Analyzing News Sentiments and their Impact on Stock Market Trends using POS and TF-IDF based approach

Sonam
School of Information Technology
Mapua University
Manila, Philippines
sonam@mymail.mapua.edu.ph

Madhavi Devaraj
School of Information Technology
Mapua University
Manila, Philippines
mdevaraj@mapua.edu.ph

***Abstract*—Since the dawn of time, investors are looking into different schemes in determining the stock trends to earn profit. Several studies have been conducted that could potentially help the investors predict the rise and fall of stocks. Most of them looked into past market pricing history in order to foresee the future. While many factors influence the fluctuation of stock market, it can be argued that the sentiments of the investors influenced by unfolding of current happenings or events has a huge impact on the stock trend. In this paper, we propose a new method in interpreting the sentiment of a given news. Through a fine-grained analysis of syntactic sentence patterns using different Part of Speech (POS) combinations, the news data inputs are preprocessed. These are then fed into Term Frequency – Inverse Document Frequency (TF-IDF) to filter only significant text in the corpus. We then conduct experiments using various classifiers to predict the sentiments. Results are fed into K-Nearest Neighbor (K-NN) classifier, along with historical stock price, to determine adjusted closing price over various time intervals. It can be observed that the results of proposed model are compatible with current researches stating about existing correlation between financial news and stock prices.**

Keywords—machine learning; natural language processing; sentiment analysis; stock markets

I. INTRODUCTION

In the quest of seeking maximized returns, stock investors attempt to foresee and predict the future trends of the market which, meanwhile, is thought-provoking due to the innate nature of stocks being highly volatile and unstable [1]. The behavior of time series data from financial markets is influenced by a mixture of quantitative information from system dynamics captured in its past behavior, such as historical prices and turnover rate, and qualitative information about underlying fundamentals, which include social media posts, annual reports, and news feeds [2]. Combining these two types of information in recognizing patterns in financial data is capturing interest in wider academic disciplines.

Recent studies are focusing on utilizing news-derived information to foresee the direction of stock movement or the precise value of a future asset price [3]. Using various methods, most of the studies are able to prove the dependence of news and stock trends. Some studies have used financial news [4], while others have used economic news [5]. Other studies, nevertheless, have used social media like Twitter [6] and Facebook [7] to investigate how people's sentiments relayed in social media posts affect shares. Yet, other researches have investigated how search queries in Google [8] and other search engines may affect the direction of stocks.

II. RELATED WORK

Numerous studies have looked into the impact of news sentiments on stock market volatility. Some researchers used lexical resources for analysis of news content and established the relevance of stocks to the resulting sentiments from the analysis, while others used machine learning classifiers in the process.

Some studies looked into using different machine learning classifiers in assigning sentiments to news articles. The headlines, including the whole news body, are used as inputs into standard text cleaning procedures for preprocessing the input dataset. The Term Frequency – Inverse Document Frequency (TF-IDF) approach is used for vector representation, followed by the application of Naïve Bayes, Random Forest, and Support Vector Machine (SVM). The resulting news sentiments are then correlated with stock trend using a time series plot [9].

Nevertheless, one study looked into a different perspective of establishing the correlation between financial news and stock market. The motive of their study is to help the managers, in media or press, in their dilemma of determining which news topics actually matter to stock market investors. The news corpus used for the study originates from German regulated ad hoc announcements. In text preprocessing, the following methods were performed: noise or unneeded content removal, stop word removal, stemming, and document-term matrix (TF-IDF). This is followed by the use of Latent Dirichlet Allocation (LDA) probabilistic model for extraction of topics with highest probabilities and assigning these to each of the announcements in our corpus. In order to find the effect to stock market, the abnormal returns for each stock are calculated to eliminate confounding effects. Based on results, those topics having no resulting impact on abnormal returns of stocks are determined, and other topics that have a large impact on stock market returns, such as drug testing [10].

Yet another study analyzed sentiments of news dataset using conventional methods, which are as follows: tokenization, data standardization, stop-word-removal, stemming, abbreviation processing, token filtering, n-gram feature extraction, and application of term frequency inverse document frequency (TF-IDF). The Naïve Bayes classifier is then applied to classify the news as either positive or negative based on TF-IDF input values [11].

The problem with the study of Khedr *et al.* is the input dataset being used. Since the whole news text is used (headline and body), there maybe a lot of words that are unnecessary but

978-1-7281-6946-0/20/$31.00 ©2020 IEEE

are still being included in latter processing. This may also take longer time in the processing. Moreover, the unnecessary words may get added as noise and affecting the overall accuracy of the proposed algorithm.

Due to this, the technique employed by a similar study [12] could be incorporated in the preprocessing part of the algorithm. Particularly, the fine-grained extraction of part of speech tags, such as verbs, adjectives, and nouns, based on syntactic patterns of sentences could be used to include only the specific words in the input dataset. This could possibly improve the performance of the algorithm in analyzing sentiments, thereby resulting into a more accurate comparison of news sentiments with regards to stock fluctuation.

The proposed research will primarily benefit the current and potential stock market investors. The investors would know how much they should rely on news feeds when making stock-related decisions. Also, if the results show a strong correlation of news with stocks, the investors can make news as one of their primary sources in predicting the future trend of stock market.

III. Proposed Model

The objective of the research is to develop an algorithm that would analyze the impact of news feeds on stock trends. The study proposes the utilization of various fine-grained syntactic Part of Speech (POS) combinations when parsing news articles. The resulting deciphered sentiments are used together with historical opening, highest, and lowest stock market prices for each day in determining the adjusted closing price of stock indices. The stock indices that are being looked into are the top 5 companies that account for 17.5% of S&P 100, which are Apple, Microsoft, Google, Amazon, and Facebook [14]. The proposed model takes two set of inputs; the first set comprises the sentiments obtained from sentiment analysis model, and the second set contains the stock market historical prices. It looks into predicting the adjusted closing price in various time intervals using historical stock prices and news sentiments, as depicted in Fig. 1. The predicted news sentiments from news articles are joined with stock market data using date, as shown in Fig. 2.

IV. Experimentation

A. News sentiment analysis

1) News data gathering: The news articles are obtained through New York Times (NYT) application programming interface (API). The study conducted by Garcia-Medina *et al.* used dataset from NYT and the results indicate a significant correlation between news and stocks [13]. A python program is developed to obtain articles through the API starting from 1st of January, 2013 until 31st of December, 2017. Those articles under Business and Finance categories are considered. Named Entity Recognition (NER) is used to filter articles to be gathered. Only those articles with headlines containing Organization NER category label are obtained. The number of articles gathered per day is limited to three (3). Annotation of news articles is done by a stock investor. Each article is tagged as positive or negative. The annotated news articles are then input into the model for data preprocessing.

2) News data preprocessing: In order to make the data suitable for further processing, data is preprocessed. The data preprocessing comprises of converting the text into lowercase, removing of numbers and punctuation marks, removing of stop words (e.g. ours, you're, himself, her, against, etc.), stemming or reducing a word into its word stem by cutting off the prefix or suffix of the inflected word, and lemmatization or converting the word into its base form through morphological analysis.

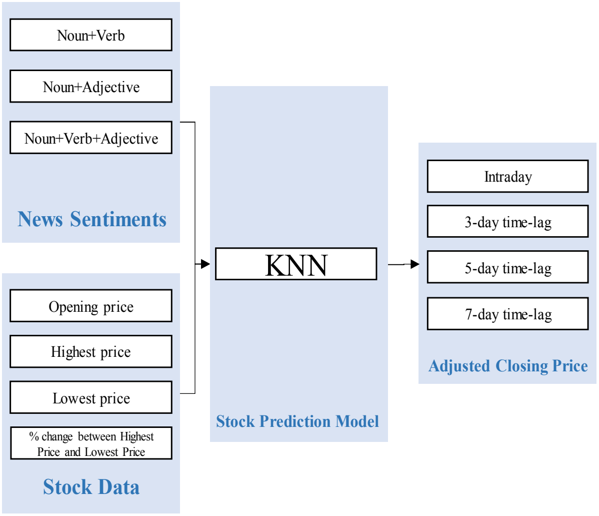

Fig. 1. Set of inputs and expected outputs of the stock prediction model.

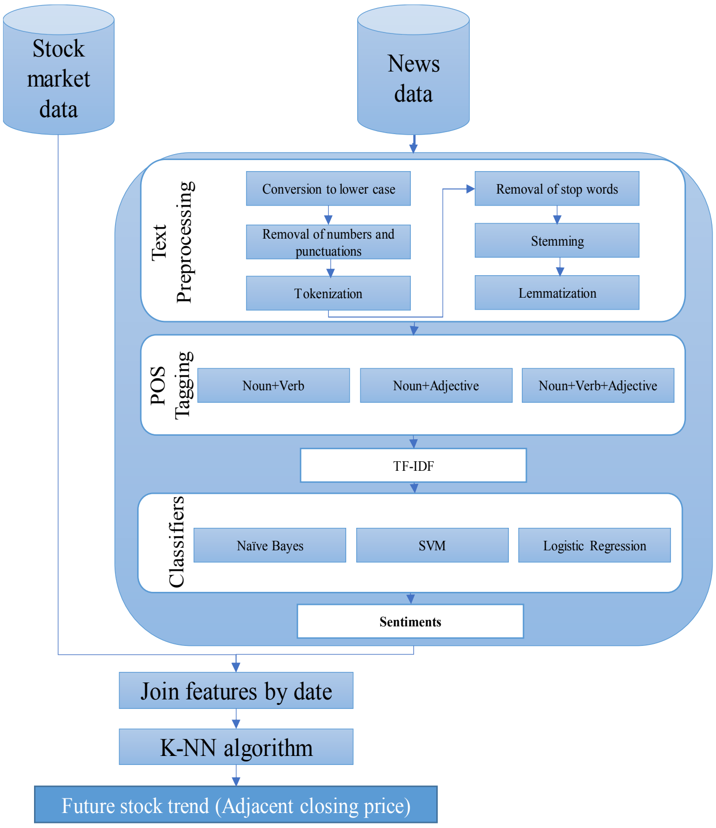

Fig. 2. Sentiment analysis model output, joined with stock market data using date, to be used as input into stock prediction model.

3) POS tagging: Each of the words in the preprocessed text are then marked to their corresponding part-of-speech. The tagging of words is based on the process employed by

Meyer *et al.* in their study. Different combinations of noun, adjective, and verb are used. The purpose of this is to determine which combination is optimum or contains the most amount of information that could be used for text analysis. In addition, this can somehow diminish the text to be processed in the next steps of the algorithm, thereby helping to reduce the overall processing time. For the purpose of this study, the following combinations are analyzed: Noun + Verb, Noun + Adjective, and Noun + Verb + Adjective [15].

4) Vector representation: In order to transform text into a vector, Term Frequency – Inverse Document Frequency (TF-IDF) is used. The vector is used to depict the important attributes of the input text. The vector representation gives high values for a term if the term occurs frequently in in the specific document, yet very seldom elsewhere. Meanwhile, if the same term occurs in all documents, the TF-IDF would be 0 [12]. Equation (1) depicts the calculation of TF-IDF, where tfi,j is the number of occurrences of i in j, dfi is the number of documents containing i, and N is the total number of documents.

$$TF - IDF = tf_{i,j}\ x\ log\left(\frac{N}{df_i}\right) \tag{1}$$

5) Sentiment anlaysis model: The resulting TF-IDF values are classified as positive or negative using machine learning models. Naïve Bayes classifier, Support Vector Machine (SVM), and Logistic Regression is used. Naïve Bayes is a probabilistic machine learning model that fits tasks in relation to classification. Equation (2) shows the Bayes theorem where A is hypothesis and B is evidence. Using the theorem, the probability of A happening can be determined given that B has occurred. The assumption of the theorem is that the features are independent such that the presence of one predictor does not affect the other; hence, the term naïve.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2}$$

SVM, on the other hand, aims to find a hyperplane that has the maximum margin in an n-dimensional space in order to distinctly classify the data points. Logistic regression produces a logistic curve which is created using natural logarithm of the 'odds' of target variable. Equation (3) depicts the logistic regression formula, where P is the probability of a 1, e is the base of natural logarithm, and a and b are the parameters of the model [15].

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}} \tag{3}$$

For training, data from 1st of January, 2013 until 31st of December, 2016 is split into training and testing. Stratified K-Folds cross-validator is used in which the number of folds or splitting iterations are specified to be 5. For validation, the news data from 1st of January, 2017 until 31st of December, 2017 is used, which is used as input for stock price prediction in the latter part.

B. Stock price analysis

1) Historical stock data gathering: As for the stocks, the stock market data is gathered using Yahoo! Finance python library, namely yfinance. The companies of which stock market data is gathered are as follows: Microsoft Corporation (MSFT), Facebook, Inc. (FB), Amazon.com, Inc. (AMZN), Alphabet Inc. (GOOG), and Apple Inc. (AAPL). The opening price, adjusted closing price, highest price, and lowest price of each day is obtained starting 1st of January, 2013 until 31st of December, 2017 [11].

2) Historical stock data preprocessing: Based on the obtained stock price features, the percent change between highest and lowest stock price for each day is calculated. Also, the input parameters, which are the opening, highest, and lowest stock prices, are scaled to standardize the dataset.

3) Prediction model: In order to predict the stocks, K-Nearest Neighbors (KNN) is used. The algorithm works by calculating the distance between test data and each of the rows in training data using a distance formula, typically the Euclidean distance as shown in (4). Based from the sorted array of training data, the top K rows are chosen and a class is assigned to the test point based on the common class of the rows.

$$d(p,q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{4}$$

V. Results and Discussion

The section describes the results from analysis of news sentiments and prediction of stock market behavior. Experimentation is divided into two phases, the first one is classification of news articles as either positive or negative, and the second phase is predicting the rise and fall of stocks based on news sentiments and historical stock price. The following subsections depict the results of experimentation in each phase.

A. Result of news sentiment analysis

The metrics to be used for evaluation of results is Area Under the Receiver Operating Characteristic Curve (AUC), Accuracy, and F1-score. The result of news sentiment analysis is depicted in Table I. SVM, Naïve Bayes, and Logistic Regression algorithms are used to analyze sentiments of input news articles. As depicted, the results from Naïve Bayes classifier has the lowest AUC, accuracy, and F-1 score for all POS combinations, in comparison to SVM and logistic regression. Meanwhile, logistic regression performed the best among the three algorithms by having an average AUC of 83.87%, average accuracy of 76.28%, and average F-1 score of 75.03%. If compared to Khedr's study, which obtained accuracies ranging from 72.73% to 86.21% using naïve bayes algorithm, the accuracy of the proposed model is lower yet the AUC score is still high enough (up to 84.38%) which is an indication that the quality of the model's predictions are good. Moreover, for the proposed model, logistic regression proved to be better than naïve bayes algorithm.

As for POS combinations, Noun+Adjective POS combination has the lowest performance across all algorithms, while the performance of Noun+Verb and Noun+Verb+Adjective is close enough as can be seen in AUC, accuracy, and F-1 metrics. In logistic regression, which

TABLE I. RESULTS OF SENTIMENT ANALYSIS MODEL

		AUC	Accuracy	F-1
SVM	Noun+Verb	84.23	68.51	73.86
	Noun+Adjective	82.73	67.63	73.35
	Noun+Verb+Adjective	84.17	69	74.25
Naïve Bayes	Noun+Verb	79.14	74	71.57
	Noun+Adjective	77.94	73.44	71.28
	Noun+Verb+Adjective	79.37	74.38	72.32
Logistic Regression	Noun+Verb	84.38	76.98	75.85
	Noun+Adjective	82.92	75.48	74.15
	Noun+Verb+Adjective	84.3	76.37	75.09

TABLE II. R^2 SCORE OF STOCK TREND PREDICTION

		Intraday	3 days	5 days	7 days
Microsoft	Noun+Verb	98.70	96.80	94.72	94.07
	Noun+Adjective	98.76	95.33	94.55	93.32
	Noun+Verb+ Adjective	98.18	95.06	95.07	91.35
Google	Noun+Verb	98.87	93.91	94.85	93.86
	Noun+Adjective	98.81	93.78	95.08	92.5
	Noun+Verb+ Adjective	98.02	96.46	95.45	91.73
Apple	Noun+Verb	98.93	96.85	95.62	93.89
	Noun+Adjective	98.41	96.7	94.21	92.61
	Noun+Verb+ Adjective	98.67	96.49	95.9	94.6
Amazon	Noun+Verb	99.07	97.15	95.85	95.68
	Noun+Adjective	98.89	98.17	97.21	96.55
	Noun+Verb+ Adjective	98.97	97.78	97.06	97.4
Facebook	Noun+Verb	99.06	97.85	97.28	96.62
	Noun+Adjective	99.14	97.72	96.88	95.74
	Noun+Verb+ Adjective	97.56	98.19	96.72	96.03

TABLE III. MAE REGRESSION LOSS SCORE OF STOCK TREND PREDICTION

		Intraday	3 days	5 days	7 days
Microsoft	Noun+Verb	0.75	1.18	1.62	1.71
	Noun+Adjective	0.69	1.43	1.64	1.89
	Noun+Verb+ Adjective	0.8	1.4	1.61	1.75
Google	Noun+Verb	8.71	20.20	21.19	23.68
	Noun+Adjective	7.84	19.74	21.21	25.21
	Noun+Verb+ Adjective	8.78	17.14	19.09	24.63
Apple	Noun+Verb	1.51	3.06	3.29	4.12
	Noun+Adjective	1.80	3.06	3.94	4.34
	Noun+Verb+ Adjective	1.77	3.04	3.49	4.03
Amazon	Noun+Verb	9.92	18.92	23.96	25
	Noun+Adjective	10.83	17.07	20.51	26.71
	Noun+Verb+ Adjective	10.56	19.75	22.19	22.22
Facebook	Noun+Verb	2.11	3.43	3.74	4.45
	Noun+Adjective	1.93	3.28	4.22	4.74
	Noun+Verb+ Adjective	2.17	3.33	4.20	4.11

performed the best among the three algorithms, the POS combination of Noun+Verb having an AUC of 84.38% outperformed the other POS combinations. This could be due to the greater number of nouns and verbs in new articles through which higher information content is deciphered in the parsed text that helps in giving more context for sentiment analysis. If looking into the findings with a different perspective, the higher AUC for Noun+Verb may imply that nouns and verbs have greater impact on human sentiments than adjectives. Meanwhile, the AUC of Noun+Verb+Adjective is 84.3%, which is not far from the AUC of Noun+Verb (84.38%).

B. Result of future stock trend prediction

The stock prediction model uses KNN algorithm for regression analysis and R^2 and Mean Absolute Error (MAE) regression loss metrics for evaluation. The sentiments obtained from sentiment analysis model, specifically from logistic regression which obtained the highest AUC, are used as input into stock trend prediction model, in addition to historical stock price parameters. In each of the different POS combinations being evaluated, the only variable input factor is the sentiment obtained from the POS combination of sentiment analysis model. The other inputs, which are historical stock prices, remain constant. Also, different time lags have been used. Intraday, 3 days, 5 days, and 7 days are used for the experiments.

The results obtained from future stock trend prediction are shown in Tables II and III for R^2 score and mean absolute error regression loss, respectively. The comparison of R^2 scores among different time spans for each of the companies can be seen in Fig. 3. As depicted, the intraday adjusted closing price is being predicted with the highest R^2 score and lowest MAE while the 7-day time-lag had the lowest R^2 score and highest MAE for all stock indices among the different time spans being evaluated. Facebook bagged the highest intraday R^2 score of 99.14% while Google had the lowest 7-day time-lag of 91.14%. As for MAE, the lowest one is achieved by Microsoft having 1.1% while the highest can be seen for Amazon with 32.39%. Higher MAE can be attributed to wider and more skewed target variable or the adjusted closing prices of stock indices.

When looking into comparison of results between sentiments obtained from different POS combinations, as shown in Fig. 4, it can be seen that the results vary. For intraday stock predictions, the results from different sentiments are close enough for all stock indices. An interesting observation deciphered is that as the time lag increases, the variation amongst R^2 scores of different POS combinations within each time span also increase. For instance, in the case of Microsoft, the intraday R^2 scores of all POS combinations range from 98.18% to 98.76% in comparison to the 7-day time-lag R^2 scores which range from 91.35% to 94.07%.

In order to have an in-depth understanding of the importance of predictors or the degree at which each of the input features affects the output, Recursive Feature Elimination (RFE) is used. The feature selection approach works by iteratively removing input features and building the model on the features that are remaining. It used the accuracy metric to recognize which combination of attributes contribute the most in predicting the desired value. Fig. 5 illustrates the ranking of input features, with 1 being the most important and 4 being the least important in predicting the target attribute. As can be seen, the highest stock price input, which is at rank 1, has the highest weight or contribution in the output, followed by the opening price, lowest price, percent change between the highest and lowest price, and finally the sentiment inputs.

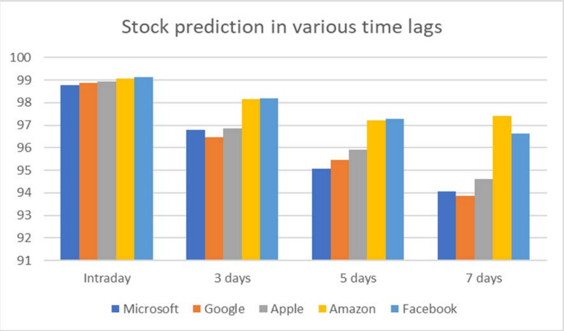

Fig. 3. Stock prediction in various time lags.

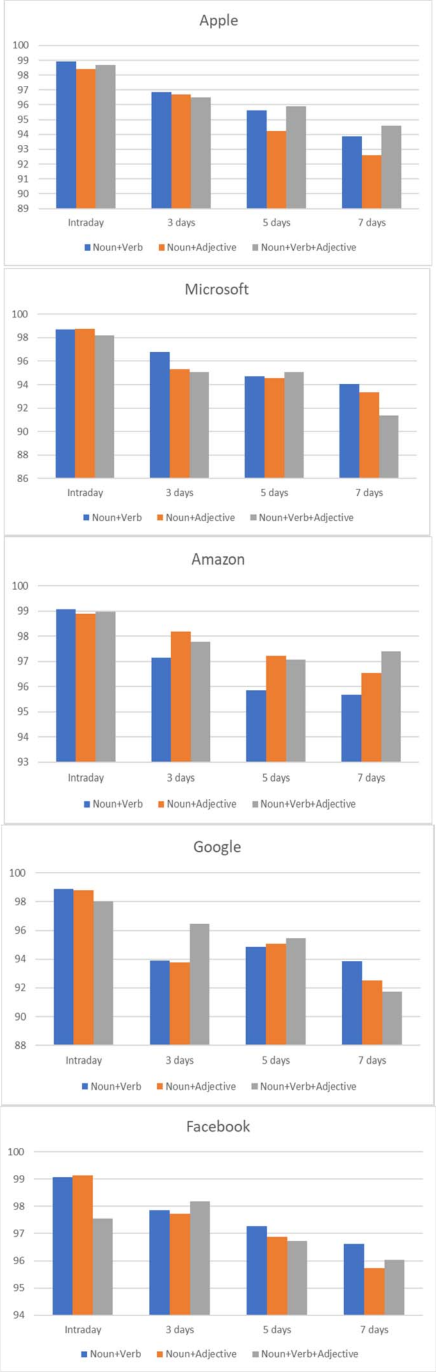

Fig. 4. Stock prediction model results for each of the POS combinations in various time lags.

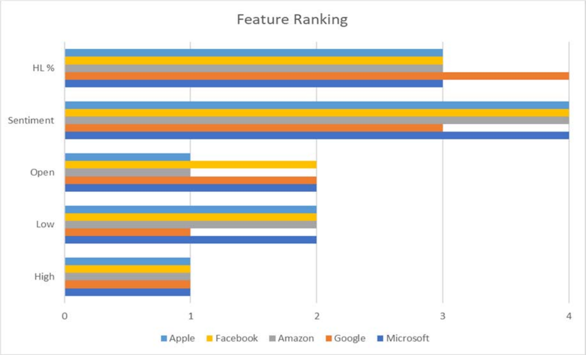

Fig. 5. Input feature ranking based on RFE for each stock index.

VI. Conclusion and Future Work

The proposed model looked into analyzing sentiments of news, and how the sentiments about a certain company can affect its stock prices, along with the company's historical stock data. The proposed model is divided into two phases, the first one involving text preprocessing and using Naïve Bayes, logistic regression, and SVM classifiers for sentiment analysis, and the second one using resulting sentiments together with historical stock opening, closing, high, and low prices for predicting future stock trends using KNN.

Our proposed sentiment analysis model reached AUC of 84.38%. If compared to Khedr's study, which obtained accuracies ranging from 72.73% to 86.21% using naïve bayes algorithm, the accuracy of the proposed model is lower yet the AUC score is still high enough (up to 84.38%) which is an indication that the quality of the model's predictions is good. Moreover, for the proposed model, logistic regression proved to be better than naïve bayes algorithm. Also, it could be observed that the text inputs used in the model has greatly been reduced since instead of using the whole text, only specific parts of speech were used, such as nouns, adjectives, and verbs. This might have reduced processing time and memory consumption to a great extent, which could be something the future studies may look into.

As for POS combinations, Noun+Adjective has the lowest performance across all algorithms, while the performance of Noun+Verb and Noun+Verb+Adjective is close enough as can be seen in AUC, accuracy, and F-1 metrics. This could be due to the greater number of nouns and verbs in new articles through which higher information content is deciphered in the parsed text that helps in giving more context for sentiment analysis. This may also possibly imply that nouns and verbs have greater impact on human sentiments than adjectives.

Based on stock prediction results, the intraday adjusted closing price is being predicted with the highest R^2 score and lowest MAE while the 7-day time-lag had the lowest R^2 score and highest MAE for all stock indices. Some stock indices had higher MAE which can be attributed to wider and more skewed adjusted closing prices. An interesting observation deciphered is that as the time lag increases, the variation amongst R^2 scores of different POS combinations within each time span also increase. The RFE results show that highest intraday stock price input parameter contributes the most while the sentiments contribute the least to the adjusted closing price output.

Based on the overall result of the stock prediction model, we can see that the findings are in line with the current studies illustrating the correlation between news sentiments and stock trends. Though based on RFE, historical stock prices have a greater contribution to the adjusted closing price than sentiments, but still there is an established correlation between stock prices and sentiments as depicted in the results. It is difficult to compare the performance of the model with studies like that of Khedr's, Ansari's, and Jaber's since these are using discretized stock price inputs and accuracy metric for assessment [14, 21, 22]; meanwhile, the proposed study is using continuous stock price inputs and has used R^2 score for evaluation. The model has achieved R^2 score of up 99.14% which surpasses the R^2 score obtained by Nguyen *et al.* of up to 98% [16].

The future studies can further improve the accuracy of the model, specifically the sentiment analysis part, by using a better vector representation other than TF-IDF. Moreover, other parts of speech can also be explored, such as adverbs, conjunctions, determiners, and the like. Aside from POS combinations, future research may also explore various permutations. This may help in determining how sentence construction can affect the derived sentiments from text.

REFERENCES

[1] W. P. Risk, G. S. Kino, and H. J. Shaw, “Fiber-optic frequency shifter using a surface acoustic wave incident at an oblique angle,” Opt. Lett., vol. 11, no. 2, pp. 115–117, Feb. 1986.

[2] Z. Hu, W. Lu, J. Bian, X. Liu, and T. Liu, “Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction,” ACM, CA, USA, pp. 261–269, 2018.

[3] X. Zhang, S. Qu, J. Huang, B. Fang, and P. Yu, “Stock market prediction via multi-source multiple instance learning,” IEEE. IL, USA, pp. 50720 – 50728, 2018.

[4] A. Atkins, M. Niranjan, and E. Gerding, “Financial news predicts stock market volatility better than close price,” ScienceDirect: The Journal of Finance and Data Science, vol. 4, no. 2, pp. 120–137, 2018.

[5] Y. Peng and H. Jiang, “Leverage financial news to predict stock price movements using word embeddings and deep neural networks,” Association for Computational Linguistics, pp. 374-379, 2016.

[6] I. Medovikov, “When does the stock market listen to eco-nomic news? New evidence from copulas and news wires,” ScienceDirect: Journal of Banking & Finance, vol. 65, pp. 27-40, 2015.

[7] V. Pagolu, K. Challa, G. Panda, and B. Majhi, “Sentiment analysis of twitter data for predicting stock market movements,” International conference on Signal Processing, Communication, Power and Embedded System (SCOPES), pp. 1345-1350, 2016.

[8] M. Siikanen, K. Baltakys, J. Kanniainen, R. Vatrapu, R. Mukkamala, and A. Hussain, “Facebook drives behavior of passive households in stock markets,” ScienceDirect, vol. 27, pp. 208-213, 2018.

[9] A. Shapiro, M. Sudhof, and D. Wilson, “Measuring news sentiment. federal reserve bank of san francisco working paper 2017-01”, 2018. DOI: https://doi.org/10.24148/wp2017-01

[10] J. Kalyani, H. N. Bharathi, and R. Jyothi, “Stock trend prediction using news sentiment analysis,” International Journal of Computer Science & Information Technology (IJCSIT), vol. 8, no. 3, pp. 67-76, 2016.

[11] S. Feuerriegel, A. Ratku, and D. Neumann, “Analysis of how underlying topics in financial news affect stock prices using latent dirichlet allocation,” 2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 1072-1081, 2016. DOI: 10.1109/HICSS.2016.137

[12] A. Khedr, S. Salama, and N. Yaseen, “Predicting stock market behavior using data mining technique and news sentiment analysis,” International Journal of Intelligent Systems and Applications, vol. 9 (7), pp. 22-30, 2017.

[13] B. Meyer, M. Bikdash, and X. Dai, “Fine-grained financial news sentiment analysis,” pp. 1-8, SoutheastCon 2017.

[14] A. Garcia-Medina, L. Junior, E. Banuelos, and A. Martinez-Arguello, “Correlations and flow of information between the new york times and stock markets,” Elsevier, vol. 502, pp. 403-415, 2017.

[15] A. Lavey and L. Konish, "The five biggest tech companies now make up 17.5% of the S&P 500 — here's how to protect yourself," CNBC, https://www.cnbc.com/2020/01/28/sp-500-dominated-by-apple-microsoft-alphabet-amazon-facebook.html. 2020.

[16] H. Nguyen, A. Rahimyar, and X. Wang, “Stock forecasting using m-band wavelet-based SVR and RNN-LSTMs models,” 2019 2nd International Conference on Information Systems and Computer Aided Education (ICISCAE) (IEEE), pp. 234-240, 2019.