**An Analysis of the Effect of News Sentiment on the Stock Market**

**Abstract**

The ready availability of data today enables investors of all levels of expertise to make better investment decisions. However data is siloed and the challenge is in determining which data is relevant, then ingesting and integrating the data into a usable format. The question is, can we use the content of current news to predict relative stock price performance? This research demonstrates that the prediction of percentage stock returns is improved by the inclusion of news sentiment as a feature for predictive model training. The results are demonstrated for both Apple and Microsoft stock using trending news articles. Six predictive models are built for comparative purposes.

**Business Problem**

Predictions on stock markets have been the object of research for many decades. The number of variables and sources of information that are to be considered is immense. This makes the predictions on the stock market a hard one. Financial news is a primary factor that investors have to consider during the process of financial decision-making. The core of stock market prediction is to forecast the opening price of the next day. Research scheme based on technical indicators analysis assumes that the behavior of stock has the property of predictability on the basis of its performance in the past and all the effective factors are reflected by the stock price. By analyzing technical indicators of the previous stock price, one could obtain important information which could be explored to forecast the following stock price. With the current advancement in computers, sophisticated forecasting methods can be implemented with ease.

Predicting the stock market is a complicated task. A particular stock has hundreds of events and preconditions to move it in a specific direction. To obtain the most reliable result, we need to capture as many of these preconditions as we can. Investors make educated guesses by analyzing the data. They analyze the news, and study the company's history and industrial trends. These data points will be used for making the prediction. The revealing theory is that stock prices are random and unpredictable. “ A blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by experts,” claimed Princeton University professor, Burton Malkiel in his bestselling book, A Random Walk Down Wall Street.

It has proven challenging to forecast stock prices using traditional time series methods. With the development of algorithmic and electronic trading, more investors have noticed that prediction of the stock market should focus not only on prediction accuracy but also on speed. Predictive accuracy for the designed model is the first and most needed requirement. This can help market traders make better choices. The speed of the real-time market transaction requires a speed of prediction. Over the past few decades, several approaches have been applied to classification and regression problems, such as ARIMA (autoregressive integrated moving average) ANN (Artificial neural network), and SVM. These approaches, however, focus mainly on prediction accuracy and are rarely concerned with speed.

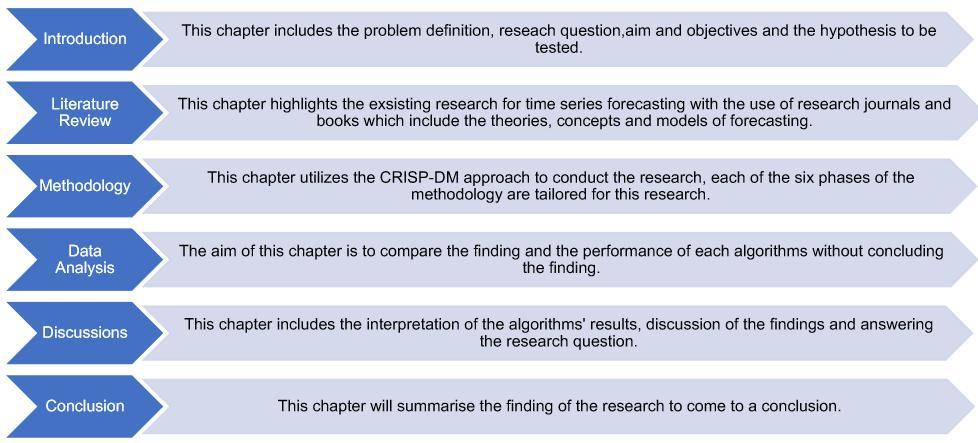
An artificial neural network is probably more suitable for this task, as no assumption of a mathematical model has to be made before the forecasting process. Also, the neural network can extract the relevant factors from big datasets. So our research question is "Is stock market and financial news related to each other? Can news articles be used for predicting the stock market?"

**Research Problem**

There are fast changes in the stock market. It's very hard to predict what's going to happen to a specific company's stock or to the stock market in general. When the stock market trends can be predicted for the day itself or the next day, a substantial sum of money can be saved. Seeing a link between news headlines and the trend in the stock market is not unusual. If a business receives adverse press, it is understandable that as a result of this adverse news, the stocks of this specific business will go down and vice versa. It may be feasible to predict the stock market trend in general or that of a specific business by evaluating news headlines.

*Roadmap*

This thesis covered forecasting models, the algorithms used within the model, and other optimization techniques used for better performance and accuracy for the stock market. The primary focus is to find the impact of news headlines in the stock price. Here we are predicting opening prices using the closing price, index, and news. We also attempted to identify trends and patterns in the data using various algorithms. For this RNN (Recurrent Neural Networks) approach, LSTM (Long Short Term Memory) and traditional machine learning algorithms Linear Regression, SVM, Random Forest, K Nearest Neighbor, AdaBoostRegressor, and Principal component analysis (PCA) were used. The following roadmap is used to develop a strategic plan to implement the thesis project.



***Goals to Accomplish***

* Collect and preprocess historical stock market data, financial news headlines, and other relevant features.
* Train and evaluate different forecasting models, including RNN-LSTM, Linear Regression, SVM, Random Forest, K Nearest Neighbor, AdaBoostRegressor, and PCA with Random Forest.
* Identify the impact of news headlines on stock market trends and assess the predictive power of the selected models.
* Optimize the selected model(s) for accuracy and speed to meet the real-time requirements of market transactions.
* Create a reliable prediction model that can forecast the opening prices of stocks based on news headlines and other factors.

**Literature Review**

The primary objective of the financial world is to forecast stock prices to guarantee the highest possible gain. Predictive speed is also a significant variable in stock price forecast. Sometimes even a mini-second lag can lead algorithmic trading to invalidate the market trading strategy. As a consequence, the substantial and challenging problem in the financial market is to develop a stock market prediction model that can deliver excellent efficiency on accuracy and speed. Data scientists have shown that business news can assist stock market prediction will provide more precision. Many studies have been done in the field of stock prediction, where a common strategy is to use Twitter texts to estimate stock price. Some of the studies conducted in this area focus on mixing stock prices with news headlines

Unlike cost information, data from business newspapers is unstructured and high-dimensional. The choice of features and weighting techniques are the essential components of the information handling of news on the market. Traditional methods of selected characteristics, such as Chi-Square and gaining information, disregard the word function frequency. Chi-square, therefore, tends to choose unusual conditions that are often less reliable. Also, the characteristics are unbalanced but essential to forecast for financial newspapers.

In this age of big data, large volumes are produced or gathered at high speed from a wide variety of valuable information of varying veracity. The stock market is a wealthy source of this significant information. People have been attempting to "beat" it for financial gain since the start of the stock market. A stock market is an exchange in which individuals trade company shares, also known as stocks. The exchange aims to facilitate transactions between buyers and sellers. The common objective of someone taking part in the stock market is to create profit by purchasing and selling stocks. The primary way individuals do this is by purchasing a stock, waiting from seconds to days to months to years, and then hopefully selling for more than they've purchased it for. Here comes the famous phrase "purchase low, sell high".

The stock price is the price at which the last trade took place. If one person sells stock to another person, it becomes the new stock price that they agree. The stock price, therefore, depends entirely on supply and demand. The higher a stock's demand, the higher its price. If more individuals try to sell the stock rather than purchase, the price will fall. Because of this, predicting the market may not be simple because it is based entirely on human choice, and these choices may not be reasonable at times. People create a choice based on many factors, including earnings, reports, news, financial factors, competition, technical analysis, and even gut feeling. It is, therefore, almost impossible to design an algorithm to take all these variables into consideration. By using the trading technique (i.e. purchasing a stock with price hopes rising and then eventually selling the stock) mentioned above, the equation can calculate its profit,

Profit = sell price – buy price

An alternative profit-making technique is brief selling or short selling. The concept is to invest as the stock is predicted to go down. One borrows stocks and sells them before purchasing them to conduct shorting. Since one has sold the stocks that he has not yet, he has to purchase the stocks to cover. That's where he's waiting and hoping the price will fall. But the operations take place in the reverse order.

Another way is through dividends to create profit from the stock market. Dividends are part of a company's earnings to its shareholders in the form of money to own its stock. The profit can thus produce irrespective of the share price. However, there is no guarantee that dividends will be consistent and subject to change at any time. If the business faces tough times, the dividend can be sacrificed to remain afloat and pay off present debts/bonds. Dividends are generally paid out of big established firms with constant to the revenue and not growth-focused. Companies with volatile revenue or attempting to develop may need to maintain their revenue to assist them to use their revenue later to support businesses afloat .Following is a discussion on previous research on sentiment analysis of text data and different classification techniques.

(Kalyani et al. 2016) collected data for Apple over three years. Also, news information about this company gathered over the same period. Several news websites, such as news.google.com and finance.yahoo.com, were used. First, the news data pre-processed, and an algorithm for sentiment identification was used to identify the sentiment of this news article. The sentiment detection algorithm worked with a positive and negative word dictionary in which the news

**Methodology**

The study was carried out using the methodology CRISP-DM (Cross-Industry Standard Data Mining Process). This methodology offers a study framework that helps the findings to be better and quicker. The CRISP-DM methodology organizes the study into six stages, these stages assist to better comprehend the process and provide a road map for planning and conducting the research. The following figure shows the CRISP-DM model phases. The phases of the CRISP-DM methodology are as follows:

1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modeling
5. Evaluation
6. Deployment

*Business Understanding*

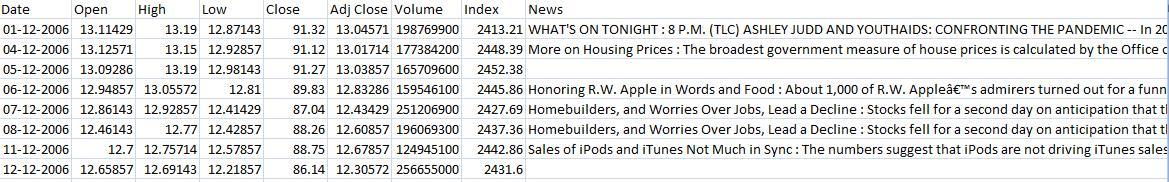
The first and most significant stage of the research project is the business understanding, this stage is aimed at understanding the goals of the project from a business viewpoint and turning this business view into a definition of study problems and then creating a plan for achieving these goals. Now, several factors can influence market movements. With the power of social media and the availability of E-news papers, the market reacts to news events almost instantly. Reacting slowly for a certain event might lead to a loss of profit or even a negative PNL (Profit And Loss). In order to maintain portability, one cannot merely rely on human labor in understanding the impact of the event.

*Data Understanding*

The data understanding phase starts with data collection. Once the required data is acquired, the data is explored to identify data quality problems if any, and insights on the data are gathered. To find the impact of news in stock market prediction, a dataset made available on Kaggle is used. The dataset contains the historical stock data for two different stocks, Apple and Microsoft. The dataset contains the historical data for two different stocks that are Apple and Microsoft here. The dataset contains over 2518 daily records of each stock over a period of ten years from 2006 to 2016 and 8 columns. From the literature review, it is understood that the composite index of the exchange will have an impact on the stock price. So the index is also considered for the prediction. The historical data was downloaded from the official website of NASDAQ.

*Data Preparation*

The data preparation phase consists of all the activities required to process the data from the initial raw form before it is fed into the model. These tasks include everything from attribute selection to transforming and cleaning the data. Text data is unstructured data. So, we cannot provide raw test data to the classifier as an input. The structure of data in the CSV file is displayed below

**Fig 3.1 The structure of data in the CSV file**

Pre-processing using NLTK

The first step is to prepare input data for the models, which involves removing stop words, punctuations, and unnecessary numbers. For the purpose of this experiment, we decided to use Natural Language Toolkit (NLTK) from the SKLearn library to help with the process. Word embedding is the collective name for a set of language modeling and feature learning techniques in natural language processing (NLP) where words or phrases from the vocabulary are mapped to vectors of real numbers.

Initially, we need to tokenize the document into words to operate on the word level. Text data will be noisy. So, we need to drop the number of words. In addition, text data may contain numbers, unwanted white spaces, tabs, punctuation characters, stop words, etc. We also need to clean data by removing the unwanted. For this purpose, an own stop-word list that specifically contains stop words related to the finance world and general English is created. This list contains general words including Generic names, Dates and numbers, Geography, Currencies, etc. The transformed data is then compared with the words in a predefined dictionary, which consisted of specific words and phrases and their associated polarity strength in terms of sentiments. To ignore words that appear in only one or two documents, we employ minimum document frequency which considers words that appear in minimum three documents. Null values in the news column were replaced by the news of the previous day.

*Data Pre-processing using Glove*

The transformed data is then compared against words in a predefined dictionary which consisted of specific words and phrases and their associated polarity strength in terms of sentiment. Also, to ignore words that appear in only one or two documents, we are considering minimum document frequency which considers words that appear in a minimum of three documents. Lemmatization is also important to reduce redundancy in words. Using the lemmatization process, all the words are replaced by their original version of the word.

For example, the words ‘developed’, ‘development’, and ‘developing’ is reduced to their stem word ‘develop’. Some of the preprocessing is done before feeding into the algorithms.

Stanford NLP GloVe Wikipedia Gigaword 3-dimension embedding set was used since it was trained on Wikipedia and would likely have vectors for many of the industry-specific words found in the reports. Glove 300d has 400000 words in it. The global word representation approach is used to capture the meaning of one word embedding with the whole observed corpus structure; word frequency and co-occurrence counts are the main measures on which most unsupervised algorithms are based. GloVe model trains word counts on global co-occurrence and makes efficient use of statistics by minimizing error in the smallest squares and thus producing a word vector space with a meaningful substructure. Such an outline preserves word similarities with vector distance sufficiently.

*Modeling*

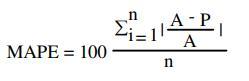
In the modeling phase, several models selected based on thorough research are applied. Each of the models has a specific requirement on how the data must be pre-processed, hence the need to step back into the data processing phase. This phase includes the selection, creation, and assessment of the models. In this research, the models are initially implemented on the training set list to forecast for the next day These forecasted values are compared with the test set list values, and the test accuracy of the models is computed. Once the test errors are computed, the entire data set list is fed into the models for the forecast. The best models for forecasting are selected based on the research conducted in this field:

*Linear Regression*

In statistics, linear regression is a linear approach to modeling the relationship between a scalar response and one or more explanatory variables. The case of one explanatory variable is called simple linear regression. The method is called multiple linear regressions for more than one explanatory variable. There may be many independent variables falling under the multiple linear regression categories. We have only one independent variable in this situation, which is the date. The date will be portrayed by an integer beginning at 1’s for the first date up to the date's vector length which may differ depending on the information from the time series. Naturally, our dependent variable will be the stock price.

*SVM (Support Vector Regression)*

SVM regression performs linear regression in the high-dimension feature space using ∈− insensitivity loss and, at the same time tries to reduce model complexity by minimizing || w || 2. This can be described by introducing slack variables ∈ and ∈ \* where i = 1., n to measure the deviation of the training sample outside ∈− sensitive zone. SVM generalization performance depends on a good setting of kernel parameters C, ∈ and kernel parameters. To evaluate the result from the SVR models, the following equation is used to measure the error rate.



Here MAPE stands for Mean average percentage error between actual share prices (A) and predicted share price (P). And ‘n’ is the number of days to take into count.

*Random Forest*

Decision trees can be used for different apps for machine learning. But trees that grow deep to learn extremely uneven patterns tend to overfit the sets of practice. A slight noise in the information may cause an entirely distinct growth of the tree. This is due to the very small bias and elevated variance of decision trees. Random Forest overcomes this issue by training multiple decision trees at the expense of mildly enhanced bias on distinct subspaces of the feature space. This means that the entire training data is not seen by any of the trees in the forest. The information is divided into partitions recursively. The split is performed at a specific node by asking a query about an attribute. The choice for the criterion of splitting is based on certain measures of impurity such as Shannon Entropy or impurity of Gini.

*K-nearest neighbor*

K-nearest neighbor is an algorithm of machine learning regarded as easy to execute. The issue of stock prediction can be mapped into a classification based on resemblance. A set of vectors maps the historical stock information and the test information. For each stock feature, each vector reflects the N dimension.

Then, to make a choice, a similarity metric like the Euclidean distance is calculated. A description of KNN is given in this chapter. KNN is considered to be a lazy learning process that does not previously build a model or function but provides the closest k record of the training data set with the highest similarity to the test (i.e. query record).

Then, among the selected k records, a majority vote is carried out to determine the class label and then assign it to the query record.

The stock market closing price forecast is calculated using kNN as follows:

* Determine the number of neighbors closest to you, k.
* Calculate the distance to the query record between the training samples.Classify all training documents by range.
* Use the majority vote for the class labels of the closest neighbors and assign it as the predictive value of the query record.

Root Mean Square Deviation (RMSD) is an accuracy metric that computes the differences between the estimated values, Y, and the actual values, X. The total of RMSD is aggregated into a single value measure. Average Estimated Error (AEE) AEE is the total sum of RMS errors for all variables in stock records divided by the total number of the records.



*AdaBoost*

PAC learning model is a theoretical framework for boosting machine learning. To fix many practical problems experienced by the previous boosting algorithms, the AdaBoost algorithm was suggested. Assume that the practice set represents samples for training ∈ is the independent variable and X denotes the domain space. ∈ variable and Y represents the value domain. For the issue of binary classification, Y is presumed to be[ -1, +1]. AdaBoost calls iteratively the weak learning algorithm for round T. While it is necessary to maintain a distribution or set of weights over during the process. Let ( ) denote the weight of the sample i distribution at round t(t=1,...,T), In the iteration method, ( ) I t will be reduced for misclassified samples in the following round, forcing the following weak learner to pay more attention to the hard samples.

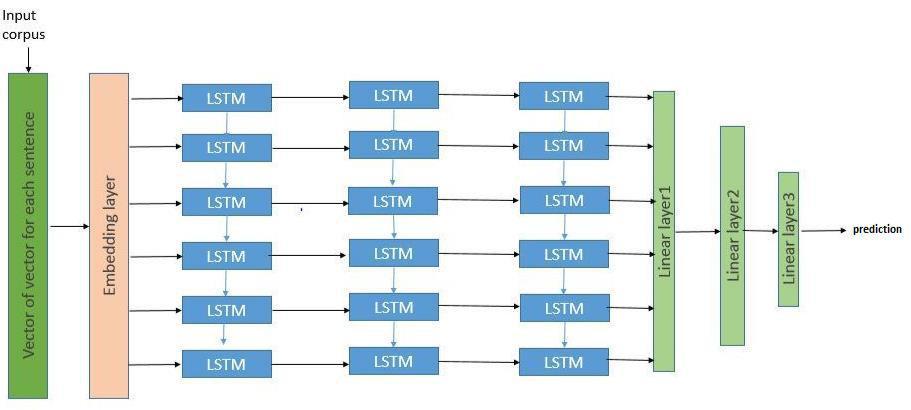
In our project, we have used an LSTM (Long Short-Term Memory) algorithm to detect fake news. LSTM is a type of recurrent neural network that is particularly good at processing sequential data, such as text. It has been successfully used in several natural language processing tasks, including sentiment analysis, text generation, and machine translation. The need for using machine learning algorithms in our project is to develop an automated tool that can detect fake news with high accuracy. Manual detection of fake news is time-consuming and prone to errors, particularly when dealing with large amounts of data. Machine learning algorithms can process data much more quickly and accurately than humans, making it possible to detect fake news in real-time.

Recurrent Neural Networks (RNNs) are a type of neural network that are designed to work with sequential data, such as time series or text. They are particularly well-suited for tasks such as language modeling, speech recognition, and sentiment analysis. In our project, we have used a variant of RNN called Long Short-Term Memory (LSTM) to detect fake news. The need for using RNN algorithms in our project is because fake news detection involves analyzing text data, which is inherently sequential. Each word in a news article is related to the previous words and can influence the meaning of the entire text. RNNs are designed to capture this temporal relationship and can model the complex dependencies between words in a text.

LSTM is a type of RNN that is specifically designed to handle the problem of vanishing gradients that can occur when training traditional RNNs. Vanishing gradients can make it difficult for RNNs to learn long-term dependencies in the data, which can limit their effectiveness. LSTM solves this problem by introducing a memory cell that can store information over long periods of time and selectively forget or update this information based on the input. In our project, we have used LSTM to classify news articles as either fake or real. We have trained our LSTM model on a dataset of labeled news articles, with the goal of learning the patterns and features that distinguish fake news from real news. To achieve this, we have preprocessed the text data to remove stop words, punctuation, and other noise, and then converted the text into a numerical representation using techniques such as word embedding.

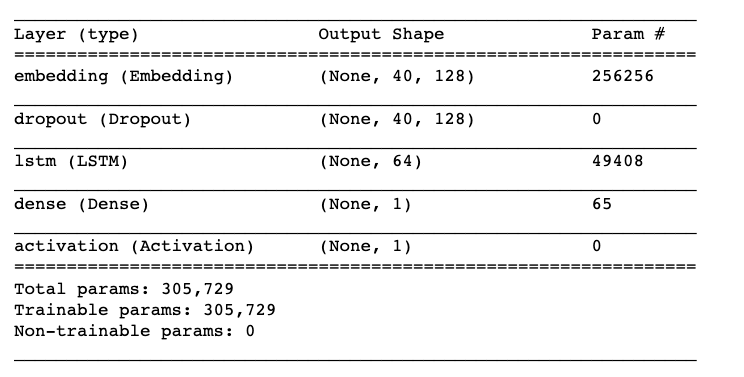
Once the text data has been preprocessed and converted into a numerical representation, it is fed into the LSTM model, which processes the data sequentially and learns the patterns and features that distinguish fake news from real news. The output of the LSTM model is a probability distribution over the two classes, which we use to classify the news article as either fake or real. One of the advantages of using LSTM in our project is that it can handle variable-length sequences of text. This means that our model can process news articles of different lengths and still make accurate predictions. LSTM also has the ability to remember long-term dependencies in the data, which is particularly useful in the detection of fake news, as some fake news articles may contain subtle cues or hints that can only be detected over longer sequences of text. the use of RNN algorithms such as LSTM is essential in the detection of fake news. These algorithms are designed to handle sequential data and can learn the complex patterns and dependencies between words in a text. In our project, we have used LSTM to develop an automated tool for detecting fake news with high accuracy. This tool has the potential to be useful for journalists, fact-checkers, and social media platforms in identifying and removing fake news from circulation.

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feed-forward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks such as nonsegmented, connected handwriting recognition or speech recognition Bloomberg BusinessWeek wrote: "These powers make LSTM arguably the most commercial AI achievement, used for everything from predicting diseases to composing music." A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.



**Fig 5: LSTM Architecture**

LSTM networks are well-suited to classifying, processing and making predictions based on time series data since there can be lags of unknown duration between important events in a time series. LSTMs were developed to deal with the exploding and vanishing gradient problems that can be encountered when training traditional RNNs. Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models and other sequence learning methods in numerous applications. LSTMs are very powerful in sequence prediction problems because they’re able to store past information. In the context of fake news detection, LSTMs can be used to analyze the sequential structure of news articles and identify patterns that are indicative of fake news. The LSTM model can be trained on a labeled dataset of news articles, with the labels indicating whether the article is real or fake news.



**Fig 6: LSTM Model Summary**

The LSTM takes in the text of the article as input, processes it sequentially, and produces an output that indicates the probability that the article is fake news. This output can be thresholded to make a binary classification decision. These texts are pre-processed and the entry removes all non-alphabetical characters. To overcome double identical phrases with distinct capitalization, all remaining characters are transformed into lowercase inputs. The top 200 words in the corpus are used for word embedding to convert the words to the correct index. All other words not in the top 200 phrases are set to a zero index. The maximum word embedding size for lower headlines is 100 and zero-padded.

**Evaluation**

In this phase, each of the models is evaluated to check if they are capable of achieving the business objectives. The data is split into two ratios and is fed into the models. Once the model’s results are evaluated, the best model with less error was used to verify whether the news is having any impact on stock market prediction. We split the data into two different proportions 80:20 and 90:10 and the results showed that the deep learning model LSTM is more capable of predicting the result. The mean square error obtained from each algorithm is tabulated in section Results Analysis and Discussion

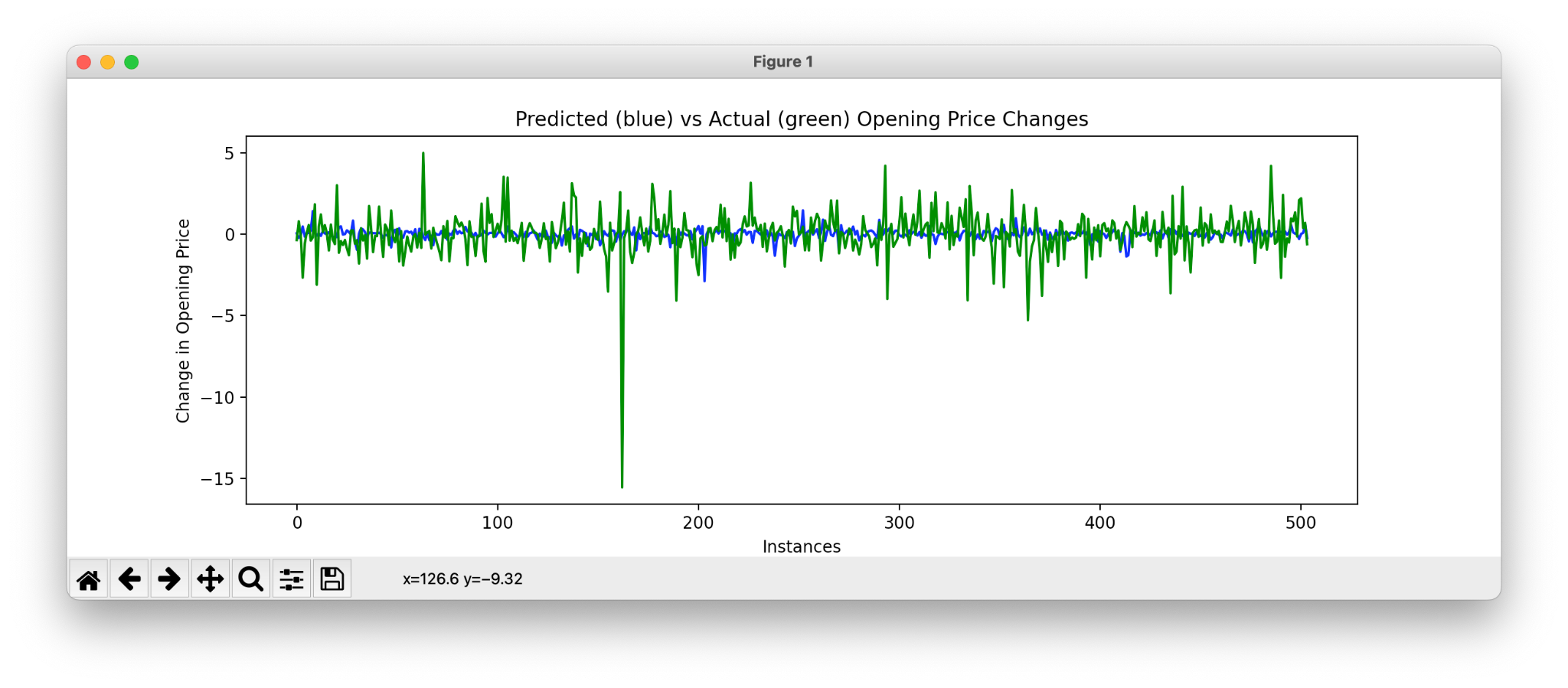
*MSE Analysis (Mean Squared Error)*

The mean squared error or mean squared deviation of an estimator (of a procedure for estimating an unobserved quantity) measures the average of the squares of the errors, that is, the average squared difference between the estimated values and the actual value. MSE is a risk function, corresponding to the expected value of the squared error loss. The fact that MSE is almost always strictly positive (and not zero) is because of randomness or because the estimator does not account for information that could produce a more accurate estimate. The MSE is a measure of the quality of an estimator. it is always non-negative, and values closer to zero are better.

Once the sample data analysis was completed then the next step was an evaluation of the models. Here, 6 models were applied which were Linear Regression, SVM, Random Forest, K Nearest Neighbour, AdaBoost, Principal component analysis. The model's result has been analyzed by MSE value. MSE values for all 6 models for Microsoft and apple are listed here.By analyzing the result, the following facts we found:

| **Algorithm** | **Mean Square Error** |
| --- | --- |
| Linear Regression | .030405 |
| SVM | .034847 |
| Random Forest | .028881 |
| K Nearest Neighbors | .030247 |
| AdaBoost | .031623 |
| Gradient boost | .029021 |

* Higher embedding dimension gives better prediction: 300d gave the best result
* News with longer length results in higher prediction accuracy: setting the truncated length of news to 250 rather than 200 results in higher accuracy because it includes more text information in the training example. But when the sentence becomes longer the time for training also increased



**Conclusions**

This study has proven that news is having a substantial effect in identifying stock market trends. Here in this research, I have gone through the extensive procedure of various machine learning algorithms for stock price prediction. The examinations of these models uncovered that the stock data set of both the two products was distinguished by each algorithm. The test results acquired with the Gradient and Adaboost model showed the capability to predict stock prices satisfactorily on a short-term basis. This could direct speculators at stock prices to settle on gainful investment decisions. With the outcomes acquired Gradient and Adaboost model can contend sensibly well with other methods in short-term prediction. The results suggest that the use of word-level embeddings is promising and competitive with more complex models which use technical indicators and event extraction methods in addition to the news articles. word embedding models are simpler and more memory efficient.

**Referencing**

1. A.E. Khedr, S. E. Salama, and N. Yaseen, “Predicting stock market behavior using data mining technique and news sentiment analysis”, IJISA , 2017.
2. Anurag Nagar, Michael Hahsler, Using Text and Data Mining Techniques to extract Stock Market Sentiment from Live News Streams, IPCSIT vol. XX (2012) IACSIT Press, Singapore
3. D. Enke and S. Thawornwong, “The use of data mining and neural networks for forecasting stock market returns,” ESWA, 2015
4. D. Kirange and R. R. Deshmukh, Sentiment analysis of news headlines for stock price prediction,2016
5. J. Bean, R by example: Mining Twitter for consumer attitudes towards airlines, In Boston Predictive Analytics Meetup Presentation, 2011
6. J. Ting, T. Fu, and F. Chung. “Mining of stock data: intra- and inter-stock pattern associative classification” IEEE ICDM 2006
7. R. Agrawal and R. Srikant, “Mining sequential patterns,” in Proc. IEEE ICDE 1995
8. R. Agrawal and R. Srikant, “Fast algorithms for mining association rules in large databases,” VLDB 1994
9. Richard C. Camara, Alfredo Cuzzocrea, Giorgio M. Grasso “Fuzzy Logic-Based Data Analytics on Predicting the Effect of Hurricanes on the Stock Market “2018 IEEE International Conference on Fuzzy Systems , 2018
10. Tong Yang,Yuxin Yang “Predict Effect of Trump’s Tweets on Stock Price Milestone“,2017
11. W.B. Yu, B.R. Lea, and B. Guruswamy, A Theoretic Framework Integrating Text Mining and Energy Demand Forecasting, International Journal of Electronic Business Management. 2011
12. Yauheniya Shynkevich, T.M. McGinnity, Sonya Coleman, Ammar Belatreche, “Predicting Stock Price Movements Based on Different Categories of News Articles, 2015 IEEE Symposium Series on Computational Intelligence”, 2015
13. Y. Wu, K. Wu, and H. Lee. “Stock trend prediction by sequential chart pattern via k-means and AprioriAll algorithm,” 2012
14. Z. Hu, J. Zhu, and K. Tse, “Stocks market prediction using support vector machine,” in Proc. ICIII 2013