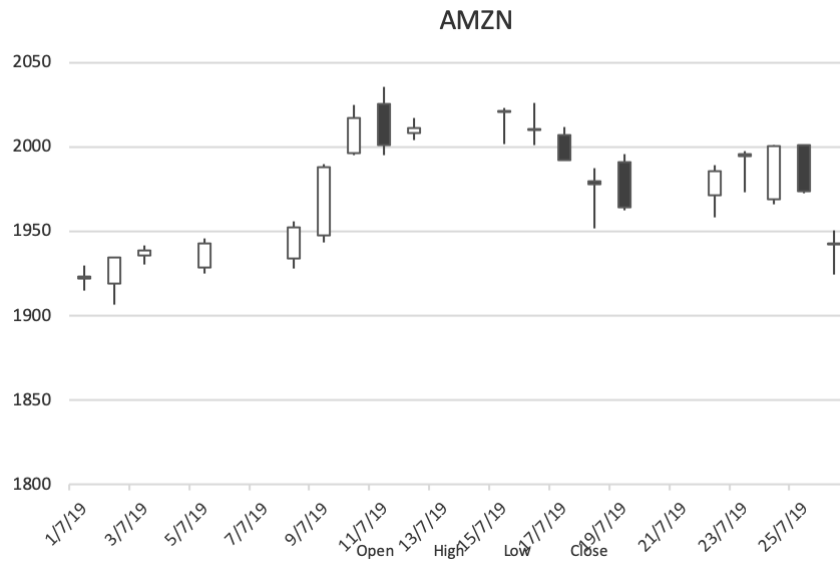


Figure 1: Graphical representation of AMZN Stock market trends.

Graph is representing a trend pattern where central dates of month showing high values. At start of month prices trend slow with central part of month it raises to high and at end of moth its showing medium rate of price changings Fig2.

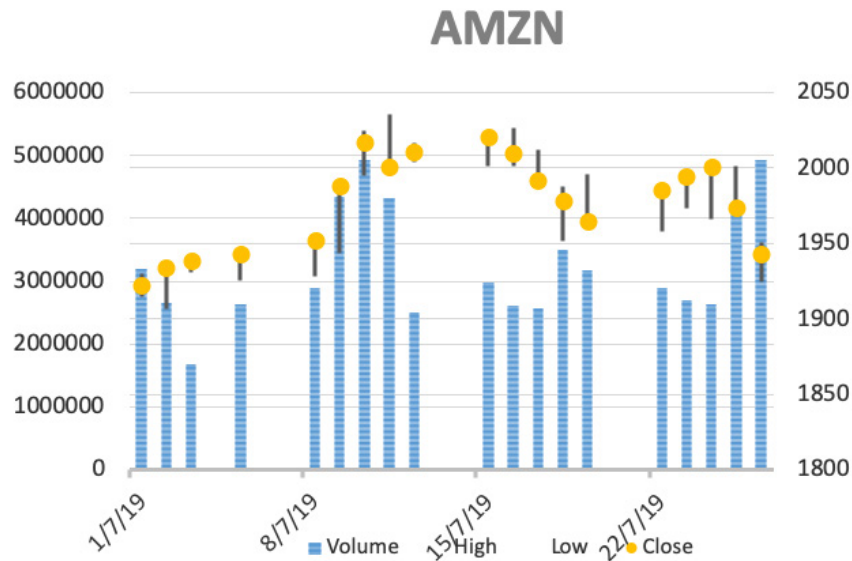


Figure 2: Amazon trend analysis graph with data adjustments.

2.2. Linear Regression (LR) introduction as ML algorithm

These algorithms can be understood easily and can be implemented easily. This algorithm runs into risky and over fitting environment easily. In some cases these algorithms are considered very much simple to solve complex problems. Linear regression runs under the relationship of two variables as one variable considered and dependent variable and other is considered as explanatory variable. A linear regression line has an equation of the form equation $Y = a + bX$, where X is the explanatory variable and Y is the dependent variable. The slope of the line is b , and a is the intercept (the value of y when $x = 0$) by (Vineet Maheshwari *et al* ,2019).

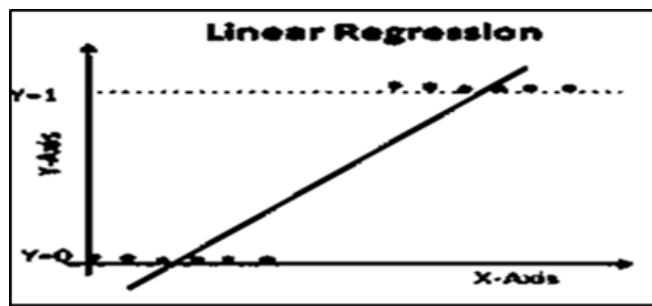


Figure 3: Linear regression/Logistic regression representation.

2.2.1. Linear Regression (LR) implementation for AMZN stock prediction

Linear regression is used for predictions with data that has numeric target variable. During prediction we use some variables as dependent variables and few considered as independent variables. In situation when there is one dependent and one independent variable, we prefer to use linear regression methodologies. Regression can be single variable or multi variable, it depends upon situation named as single variable or multi variable regression (FengmeiYang *et al* , 2019). Here in stock market prediction process, we have one date variable and one closing price variable. Closing price variable is our independent variable which also be considered as target variable. In this processing we will generate a prediction equation using liner regression method. We will generate prediction as $y=c+bx$, then we can say 'Y' is our predicted stock price and x is actual price.

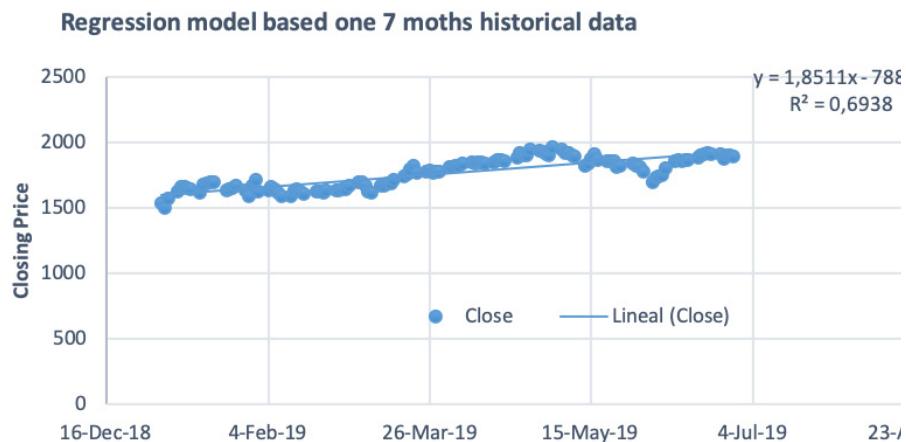


Figure 4: Regression Line for historical data of closing prices with respect to date.

In Fig4, linear regression line drawn between date and historical closing prices from Jan 2019 to July 2019. When we used linear regression graph in MS Excel it given us an equation, where ‘Y’ representing predicted value and ‘X’ represents actual value. By using this equation, we can predict future stock prices.

$$\text{Linear Regression Equation } Y=1.8511X - 78860 \quad \text{Eq1}$$

$$R^2 = 0.6938$$

R^2 represents root mean square error that is considered as classification error, we can say 0.6% classification performed wrongly by linear regression and remaining working is correct and error free. In this processing data of closing prices from Jan 2019 to July 2019 used. We considered only date and closing prices for prediction. Equation 1 is our prediction with respect to actual values. By using this equation, we can predict next month closing prices that can be great beneficial to stockiest. Now let us consider predictions based on linear regression and generated data.

Table 3: Linear regression classification for stock market prediction

JULY	Actual Prices	Predicted Prices	ABS Error	Error^2	% Error
1-Jul-19	1922.19	1,934.96	12.77	163.12	0.66%
2-Jul-19	1934.31	1,936.81	2.50	6.26	0.13%
3-Jul-19	1939.00	1,938.66	0.34	0.11	0.02%
25-Jul-19	1973.82	1,979.39	5.57	31.00	0.28%
		Average=	24.31	1065.03	1%
			MAD	MSE	MAPE
		r-square=	0.6938		
		r=	-0.83295		

Now let’s consider Table 3 for prediction explanation, In this table we considered July as our actual data. When we used this data there was only date and closing actual price columns were present. We used linear regression curve to obtain all other columns of table and predicted values. In this Table predicted values based on equation 1, we used MS Excel for the purpose of our statistical as well as linear regression prediction analysis.

2.2.2. Graphical representation of actual and predicted prices by line graph

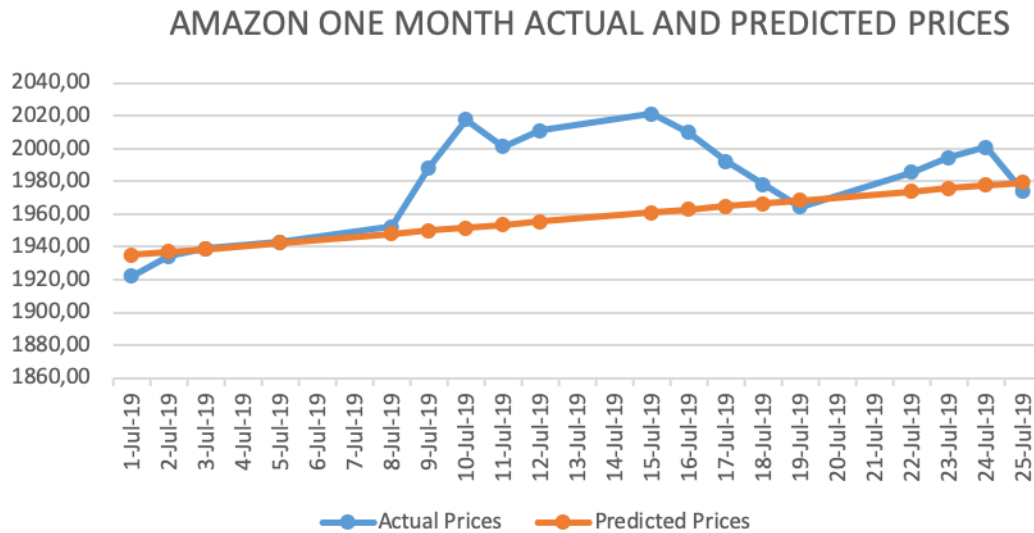


Figure 5: Line graph between actual and predicted prices with respect to date.

There can be different types of error present in our predictions that are explained by Table 3.

- **Absolute Error:** If we want to calculate absolute error in our prediction we can perform calculations by predicted values-actual close prices. In MS Excel it can be calculated by following formula

$$\text{ABS Error} = Fx = \text{abs}(\text{forecast} - \text{actual})$$

Similarly, we can calculate other types of errors in our predictions

- **Square Error:** It can be defined as square of absolute error = $(\text{ABS Error})^2$
- **%age Error:** When absolute error is divided by actual close then we obtain percentage error

$$\% \text{ Error} = \text{ABS Error} / \text{Actual Close price}$$

- **Average absolute error** = Mean absolute deviation = **MAD**
- **Average square error** = **Mean of square error** = **MSE**
- **Average percentage error** = **Mean absolute %age error** = **MAPE**

As r-square is 0.6938, this represents root square value or it can be said as variability in closing prices of AMZN stock market.

- R-value is $-\sqrt{\text{r-square}} = r = -0.8329$

When size of data increases the accuracy of results also increases. Here we used only seven months data for predictions on the basis of regression equation. If we check only one-month data on the basis of linear regression, there will be large amount of variability and all types of error will increase.

2.2.3. Linear Regression based on One-month AMMZ Stock data

Here we used only June month data to draw linear regression curve, we can see all types of error values increases. We can say as SIMPLE size will be larger the prediction results will be accurate and vice versa.

Regression Model based on one month data

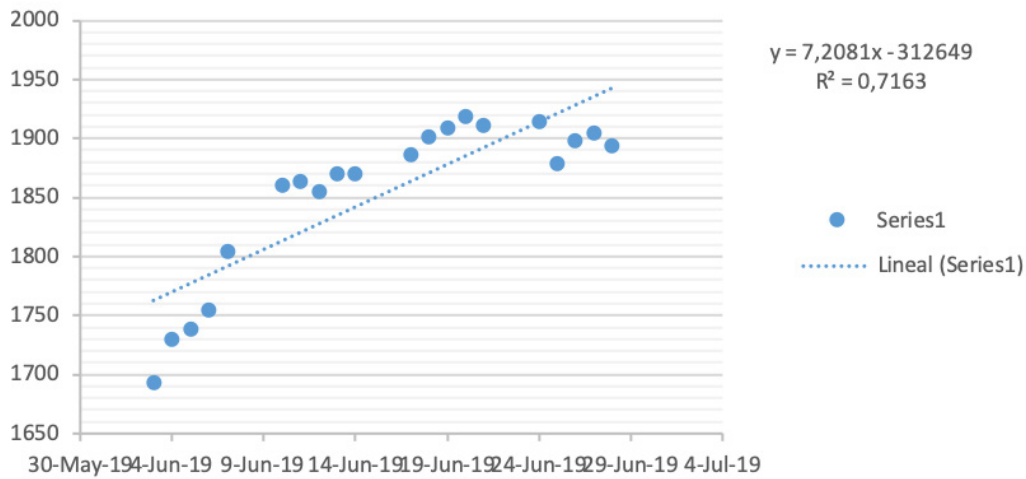


Figure 6: Regression equation generated on one month data as on JUNE.

Table 4: Predictions based on one-month regression line for AMZN Stock

JULY	Actual Prices	Predicted Prices	ABS Error	Error^2	% Error
1-Jul-19	1922.18994	1,962.94	40.75	1660.62	2.12%
2-Jul-19	1934.31006	1,970.15	35.84	1284.42	1.85%
3-Jul-19	1939	1,977.36	38.36	1471.25	1.98%
5-Jul-19	1942.91003	1,991.77	48.86	2387.60	2.51%
8-Jul-19	1952.31995	2,013.40	61.08	3730.46	3.13%
10-Jul-19	2017.41003	2,027.81	10.40	108.23	0.52%
15-Jul-19	2020.98999	2,063.85	42.86	1837.33	2.12%
16-Jul-19	2009.90002	2,071.06	61.16	3740.81	3.04%
17-Jul-19	1992.03003	2,078.27	86.24	7437.38	4.33%
18-Jul-19	1977.90002	2,085.48	107.58	11573.11	5.44%
19-Jul-19	1964.52002	2,092.69	128.17	16426.65	6.52%
22-Jul-19	1985.63001	2,114.31	128.68	16558.75	6.48%
23-Jul-19	1994.48999	2,121.52	127.03	16136.34	6.37%
24-Jul-19	2000.81006	2,128.73	127.92	16362.74	6.39%
25-Jul-19	1973.81995	2,135.94	162.12	26281.32	8.21%
		Average:	72.47	7231.59	4%
			MAD	MSE	MAPE
		r-square=	0.7163		
		r=	-0.84635		



Here we can compare the results of Table3 and Table4, the values of error percentage increases from 1% to 4% that is prove of our argument. In other words, we can say as SAMPLE size increases the chances of errors in predictions decreases.

If we want to improve accuracy in our predictions there are two more linear regression techniques that supportively work to decrease variability in actual and predicted values of closing prices. If we check Fig5, there is less symmetry between actual and predicted stock price line. The amount of variability in actual and predicted values of stock prices indicate accuracy of classifier. Let's decrease variability and increase accuracy by following methods.

2.2.4. Stocks prediction using 3-month average measurements (AMZN)

Table 5: AMZN stock prediction using 3-month average method

JULY	Actual Prices	Predicted Prices	ABS Error	Error^2	% Error
26-Jun-19	1897.83				
27-Jun-19	1904.28				
28-Jun-19	1893.63				
1-Jul-19	1922.19	1,898.58	23.61	557.43	1.23%
2-Jul-19	1934.31	1,906.70	27.61	762.32	1.43%
3-Jul-19	1939.00	1,916.71	22.29	496.84	1.15%
5-Jul-19	1942.91	1,931.83	11.08	122.69	0.57%
8-Jul-19	1952.32	1,938.74	13.58	184.41	0.70%
9-Jul-19	1988.30	1,944.74	43.56	1897.19	2.19%
10-Jul-19	2017.41	1,961.18	56.23	3162.19	2.79%
11-Jul-19	2001.07	1,986.01	15.06	226.80	0.75%
15-Jul-19	2020.99	2,009.83	11.16	124.62	0.55%
18-Jul-19	1977.90	2,007.64	29.74	884.47	1.50%
19-Jul-19	1964.52	1,993.28	28.76	826.95	1.46%
22-Jul-19	1985.63	1,978.15	7.48	55.95	0.38%
23-Jul-19	1994.49	1,976.02	18.47	341.26	0.93%
24-Jul-19	2000.81	1,981.55	19.26	371.08	0.96%
25-Jul-19	1973.82	1,993.64	19.82	392.97	1.00%
		Average	21.08	609.22	1%
			MAD	MSE	MAPE
		r-square=	0.6938		
		r=	-0.83295		

In this method we used previous three days average to predict next day price. If we compare Table3,4,5, It can be said the values of average absolute error decreases by 3-month average regression method. The value of MAE in Table3 was 24.31, in Table 4 was 74.47 but in Table5 it is 21.08 that represents less error chances by using 3-month average method. Below is graphical representation of actual vs predicted closing prices that can be considered better prediction.

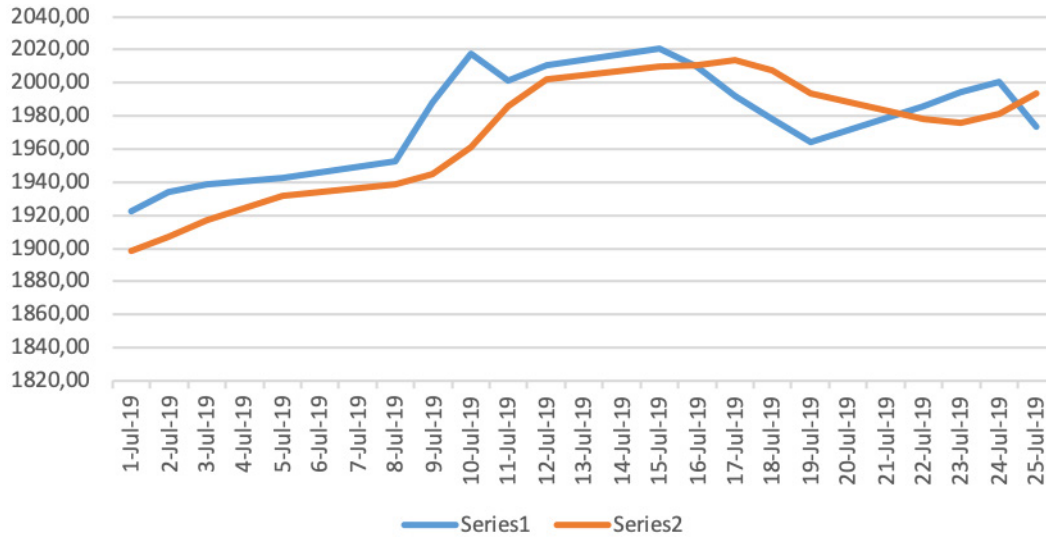


Figure 7: AMZN stock prediction using 3-month average method (Table 5).

Here in Fig7 blue line represents actual stock prices of AMZN products and orange line representing predicted closing prices of AMZN stocks. If we compare Fig 5 and Fig7 we can imagine about prediction accuracy.

2.2.6. AMZN-Stock prediction using Exponential Smoothing method

In this method we use an alpha value as alpha can be 0.4 or 0.8 then we consider first three close prices as seed values and from first three values central value considered as seed of prediction value, then we apply formula for prediction calculation.

Following formula work for exponential smoothing measurements

Prediction with exponential smoothing=

$$\text{Exp smoothing} = \text{previous forecast} + \alpha * (\text{actual value} - \text{forecast value})$$

For prediction test we change values according to need, first we put alpha=0.4 and check the values of error measures, then we place alpha=0.8 and the changings in errors measurements considered. Whole processing, we did in MS excel, let us consider data below for further explanation.

Table 6: AMZN stock prediction using exponential smoothing

	alpha=	0.8			
JULY	Actual Prices	Predicted Prices	ABS Error	Error^2	% Error
26-Jun-19	1897.83				
27-Jun-19	1904.28				
28-Jun-19	1893.63	1904.28			
1-Jul-19	1922.19	1,895.76	26.43	698.54	1.37%
2-Jul-19	1934.31	1,916.90	17.41	302.97	0.90%
17-Jul-19	1992.03	2,011.65	19.62	385.08	0.99%

	alpha=	0.8			
JULY	Actual Prices	Predicted Prices	ABS Error	Error²	% Error
19-Jul-19	1964.52	1,981.51	16.99	288.69	0.86%
22-Jul-19	1985.63	1,967.92	17.71	313.71	0.89%
23-Jul-19	1994.49	1,982.09	12.40	153.82	0.62%
24-Jul-19	2000.81	1,992.01	8.80	77.45	0.44%
25-Jul-19	1973.82	1,999.05	25.23	636.55	1.28%
		Average	16.62	363.83	1%
			MAD	MSE	MAPE
		r-square=	0.6938		
		r=	-0.83295		

Here in Table 6, we can check accuracy improvements absolute error value decreases up to 16.62 that is less from all other methods. In short, we can say the prediction values by exponential smoothing is best in results. Error accordance decreases in this method and as the value of errors decreases the accuracy of predicted results can be considered best.

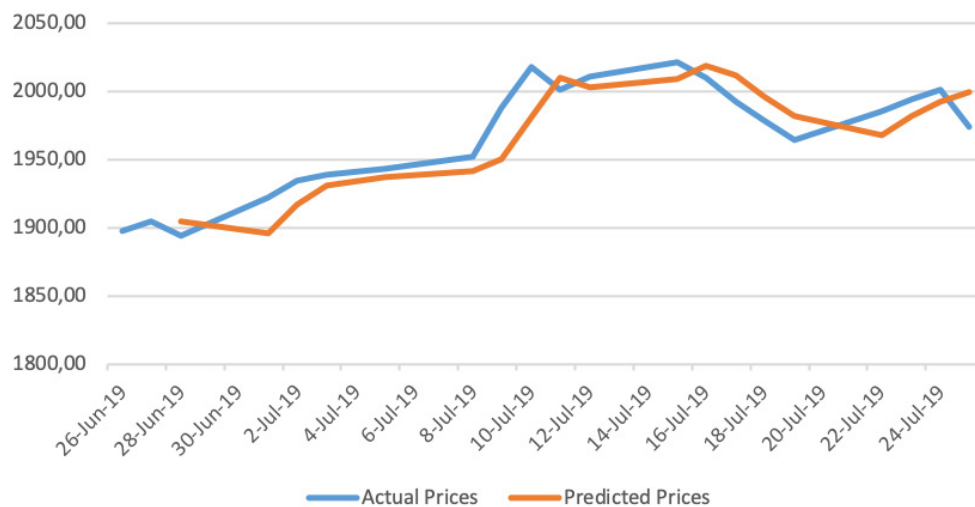


Figure 8: Exponential smoothing results for AMZN Stock prediction.

In this previous study we considered 3 different methods for stock prediction, firstly we considered linear regression and results were presented with 24.31 average absolute error, secondly we used only one month sample for linear regression equation then the value of average absolute error was 72.43,

Thirdly we used 3-month average regression methodology and obtained 21.08 average absolute error value, at last we used exponential smoothing measurements based on linear regression methodology and obtained 16.62 average absolute error values that was most significant results that were predicted on the basis of AMZN stock market data from yahoo finance market. At last we can say exponential smoothing proved its prediction best as the value of average absolute error was smallest than all others.

2.2.7. AMZN One-month prediction on the basis of linear regression equation

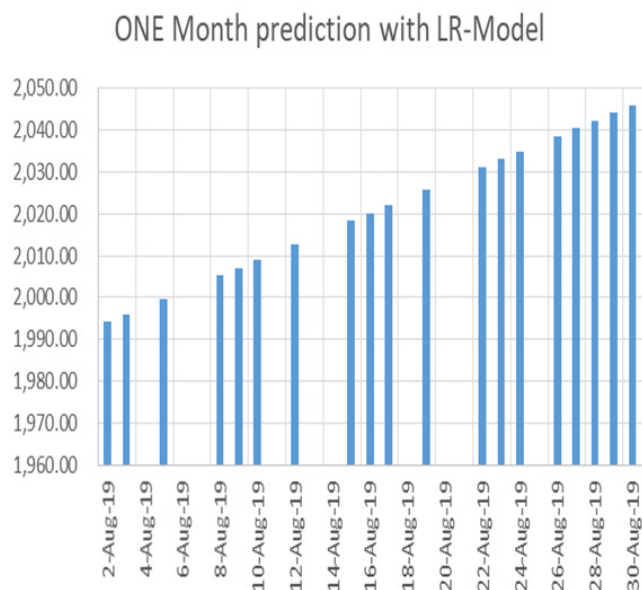


Figure 9: One-month prediction graph for AMZN Stocks.

After apply four different working methodologies we are able to predict next month prices of Amazon stock market, below is next month prediction as well as graphical representation of predicted prices. In linear regression equation we used seven months data as historical record and after generating equation we evaluated next month predictions.

Table 7: One-month prediction for AMZN Stock prices

August	Predictedcloseprice
2-Aug-19	1,994.20
3-Aug-19	1,996.05
5-Aug-19	1,999.75
8-Aug-19	2,005.30
9-Aug-19	2,007.15
12-Aug-19	2,012.71
16-Aug-19	2,020.11
17-Aug-19	2,021.96
19-Aug-19	2,025.67
22-Aug-19	2,031.22
23-Aug-19	2,033.07
24-Aug-19	2,034.92

August	Predictedcloseprice
26-Aug-19	2,038.62
27-Aug-19	2,040.47
28-Aug-19	2,042.33
29-Aug-19	2,044.18
30-Aug-19	2,046.03

From Figure 9, it is cleared that AMZN stock prices increasing day by day, that is why graph is presenting trending scenario. Prediction give estimation idea of future, by analyzing prediction results people can value their business.

2.3. Time Series Forecasting method for AAPL Stock Market Prediction

For time series forecasting method, we used AAPL Stock market data from Jan 2019 to July 2019. The data was obtained from yahoo finance and before stock market prediction measurements; we applied some preprocessing to organize data for better prediction results. We converted each month into three quarters, one quarter was ten days and total thirty days for each month become equal to 1,2,3 values. Now we can say for seven months there was considered 21 days or 21 quarters. For each quarter, we obtained average of ten working stock market days.

Now let us examine the AAPL data before and after preprocessing.

Table 8: General representation of AAPL data from yahoo finance

Date	Open	High	Low	Close	Adj Close	Volume
2-Jan-19	154.89	158.85	154.23	157.92	156.6424	37039700
3-Jan-19	143.98	145.72	142	142.19	141.0396	91312200
4-Jan-19	144.53	148.55	143.8	148.26	147.0605	58607100
7-Jan-19	148.7	148.83	145.9	147.93	146.7332	54777800
8-Jan-19	149.56	151.82	148.52	150.75	149.5304	41025300
9-Jan-19	151.29	154.53	149.63	153.31	152.0697	45099100
10-Jan-19	152.5	153.97	150.86	153.8	152.5557	35780700
11-Jan-19	152.88	153.7	151.51	152.29	151.0579	27023200
14-Jan-19	150.85	151.27	149.22	150	148.7864	32439200
15-Jan-19	150.27	153.39	150.05	153.07	151.8316	28710900
16-Jan-19	153.08	155.88	153	154.94	153.6865	30569700
17-Jan-19	153.08	155.88	153	154.94	153.6865	30569700
18-Jan-19	153.08	155.88	153	154.94	153.6865	30569700
19-Jan-19	153.08	155.88	153	154.94	153.6865	30569700
-	-	-	-	-	-	-
25-Jul-19	208.89	209.24	206.73	207.02	207.02	13909600
26-Jul-19	207.48	209.73	207.14	207.74	207.74	17601200

Only dates and closing prices of stock market AAPL product was our actual need to predict future, then we considered stock values as follow

Table 9: AAPL data before and after preprocessing

Used AAPL data for predictions Data before preprocessing		Used AAPL data for predictions Data after preprocessing			
Date	Close	T	Month	Quarter	Close Price
2-Jan-19	157.92	1	January	1	150.59
3-Jan-19	142.19	2		2	153.83
4-Jan-19	148.26	3		3	157.54
7-Jan-19	147.93	4	February	1	171.26
8-Jan-19	150.75	5		2	170.67
9-Jan-19	153.31	6		3	173.43
10-Jan-19	153.8	7	March	1	174.38
11-Jan-19	152.29	8		2	184.26
14-Jan-19	150	9		3	189.83
15-Jan-19	153.07	10	April	1	196.69
16-Jan-19	154.94	11		2	200.55
-	-	12		3	204.86
-	-	13	May	1	205.44
25-Jul-19	207.02	14		2	187.72
26-Jul-19	207.74	15		3	178.63
		16	Jun	1	183.90
		17		2	195.69
		18		3	198.40
		19	July	1	202.49
		20		2	203.77
		21		3	207.90
		22	August	1	
		23		2	
		24		3	

In Table 9, there is present on month as august without values that will be predicted after all type of time series forecasting process. As we divided each month into three quarters then the approach that need to be predict is august three quarters. We will perform different processing methods for the finding of different forecasting values. Before further processing let's consider graphical representation of data after preprocessing.

2.3.1. Graphical representation of data after preprocessing

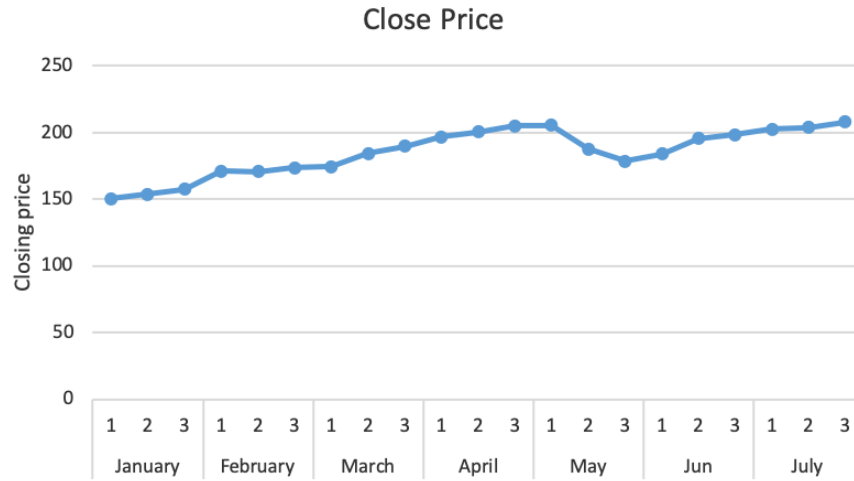


Figure 10: Quarter based graphical data representation for AAPL data.

Now according to Fig 10 each month is represented by three dots and we can say it is one quarter. For further quarter consideration all data is present in Table 9. If we consider this graph there is no regularity in data some time it's going straight sometime its going up sometime its going down then we can say there is no presence of seasonality.

Seasonality can be considered when continuously data changing with a trend in each month. As we divided each month into three quarters with each quarter of ten days now let's consider moving average with MA(3). For each quarter value of MA can be considered as average of three quarters used in each month.

For first value the required result for MA can be predicted with average of 1,2,3=2nd value of the month, similarly we can find out all other patterns for whole data by using MS excel.

2.3.2. MA values for closing prices with AAPL stock data

Table 10: MA values for closing prices with AAPL stock data

T	Month	Quarter	Close Price	MA(3)
1	January	1	150.59	
2		2	153.83	153.99
3		3	157.54	160.88
4	February	1	171.26	166.49
5		2	170.67	171.79
6		3	173.43	172.83
7	March	1	174.38	177.36
8		2	184.26	182.82
9		3	189.83	190.26
10	April	1	196.69	195.69
11		2	200.55	200.70

T	Month	Quarter	Close Price	MA(3)
12		3	204.86	203.62
13	May	1	205.44	199.34
14		2	187.72	190.60
15		3	178.63	183.42
16	Jun	1	183.90	186.08
17		2	195.69	192.67
18		3	198.40	198.86
19	July	1	202.49	201.55
20		2	203.77	204.72
21		3	207.90	205.83
22	August	1		
23		2		
24		3		

In Table 10, MA is considered as central value of three quarters of a month. Now let us consider other terms for this data that completely represented in Table 11. Before the presentation of other terms we will consider there meanings and use in whole process.

Let us now consider central moving average of three quarters of a month with CMA(3). CMA is considered as center of two MA values for first moth and it can be said as central value of MA values in three quarters of a month.

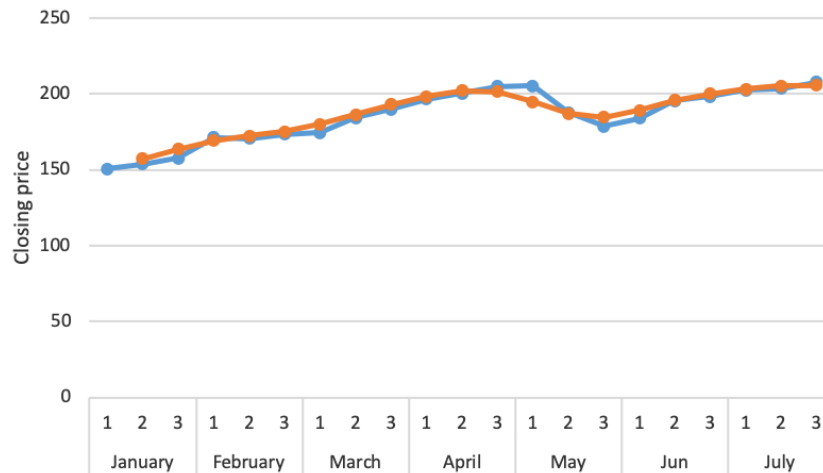


Figure 11: Line graph for actual closing and predicted central moving average (CMA).

In Fig 11, blue line representing actual closing prices and orange line representing predicted central moving average CMA(3) values that can be checked in Table 17.

Now for further accuracy and other parameters finding we will do more steps. Next step is called smoothing, for this purpose we need two variables one is called sessional component(S_t) and other is called irregular component(I_t), final result for smoothing can be considered by product of both of these variables($S_t I_t$).

$$\text{Smoothing} = S_t I_t = \text{actual closing price}/\text{CMA}(3) \quad \text{Eq3}$$

Let us consider **classical model for time series** that can be considered as

$$\text{Actual Close Prices} = Y_t = S_t * I_t * T_t \quad \text{Eq4}$$

Let us now consider full output by linear regression before further consideration of Time series forecasting. **T_t** can be obtained by addition of coefficients present in Table 16.1, and then multiplying with T values as serial number of stock market.

Table 11: Summary output linear regression

SUMMARY OUTPUT LINEAR REGRESSION						
<i>Regression Statistics</i>						
Multiple R		0.810585				
R Square		0.657048				
Adjusted R Square		0.637995				
Standard Error		9.829845				
Observations		20				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	3332.184	3332.184	34.48543		1.47E-05
Residual	18	1739.265	96.62585			
Total	19	5071.449				

Table 11.1: Time series forecasting summary with output results

	<i>Coefficients</i>	<i>Standard Error</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	162.6366	4.903824	1.36E-17	152.3341	172.93918	152.3341	172.9392
1	2.238482	0.381185	1.47E-05	1.437642	3.0393223	1.437642	3.039322

Seasonality and irregularity components tell about the distance closeness or away position from actual price line of graph. S_t in Table 17 is sessional component, let us consider its method of working. Now let us consider depersonalize values for data, this can be done actual prices (Y_t)/ S_t that is also necessary for stock values prediction. Now for forecasting we can obtain results of our predictions by dividing seasonality with T_t , that is considered our actual prediction obtained from whole process. Below is the obtained components table that is obtained by methodology for stock values prediction of AAPL stock for next month that is represented by yellow color.

Table 12: Time series forecasting with predicted stock prices for next month august

Month	CMA(3)	$S_t I_t$	S_t	DESEZ	T_t	Forecasting Predicted prices
January			1.00	150.59	164.88	164.88
	157.43	0.98	0.99	155.38	167.11	165.44
	163.68	0.96	0.99	159.14	169.35	167.66

Month	CMA(3)	$S_t I_t$	S_t	DESEZ	T_t	Forecasting Predicted prices
February	169.14	1.01	1.00	171.26	171.59	171.59
	172.31	0.99	0.99	172.39	173.83	172.09
	175.09	0.99	0.99	175.19	176.07	174.31
March	180.09	0.97	1.00	174.38	178.31	178.31
	184.75	0.97	0.99	180.43	196.21	194.25
Jun	189.37	0.97	1.00	183.90	198.45	198.45
	195.76	1.00	0.99	197.67	200.69	198.68
	200.21	0.99	0.99	200.40	202.93	200.90
July	203.13	1.00	1.00	202.49	205.17	205.17
	205.27	0.99	0.99	205.82	207.41	205.33
	205.83	1.01	0.99	210.00	209.64	207.55
August			1.00		211.88	211.88
			0.99		214.12	211.98
			0.99		216.36	214.20

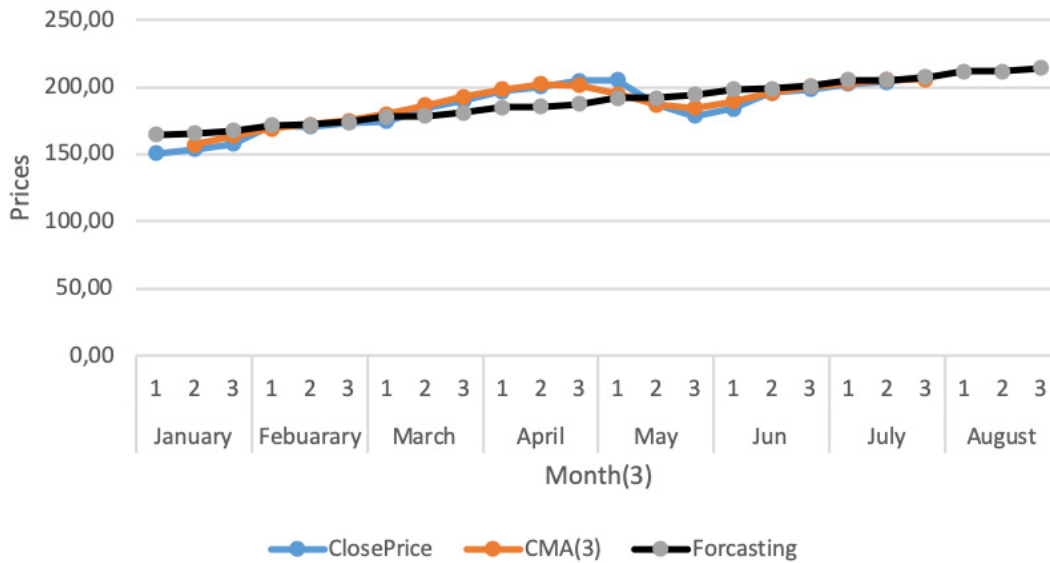


Figure 12: Graphical representation of actual and predicted stock prices for AAPL.

3. Conclusion

Stock market prediction is actual demand for beneficial business. Predictions always helpful to decrease risk factor in any business environment. Risk factor can be analyzed on the basis of historical data and previous business trends. This research based on several results and we used machine learning algorithm (ML) as Linear Regression (LR) with respect relations to business priority. Linear regression applied on different data sets that were obtained

from stock market place (Yahoo finance). Yahoo Finance ever considered as best market place for obtaining stock market data about any product. In our research we used Amazon (AMZN) and Apple (AAPL) datasets for our practical approaches. Before applying ML on datasets, we analyzed stock market trends for both products. Trend analysis also provide predictions about future business plan. In next step first we used AMZN dataset and after analysis of stock market trend we applied linear regression with the help of Excel statistical graphs. Secondly, we applied three month moving average(3MMA) method to predict stock market prices of AMZN products. Thirdly we applied exponential smoothing (ES) for predictions. After comparing all results, we obtained hypothesis that exponential smoothing prediction results given less error and greater accuracy and we considered it best stock market predictor with general trend analysis. Similarly, we applied these three methods on AAPL data and obtained results about predictions. After applying these methodologies, we capable to predict one-month forward stock market trend and we presented August prices as founded throughput. At end of previous chapter, we applied Time Series Forecasting methodology and predicted AAPL stock prices for next month. Time Series Forecasting method also introduced new ways for stock market trend analysis. At last we can say by applying this research methodology we are able to predict future stock market trends easily. Several Machine learning algorithms can be used for stock market prediction but in this research we used few algorithms like Linear regression(LR), Three Months Moving Average(3MMA) and Exponential Smoothing and if we further consider many other algorithms can also be used for Stock Market Prediction(SMP).In Whole research we found Exponential Smoothing predictions results are best rather than Linear Regression(LR) and Three Months Moving Average(3MMA).

4. References

- Tae Kyun Lee *et al.* “Global stock market investment strategies based on financial network indicators using machine learning techniques”, *Expert Systems with Applications*, 117(2019):228-242.
- Bruno *et al.* “Literature review: Machine learning techniques applied to financial market prediction”, *Expert Systems with Applications*, 124(2019): 226-251.
- Thomas Fischer *et al.* “Deep learning with long short-term memory networks for financial market predictions”, *European Journal of Operational Research*, 270.2(2018): 654-669.
- Kang Zhang *et al.* “Stock Market Prediction Based on Generative Adversarial Network”, *Procardia Computer Science*, 147(2019):400-406
- Ben Moews *et al.* “Lagged correlation-based deep learning for directional trend change prediction in financial time series”. *Expert Systems with Applications*, 120(2019):197-206.
- Fische Mospel *et al.* “Deep learning with machine learning algorithms for financial market predictions”, *European Journal of scientific Research*, 220.2(2019):654-669.
- Markku Karhunen *et al.* “Algorithmic sign prediction and covariate selection across eleven international stock markets”, *Expert Systems with Applications*, 115(2019):256-263.
- Eunsuk Chong *et al.* “Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies”, *Expert Systems with Applications*, 8 (2017): 187-205.
- WangQili *et al.* “Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles”, *Neurocomputing*, 299(2018):51-61.
- ChangjuLee *et al.* “Explaining future market return and evaluating market condition with common preferred spread index”, *Physica A: Statistical Mechanics and its Applications*, 220(2019):220-229.
- Mu-YenChen *et al.* “Modeling public mood and emotion: Stock market trend prediction with anticipatory computing approach”, *Computers in Human Behavior* (2018).
- Hiransha *et al.* “NSE Stock Market Prediction Using Deep-Learning Models”, *Procedia Computer Science*, 132(2018): 1351-1362.
- FengmeiYang *et al.* “A novel hybrid stock selection method with stock prediction”, *Applied Soft Computing* (2019).
- Yue-gang Song *et al.* Corrigendum to “Towards a new approach to predict business performance using machine learning” [*Cogn. Syst. Res.* 52 (2018): 1004-1012].