AI-Driven Intelligent Models for Business Excellence

Samala Nagaraj
Woxsen University, India

Korupalli V. Rajesh Kumar
Woxsen University, India

A volume in the Advances in Business Information Systems and Analytics (ABISA) Book Series

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

Published in the United States of America by
IGI Global
Business Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Nagaraj, Samala, 1980- editor. | Kumar, Korupalli V. Rajesh, 1989- editor.
Title: AI-driven intelligent models for business excellence / Samala Nagaraj, and Korupalli V. Rajesh Kumar, editors.
Description: Hershey, PA : Business Science Reference, 2022. | Includes bibliographical references and index. | Summary: "As digital technology is taking the world in a revolutionary way and business related aspects are getting smarter this book is a potential research source on the Artificial Intelligence-based Business Applications and Intelligence"-- Provided by publisher.
Identifiers: LCCN 2022015764 (print) | LCCN 2022015765 (ebook) | ISBN 9781668442463 (hardcover) | ISBN 9781668442470 (paperback) | ISBN 9781668442487 (ebook)
Subjects: LCSH: Business intelligence--Data processing. | Management--Data processing. | Artificial intelligence--Industrial applications.
Classification: LCC HD38.7 .A3898 2022 (print) | LCC HD38.7 (ebook) | DDC 658.4/038028553--dc23/eng/20220408
LC record available at https://lccn.loc.gov/2022015764
LC ebook record available at https://lccn.loc.gov/2022015765

This book is published in the IGI Global book series Advances in Business Information Systems and Analytics (ABISA) (ISSN: 2327-3275; eISSN: 2327-3283)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 13
Machine Learning-Based Stock Price Prediction for Business Intelligence

Bhavya K. R.
Reva University, India

L. N. C. Prakash K.
CVR College of Engineering, India

Malla Sudhakara
Reva University, India

Rupa Devi B.
AITS, India

G. Ramasubba Reddy
Sai Rajeswari Institute of Technology, India

Sangeetha M.
Reva University, India

ABSTRACT

The act of digital marketing uses a variety of traditional methods such as analyst consensus, earnings per share estimation, or fundamental intrinsic valuation. Also, social media management, automation, content marketing, and community development are some of the most popular uses for digital marketing. Stock price prediction is a challenging task since there are so many factors to take into account, such as economic conditions, political events, and other environmental elements that might influence the stock price. Due to these considerations, determining the dependency of a single factor on future pricing and patterns is challenging. The authors examine Apple's stock data from Yahoo API and use sentiment categorization to predict its future stock movement and to find the impact of "public sentiment" on "market trends." The main purpose of this chapter is to predict the rise and fall with high accuracy degrees. The authors use an artificial intelligence-based machine learning model to train, evaluate, and improve the performance of digital marketing strategies.

DOI: 10.4018/978-1-6684-4246-3.ch013

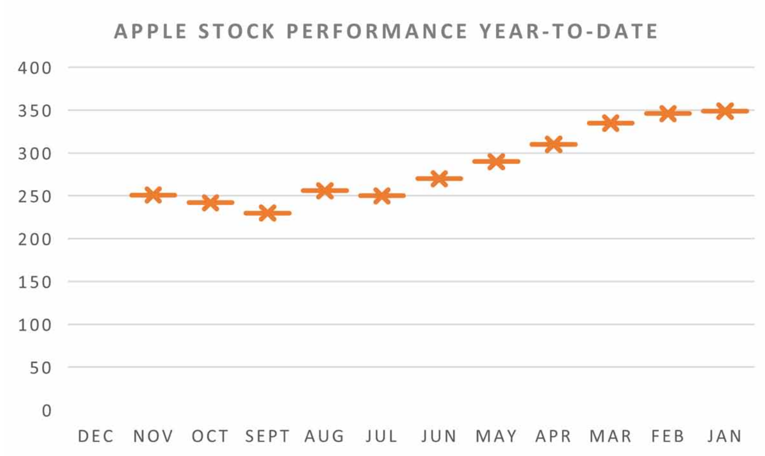
INTRODUCTION

Forecasting of stock prices is vital for financial backers and is quite possibly the most fascinating issues for scientist. As per the proficient market speculation (EMH) and irregular walk hypothesis, stock costs are considered to not have anything to do with chronicled patterns Notwithstanding, according to the viewpoint of conduct finance, financial backers' way of behaving and independent direction is frequently impacted by silly factors and commotion. The market by and large shows consistency. The forecast of monetary time series is a vital assignment. Analysts have directed a great deal of chips away at stock cost prediction. From one perspective, the customary technique is to utilize verifiable cost information inside the business sectors to foresee stock costs (Yang and Parwada, 2012). Then again, More and more peculiarities show that data outside the exchanging market ay essentially affect resource costs; Twitter's opinion condition of financial backers frequently influences stock returns (Nofer and Hinz, 2015). Prices are not set in stone by the essential worth and deviations brought about by financial backers' silly way of behaving (Szyszka, 2007). In these examinations, an ordinary strategy is to utilize monetary information, declarations, budget summaries, and other data to foresee stock costs (Shang and Wang, 2020). Nonetheless, news, reports, and declarations, as a rule, happen haphazardly, so the coherence of such sort of data is effectively impacted by time stretches, bringing about incomplete loss of significant data. Another technique is to utilize online information sources via web-based media stages to foresee the stock costs.

These examinations utilize online information from interpersonal organizations, for example, Twitter to assess financial backer feeling and dissect the connection among opinion and securities exchanges (Renault, 2020). Yet, the characterization of feeling is typically outrageous, like good and pessimistic. Additionally, it isn't permitted to quantify financial backer' social connection data in manners aside from the particular opinion aspect. Interpersonal organizations contain a great deal of significant data; however, the financial backer' social association information actually needs proper devices and innovations to change it into important data.

As of the end of October 2021, Microsoft posted revenue of \$45.3bn for the first quarter to September - 22% higher than the same period last year. Operating income rose 27% to \$20.2bn, while profits totalled \$20.5bn in GAAP and \$17.2bn in non-GAAP and \$17.2lbn non-GAAP increased by 48% and 24%. Microsoft 365 commercial revenue increased by 23%, driving an 18% increase in Office commercial products and cloud services. Microsoft 365 consumer subscribers increased to 54.1 million, driving a 10% increase in Office consumer products and cloud services. As a result, LinkedIn revenue rose by 42%, fuelled in part by Marketing Solutions growth of 61%. Up to date apple stock performance is shown in Figure 1.

Figure 1. Up to date stock performance



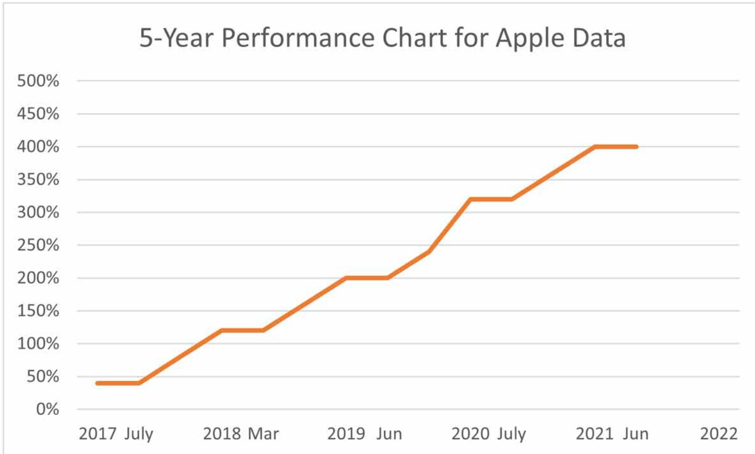
More normal service at software and cloud giant Microsoft (MSFT) has been resumed after a rocky start to 2022. Its stock had dropped around 20% over January but today, is showing a 10% improvement from that low at $304.56.In fairness the tech sector in general had taken a pounding in the new year but Microsoft was able to dispel some of the gloom by posting strong results for the second quarter. Microsoft projected third-quarter 2022 revenue growth of 17%, the mid-point of the company's range, while the projected mid-point for operating income was for $19.9bn both of which are ahead of analyst forecasts. At the time of writing (9 February), Microsoft had a market cap of $2.28trn, just behind Apple, which had a market cap of $2.85trn, according to Companies Market Cap.

Microsoft's stock is currently trading at $305. Major challenge is that it regains its upward momentum? According to the algorithmic forecasts of Wallet Investor, MSFT stock could rise to $377 within the next year investment. It doesn't provide a 10-year forecast for MSFT stock, but it predicts that MSFT could reach $669 within five years and the graphical representation of the corresponding statistics is shown in Figure 2.

According to Market Beat, 33 analysts rated MSFT stock as a 'buy', with just one saying it is a 'hold.' Analysts expect MSFT stock to rise 17% to $358 over the next year. The highest analyst forecast is $411, while the lowest is $290.

To analyse the strength of the correlation between Twitter sentiment and stock performance, we are collecting Apple's stock data from yahoo API in different sectors. Due to practical reasons, like the limited time interval and relevant data, we are focusing to only analyse data from social media Twitter during a six-week period

Figure 2. Year wise apple stock performance



The objective of this study is to utilize advanced AI method to work on the precision of offer cost prediction in industry. The Implementation of the interpersonal organization variable to expand the impact of stock value forecast is our key commitment. We study in the event that the interpersonal organization variable coming about because of financial backers' unconstrained consideration could give new explanations behind stock cost expectation. We concentrate on social networks, which use the "word of mouth" impact to propagate ideas, views, and trends. Other internet data sources, on the other hand, are unable to assess investor social interaction information in any way other than the unique dimension.

RELATED WORK

A stock's price depends on supply and demand, which in turn depends on fundamental factors and market sentiment and examples of fundamental aspects include the stock's perceived risk, discount rate and expected growth, inflation the strength of the economy, trends and stocks liquidity are among the technical factors while market sentiment refers to the psychology of the stock market participate.

Major examination is the method involved with checking out at a business at the most essential or key monetary level. This will inspect the critical proportions of a business to decide its monetary wellbeing. Key examination can give a thought of the worth of what an organization's stock ought to be. Henceforth, crucial investigation concerns the actual stock, like the resources, liabilities, and salaries of the objective organizations by examining their fiscal reports, as well as the proportions of past execution, like Price-to-Earnings Ratio (P/E proportion) (Harrington, 2009), which is the proportion for esteeming an organization that actions its present offer value, connected with its profit per-share (EPS). Principal examination depends on the faith in the business needs for the cash-flow to continue to work. On the off chance that the organization runs well, it ought to get extra capital honours which will prompt stock cost took off. Essential examination is directed from the worldwide economy initially and afterward public economy prior to breaking down a particular industry and a particular organization. It is a hierarchical interaction. As the essential examination is a moderately sensible and technique, it is utilized broadly (Harrington, 2010).

Crossover models generally perform better compared to single AI models, and the brain organizations and SVMs are regularly a significant part in numerous mixture models, for instance, the mix of brain organization and choice tree for the expectation of advanced game substance stocks value (Chang, 2011), the reconciliation of dim calculation and RBF brain organization, prepared by 4 different learning methodologies (Lei, 2018), fluffy time series investigation with brain networks for the estimate of the Taiwan Stock Exchange Capitalization. Developmental improvement with covering strategies is frequently used to advance AI models or select elements for stock cost (pattern) forecast. For instance, Harmony Search and Genetic Algorithm was utilized to track down the best design of brain organization (G¨oçken et al., 2016), brain network with avaricious calculation in light of the component decrease with covering procedures at the expectation of stock cost pattern (Lertyingyod and Benjamas, 2017), recreated tempering calculation was utilized to streamline include space and model boundaries (Torun, 2011), and a Markov choice interaction was joined on hereditary calculations to foster stock exchanging methodologies (Chang and Lee, 2017).

Following quite a while of stock exploration, in excess of 100 pointers and proportions have been produced for crucial and specialized investigation, individually. In this exploration, information is gathered from the data sets, it is the most famous stock data set. 30 elements are chosen, as they are the most prominently utilized in existing exploration. For the essential investigation, 15 highlights in 6 classes are chosen, though for the specialized examination, 15 elements in 10 gatherings are chosen. The gathered crude information will go through two phases of information handling. For the essential elements, information is gathered each quarter after the organizations distributing their monetary reports. For the specialized highlights, information is gathered each exchanging day, synchronized with the nearby cost. An example is the information on each exchanging day. To analyse the effect of two arrangements of elements on specialized investigation, the quarterly essential information is changed to day-to-day information by utilizing the univariate straight handling, which is chosen by the start and the closure upsides of the element in the quarter, as displayed in Fig. 1, where accept all the difference in a central component in a quarter is straight. While handling the specialized information, to keep the uprightness of the example information, the examples that miss some component values are erased. Simultaneously, the relating exchanging day tests moved from the principal information are likewise erased to keep up with the consistency of the key information and specialized information.

Recent studies show that investors can get early indicators of news by scraping different data sources such as blogs, social media and search engines. For example, Google used their software called Google Trends and showed that Google search queries could be used in order to predict consumer spending (Torsten Schmidt et al.2009). A previous paper by Bollen, Mao and Zeng show that mood and emotions are difficult in financial decision making,therefore it is safe to assume that mood and emotions also may impact asset pricing (Pagolu et al. 2018). Mao, Wang found that it is a important correlation between the daily number of tweets that mentions stocks in the stock index Standard & Poor 500 and stock pointers that are related to the price and volume. They also identified correlation between the daily number of tweets and the daily traded volume for 8 out of 10 industry sectors. (Shuyuan Deng et.al 2018).

Machine Learning techniques have becoming widely used in stock research due to their intrinsic potential to extract information from vast amounts of data (Kumar, Dogra, Utreja, & Yadav, 2018). Supervised Learning, Unsupervised Learning, and Reinforcement Learning are the three types of machine learning, according to Khadka (2019). Analysing and extracting meaningful patterns from massive raw input data is at the core of machine learning. Higher degrees of understanding for decision-making and trend prediction are the consequence. As a result, many companies place a high value on obtaining these insights and information from data since it allows them to achieve competitive advantages (Mohammadi, Al-Fuqaha, Sorour, & Guizani, 2018). Machine learning approaches, according to Sodhi, Awasthi, and Sharma (2019), can address difficulties.

Efficient-Market Hypothesis

Stock market prediction has been on the table for decades and has brought attention both from business and in science. One of the first theories on the subject was the Efficient-Market Hypothesis (EMH) which states that asset pricing reflects all available information. According to the EMH assets always trade on a fair value and markets only reacts on new information such as news, therefore is it impossible to beat the market over time on a risk-adjusted basis. News are unpredictable and this suggest that asset pricing follows a random walk with a 50% chance of going either up or down (Kenton 2018). EMH and The Random Walk Theory has received a lot of criticism and evidence show that asset pricing does not follow a random walk and in other words asset pricing can be predicted to some degree. For example, Warren Buffett wrote a article arguing that if several funds managed to beat the index year after year it cannot be an random event (Buffett, Warren 1984). The NYU Stern School of Business professor Aswath Damodaran referred this as a proof that the market is not always efficient (Damodaran 2014).

Correlation and Causality

Correlation is commonly defined as how close two variables are to having a linear relationship with each other. The co-variance of two events A and B is defined below and E[A] is the expected value:

$$C(A, B) = E[(A - E[A])(B - E[B])$$

If C (A, B) = 0 then they are said to be uncorrelated [7].

The correlation is defined as:

$$\boldsymbol{\rho}(A, B) = \frac{C(A;B)}{D(A)D(B)} C\left(A;B\right) \text{€} \left[-1,1\right]$$

where D(A) and D(B) represent the standard deviation of A and B.

If **ρ** (A, B) = 1 there is a true positive relationship between both A and B and if A increases B also increases.

If **ρ** (A, B) = -1 there is a true negative relation between both A and B and if A increases Y decreases.

Even-though the correlation coefficient is -1 or +1 it does not have to imply that one causes the other. Causality refers to when one thing, often referred to the cause, gives rise to another cause which is often called the effect. It is important to differentiate association with causation, association refers to event which happen together more regularly than others, but that does not mean that the relationship is meaningful. In this chapter we will identify the correlation between public sentiment and stock price in general and if any of the two causes the other in particular.

Figure 3. Year wise apple stock performance

1	Date	Open	High	Low	Close	Adj Close	Volume
2	2012-01-03	58.485714	58.928570	58.428570	58.747143	50.857235	75555200
3	2012-01-04	58.571430	59.240002	58.468571	59.062859	51.130558	65005500
4	2012-01-05	59.278572	59.792858	58.952858	59.718571	51.698215	67817400
5	2012-01-06	59.967144	60.392857	59.888573	60.342857	52.238651	79573200
6	2012-01-09	60.785713	61.107143	60.192856	60.247143	52.155792	98506100
7	2012-01-10	60.844284	60.857143	60.214287	60.462856	52.342537	64549100
8	2012-01-11	60.382858	60.407143	59.901428	60.364285	52.257195	53771200
9	2012-01-12	60.325714	60.414288	59.821430	60.198570	52.113747	53146800
10	2012-01-13	59.957142	60.064285	59.808571	59.972858	51.918343	56505400
11	2012-01-17	60.599998	60.855713	60.422855	60.671429	52.523098	60724300
12	2012-01-18	60.994286	61.352856	60.900002	61.301430	53.068489	69197800
13	2012-01-19	61.450001	61.624287	60.930000	61.107143	52.900303	65434600
14	2012-01-20	61.070000	61.071430	59.964287	60.042858	51.978935	103493600
15	2012-01-23	60.381428	61.207142	60.328571	61.058571	52.858246	76515600
16	2012-01-24	60.728573	60.728573	59.935715	60.058571	51.992550	136909500

METHODS

Data Collection

To be able to perform the sentiment analysis and calculate the correlation, Twitter and stock data sets is necessary. The collection of these data sets is described and shown in figure 3 below

Stock Data Collection

The stock information has been gathered pandas_datareader library permits us to interface with the site and concentrate information straightforwardly from web sources, for this situation we are removing information from Yahoo Finance API AAPL US Equity.

In order to ensure a varied set of companies in different industries, the companies presented below were selected. American International Group AIG is a worldwide insurance agency with tasks in excess of 80 countries and locales. They provide a range of insurance products in business and in life, including: life insurance, general property and financial services. AIG's industry is Insurance (AIG 2019). Apple Inc. Apple designs and manufactures electronics such as phones, computers and smart watches and sells a range of related software. Apple's segments include the Americas, Europe, Japan, Greater China and Rest of Asia Pacific. Apples industry is Consumer Electronics [About Apple 2019].

World's biggest aviation organization is the Boeing Company. They are a creator of business jetliners, guard, space and safety frameworks, and high-quality group of subordinate vending support. Boeing upholds aircrafts and U.S. what's more, partnered government clients in excess of 150 nations. The Boeing's industry is Aerospace and Défense (About Boeing 2019). Facebook Inc. Facebook is a web-based web-based entertainment and informal communication administration organization. They are building items that empower individuals to interface through PCs, cell phones and different surfaces. Facebook's industry is Internet Content and Information (About Facebook 2019). Netflix Inc. Netflix is a web diversion administration and are usable in excess of 190 nations. Clients can sit in front of the TV series, narratives and element films across a wide assortment of types and dialects on any web associated screen. Netflix's industry is Media (About Netflix 2019). Home Depot Inc. Home Depot is a home improvement retailer. They sell Home Depot's industry is Home Improvement Stores (About Home Depot 2019).

Data Pre-Processing

The pre-processing stage is critical in order to obtain a high accuracy in the sentiment analysis [10]. Before the data is cleaned there are a lot of spam, special characters and other unnecessary data which not are required or can be classified in the analysis. Tweets were classified as spam if a user tweeted the same content several times. These tweets were deleted during this stage. Then blank spaces at the beginning and the end, usernames, punctuations, links, tabs and tweets made in other languages than English were deleted.

Sentiment Analysis

After the pre-processing stage, the data is stored as a Jupiter data frame since the following work also was done in python. Sentiment was selected to perform the sentiment analysis. The package is well used to calculate text polarity on a row-level. Sentiment also provides the opportunity to take valence shifters in consideration, e.g., negators, adversative conjunctions, amplifiers and de-amplifiers. The words of each sentence are compared to a polarity dictionary and then tagged with a number between -1 and +1. Sentiment have weights on valence shifters and the standard weights provided by the creator was chosen. The final result is the divided by the number of words in order to get a value which is not dependent on the length of the sentence.

Data Aggregation

When the sentiment analysis was completed, the daily sentiment was calculated for each stock during the period. This was done by calculating the mean value of all sentiments from all tweets during each day. One day is defined as one minute after the closing of the stock exchange until the next day's closing. The sentiment values were inserted into a matrix with two columns where column one contained the stock's daily sentiment, and where column two contained the close price of the stock on the corresponding day.

Correlation Calculation

Correlation is a proportion of affiliation or relation between two highlights for example the amount B will shift with a variety in A. The connection technique will make use of Pearson Correlation and it is shown in figure 4.

Using PC coefficient:

corr=df. corr(method='pearson')

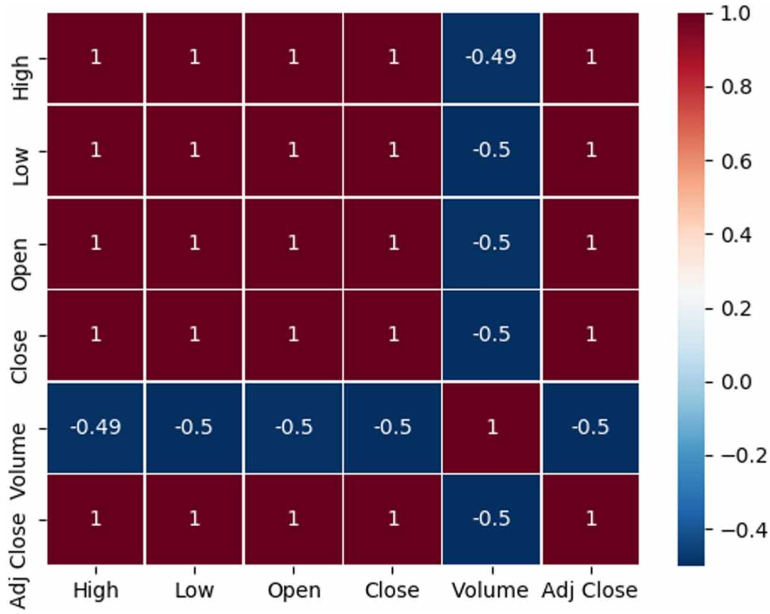
PC Coefficient is the most well-known method for estimating connection, the scope of values shifts from - 1 to 1. In arithmetic/physical science terms it tends to be perceived as though two elements are emphatically associated then they are straightforwardly relative and in the event that they share negative connection, they are contrarily corresponding.

EDA (Explanatory Data Analysis)

EDA indicates to the basic sequence of performing beginning inspections on data in order to find designs, to detect peculiarities, to test theory and to actually take a look at suspicions with the support of synopsis insights and graphical portrayals. The fair practice to fathom the data first and endeavour to collect as various encounters from it. Hence EDA is tied in with figuring out information close by, prior to getting them messy with it and it is certifiably not a conventional interaction with a severe arrangement of rules. More than anything, EDA is a perspective. During the underlying periods of EDA, we should go ahead and explore each thought that happens. A portion of these thoughts will work out, and some will be impasses. As your investigation proceeds, you will home in on a couple of especially useful regions that you'll ultimately review and convey to other people.

EDA is also a significant piece of any information examination, regardless of whether the inquiries are given to we on a platter, since we generally need to explore the nature of our information. Information cleaning is only one use of EDA. To do information cleaning, we want to convey every one of the devices of EDA: perception, change, and displaying.

Figure 4. Pearson correlation calculation



The detailed pictorial approaches used in EDA are regularly very forthright, comprising of different procedures of:

1. Charting the simple data, (for ex, data follows, histograms, bi-histograms, likelihood plots, slack plots, block plots, and Youden plots.
2. Charting upfront insights, for ex, mean plots, standard deviation plots, box plots, and primary impacts plots of the crude information.
3. Aligning charts in order to enlarge capacities, for example, utilizing various plots per page.

Below EDA shows the year wise Apple price prediction with dependent and independent variable.

Visualize the Dependent Variable With Independent Variable

See Figure 5.

Bar Plot of Open Price vs. Close Price (Year 2012)

Let us take a look at the bar plot of top 50 data which is from 2012 year

Figure 5.

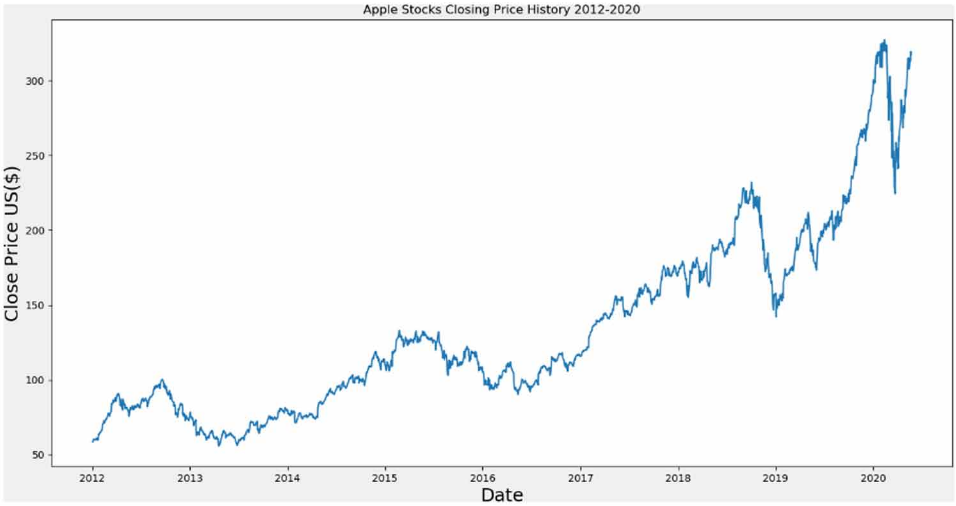
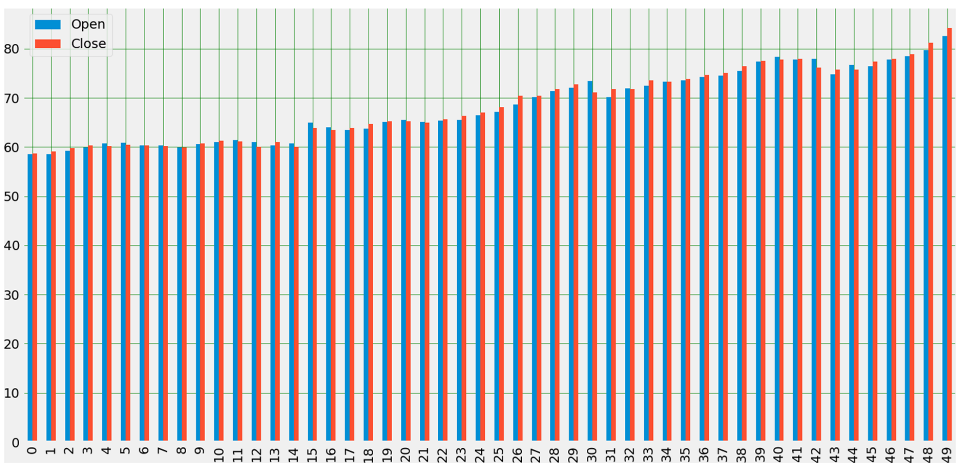


Figure 6.

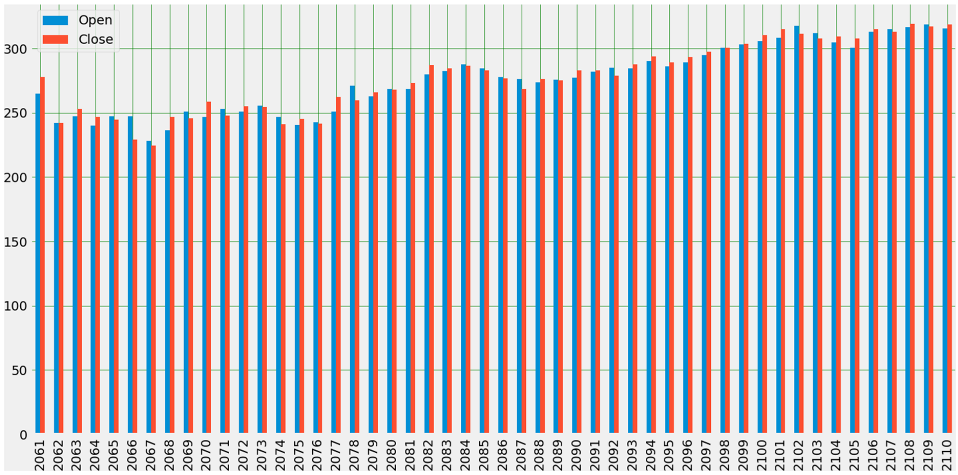


Bar Plot of Open Price vs. Close Price (Year 2020)

Let us take a look at the bar plot of top 50 data which is from 2020 year

All of the above bar plot represents year-wise open price and close price for the apple stock from 2012 to 2020. Now by applying a machine-learning algorithm to identify the rise and fall with high accuracy degrees.

Figure 7.



LINEAR REGRESSION MODEL

Separation of two variables to examine the one relationship and also it is used to measure the technical and quantitative analysis in market fields. Initially, we use Explanatory variables to predict the apple stock price along with average and correlation for making the prediction. Next is to define a dependent variable and it is the outcome for which the ML model will predict the stock price based on the explanatory variables.

The next step is to split the data into train and test. Train data is to create the linear regression by grouping the input and expected output. Test data is to show how well the model has been trained. Finally, cross-validation has been done for the model. Basically, Cross Validation is a procedure utilizing which a Model is assessed on the dataset on which it isn't prepared for example it tends to be test information or can be one more set according to accessibility or possibility.

number of splits: 20 and Accuracy: 99.99743780203187

Plot Actual vs. Predicted Value of Linear Regression Model

See Figure 8.

Figure 8.

KNN: K-nearest Neighbor Regression Model

KNN is used for both regression and classification, it finds the association between the independent variable and continuous result by averaging the observations.

It chooses the feature similarity to predict the values from new data points.

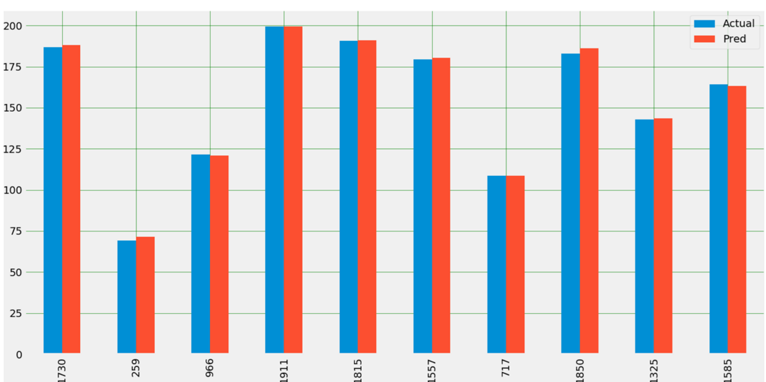
k neighbours = 4

Accuracy: 99.91435220285842

Plot Actual vs. Predicted Value of kNN Model

See Figure 9.

Figure 9.



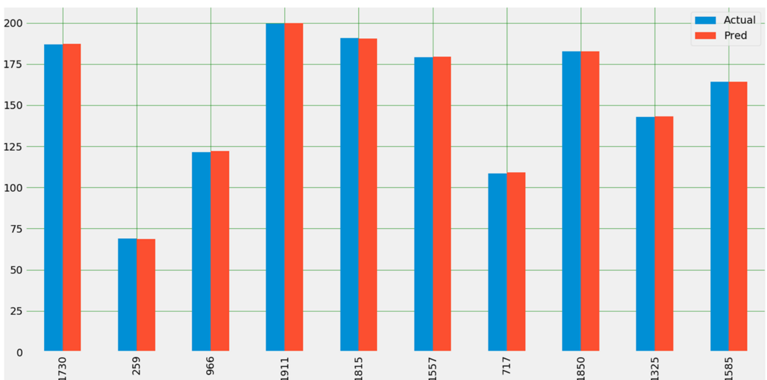
SVM Support Vector Machine Regression Model

Accuracy: 99.99301338392715

Plot Actual vs. Predicted Value of SVM

See Figure 10.

Figure 10.



RMSE (Root Mean Square Error)

RMS is the Standard Deviation of residuals, which are a proportion of how far information focuses are from the relapse. Or then again in straightforward terms how thought the information focuses are around the best fit line.

Linear Model RMSE: 3.0534992716871643e-14

KNN Model RMSE: 1.191675778610913

SVM Model RMSE: 0.5182098703394772

R-Squared Error

R-Squared score differs between 0 to 100%.

Mathematical Formula for R-squared score:(y_test[i] — y_pred[i]) **2

Linear R-Squared: 1.0

KNN R-Squared: 0.999629726665711

SVM R-Squared: 0.9999299807307482

CONCLUSION

One cannot say that there is a clear correlation between the Apple data and the stock price. In order to get more reliable results much more data would be needed and also a better method to clean up the data. There are some evidences suggesting that if the quality of the data is good then there is a correlation between the Twitter sentiment and the stock price. But in most cases, there are other factors driving the stock price. Also, the results suggests that the importance of sentiment is fluctuating and could increase if a significant event has occurred which might affect the stability of the correlation. There are no or little evidence showing that the correlation of Twitter sentiment and stock price has any relation to a specific industry or product type. Regarding the certainty of our results, the correlation seems to be significant, but the found results does not prove causality between the two. An investor should not solely base their analysis on these results but it could be used as an extra layer in their analysis improving accuracy or to monitor company sentiments in long term, considering the high correlation on the clean data sets.

REFERENCES

About Apple Inc. (2019). https://www.reuters.com/finance/stocks/company-profile/AAPL.O

Boeing, A. (2019). https://www.boeing.com/company/general-info/

Chang, Y.-H., & Lee, M.-S. (2017). Incorporating Markov decision process on genetic algorithms to formulate trading strategies for stock markets. *Applied Soft Computing*, *52*, 1143–1153. doi:10.1016/j.asoc.2016.09.016

Damodaran, A. (2014). *Session 7: Market Efficiency - Laying the Groundwork*. Academic Press.

Depot, A. H. (2019). https://www.reuters.com/finance/stocks/company-profile/HD.N

Facebook, A. (2019). https://www.reuters.com/finance/stocks/company-profile/FB.OQ

Focus, F. (2011). *Profit Margin Analysis. Third Quarter 2011*. http://search. credoreference.com.ezproxy.westminster.ac.uk/content/entry/ultimatebusiness/pr ofit_margin/0/

Harrington, C. (2003). *Fundamental vs. Technical Analysis*. *CFA Institute Magazine*. doi:10.2469/cfm.v14.n1.2789

Huang, Sinha, & Zhao. (2018). The Interaction between croblog. *Sentiment and Stock Return: An Empirical Examination*.

Netflix, A. (2019). https: / / www. reuters. com / finance / stocks / company -profile/NFLX.OQ

Nisar, T. M., & Yeung, M. (2018). Twitter as a tool for forecasting stock market movements: A short-window event study. *J. Finance Data Sci.*, *4*(2), 101–119. doi:10.1016/j.jfds.2017.11.002

Nofer, M., & Hinz, O. (2015). Using twitter to predict the stock market. *Business & Information Systems Engineering*, *57*(4), 229–242. doi:10.100712599-015-0390-4

Nofsinger, J. R. (2005). Social Mood and Financial Economics. *Journal of Behavioral Finance, 6*(3), 144–160.

Pagolu, V. S. (2016). *Sentiment Analysis of Twitter Data for Predicting Stock Market Movements*. https://arxiv.org/pdf/1610.09225.pdf

Renault, T. (2020). Sentiment analysis and machine learning in finance: A comparison of methods and models on one million messages. *Digit. Finance*, *2*(1), 1–13. doi:10.100742521-019-00014-x

Rowe, M., Stankovic, M., Dadzie, A.-S., & Hardey, M. (Eds.). F. °A. Nielsen. A new ANEW: evaluation of a word list for sentiment analysis in microblogs. In M. Rowe, M. Stankovic, A.-S. Dadzie, & M. Hardey (Eds.), *Proceedings of the ESWC2011 Workshop on Making*. Academic Press.

Schmidt, T., & Vosen, S. (2009). *Forecasting Private Consumption: Survey-based Indicators vs. Google Trends*. https://pdfs.semanticscholar.org/66f2/b2cfe56fc2414d3036cb6e28098bfa55216f.pdf

Sense of Microposts. (2011). Big things come in small packages. *CEUR Workshop Proceedings, 718*, 93–98.

Shang, Y., & Wang, Y. (2020). Study of CNN-based news-driven stock price movement prediction in the A-share market. In *International Conference of Pioneering Computer Scientists, Engineers and Educators*. Springer. 10.1007/978-981-15-7984-4_35

Smith, A., & Andersson, M. (2018). Social Media Use in 2018. Academic Press.

Szyszka, A. (2007). *Wycena papierów wartościowych na rynku kapitałowym wświetle finansów behawioralnych. Prace Habilitacyjne/Akademia Ekonomiczna w Poznaniu 35* [Valuation of Securities on the Capital market in the light of Behavioral Finance. Habilitation Theses]. University of Economics in Poznah.

Thelwall, M. (2009). MySpace Comments. *Online Information Review*, *33*(1), 58–76. doi:10.1108/14684520910944391

Tumasjan, A. (2010). *Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment*. Academic Press.

Yang, J. W., & Parwada, J. (2012). *Predicting stock price movements: An ordered probit analysis.* Academic Press.