The University of Texas at Austin

Department of Computer Science

Sentient Stocks: Discovering Market Mysteries Through Sectorized Sentiment Analysis and Predictive Modeling

Diego Renzo Sariol

Case Studies in Machine Learning

December 2023

Table of Contents

Abstract

The influence of the media on the stock market comes as no surprise and has been made evidently more aware these last few years given the traumatic global events that have occurred. The last five years of both natural and man-made disasters have had detrimental impacts to the United State's economy, most recently being the Covid-19 influenza pandemic that began in 2019. This research paper sets forth to delve deeper into understanding the intricate relationship between global headlines and corresponding price fluctuations for several stocks across diverse sectors. Through the use of Natural Language Processing, Sentiment Analysis, Logistic Regression, and Time Series Forecasting, the value of the media on different markets will be better understood for the last five years starting in May of 2018. This analysis will be aided through the creation of several different types of Machine Learning models used for NLP label forecasting. The predictive power of the media will be uncovered across the following markets: Technology, Transport, Medicine, and Finance.

1) Introduction

The advancement of technology has grown exponentially and thus the facilitation of communication has become easier and more accessible as a result. The power of information can have either a positive or negative effect and this can directly be observed as a result of media publications. The last five years of global events have had great consequences on the United States and its economy. The Covid-19 pandemic resulted in commercial businesses shutting down and enforced lockdowns which both had severe impacts to the economy and this is only one major global event that had a major influence on the market and the economy (Wen et al., 2021). This paper delves into the intricate relationship between the media's coverage of events and overall stock market trends.

Through the use of advanced Natural Language Processing techniques and machine learning methods, the subtle embedded relationships between media headlines and different stock markets will be revealed. The growth of Natural Language Processing has dynamically changed how large volumes of textual data can be restructured, pre-processed, and normalized into a set format (Souma et al., 2019). This will allow for a sentiment analysis to be performed on each day's top twenty-five new's headlines by two pre-trained sentiment analysis tools. A standard social media model and a variation of the famous language model "BERT", ironically named "finBERT" for it is a sentiment analysis model trained on financial news. Extracting the most valuable features from the text will allow for a better understanding of the most frequently occurring pieces of information and the overall compound sentiment score for each headline (Dong et al., 2020).

Four different sectors in the United States' stock market will be further analyzed alongside Reddit's *r/WorldNews* subreddit's last five years of new's headline posts. This includes the following stocks: *Microsoft*, *Dow Jones Transportation Average*, *Johnson & Johnson*, and lastly *VISA*. A total of three models will be constructed, trained, and evaluated for each of the four stock tickers. This will allow for a sectorized stock sentiment analysis on the media's top twenty-five headlines for each day. The predictive power of the media and its influence on the different sectors of the stock market will be revealed. The models created in these experiments and the insights derived from the following analysis will not be used for market trading, nor are they recommended. The value of the research and results of the sentiment analysis, logistic regression, and time series forecasting will allow for a foundation to be built for understanding and predicting market behavior and the concurrent media narratives (Pagolu et al., 2016).

2) Research Background

The study of Natural Language Processing in stock market prediction has been gaining more traction and support as both technologies and insights into deep learning have evolved. Investigating the relationship between the overall sentiment of new's headlines and different sectorized stocks comes as a result of previous research studies utilizing only one stock and eight years of new's headlines (Vicari et al., 2020). The study of capturing and analyzing sentence semantics and sentiment context has developed itself into the established field of "Natural Language Based Financial Forecasting" (NLFF) (Xing et al., 2017). Utilizing the latest available resources is what sets this research study apart from previous attempts which incorporated world new's headlines from the year 2008 all the way through 2016. The most popular type of model used throughout these studies were chosen due to their ability to handle time series forecasting and textual information (Puh et al., 2023).

With the stock market being observably influenced by different forms of media, there exist methods to patternize these trends and sentiment based movements. In the case of time series forecasting, future predictions are based on previously known values. This is known as an autoregressive time series problem which allowed for one study to utilize a statistical analysis test known as the "Autoregressive Integrated Moving Average", ARIMA, in order to attempt to make financial forecasts (Mehta et al., 2021). These studies continue to delve deeper into the world of Natural Language Processing and Sentiment Analysis as text corpuses exponentially increase with the endless input of social media and new's information that are published every hour of each day. Analyzing and predicting stock market price fluctuations based on new's headlines reveals the influence of the global media on the market; however, many such studies have been conducted on other social media networks, such as Twitter (Bollen et al., 2010). This provides an interesting approach to a well known problem as Twitter allows for almost anyone to publish a Tweet whether it be factual or not.

3) Methods

3.1) N-Gram Logistic Regression

Logistic Regression is a supervised machine learning technique used for classification problems. In the case of this research paper, the classification problem is binary as the label for a stock's price will be a "1" if the value stayed the same or increased and "0" if the value decreased. The value of this classification algorithm is that it allows for the creation of a generalized linear baseline model that utilizes previously known information and labels in order to make a classification prediction for the expected value (Pant, 2019). Logistic Regression is known as a sub-type of Linear Regression; however, the difference comes as a result of the *Sigmoid* cost function. This cost function allows for the mapping of probabilities to actual concrete binary values to be used in classification (Coursesteach, 2023). The difference between the baseline model and the intermediate Logistic Regression model is the use of bi-grams. This allows for the model to understand the contextual and sequential relationships found in each of the daily new's headlines. N-Gram models can be expanded to any value of N; however, this will not necessarily increase the accuracy output, but it will increase the computation complexity.

3.3) LSTM Recurrent Neural Network

Long short term model, better known as LSTM, is a type of Recurrent Neural network model that also utilizes supervised learning in order to produce a binary classification label, depending on the cost function as mentioned in the previous section (Saxena, 2023). Unlike the Logistic Regression model, the LSTM is a deep learning neural network that is particularly adept at handling sequential data and understanding both short and long term dependencies. The creation of LSTM was to handle the problems brought forth by the standard feed forward neural network which suffers from the "vanishing gradient problem." This issue comes as a direct result of both the gradient-based learning techniques exhibited by the model and the use of backpropagation, but if the gradient calculation is incredibly small, it essentially "vanishes" (Dolphin, 2020). The predictive capabilities of the LSTM model come as a result of its ability to process entire sequences of text and still be able to retain the temporal relationships between the context and sequences (Chandola et al., 2022).

4) Data

4.1) Retrieval, Creation, and Modification

The collection of the data can be divided into two separate tasks as this original research analysis utilizes the latest *Kaggle* dataset for *Reddit's* top new's headlines and more than one stock sector price sheet. This means that the new's headline data can simply be sourced from the well known dataset database, *Kaggle*, and can be seen below in **Figure 1** (Arora, n.d.). The stock ticker's data must be retrieved through *Yahoo Finance*. This process can be achieved through a Python library coined *"yfinance"* as it can directly access the publicly available *Yahoo* API's and retrieve all listed price information for a particular ticker between two different dates and can be seen below in **Figure 2**. This process can be repeated several times in order to retrieve all of the necessary dataset information to be combined with the previously mentioned textual new's dataset.

In order to create the combined dataset which could then enter the textual pre-processing stage, the stock ticker dataset needs to be modified in order to include an additional column which will denote whether or not that particular stock's adjusted closing price changed from the previous day. This can be done by applying a lambda function to a specific dataset's column in order to produce the new column and values. The two above datasets must then be merged through the use of an inner join on the "Date" values in the retrieved stock dataset. Each row of data that exists in the stock dataset would be merged with the corresponding row in the new's headline dataset. This means that there will not be any rows in the joined dataframe that do not contain a stock price label change. Lastly, the columns containing a new's headline could be textually aggregated and a new column could be created of the top twenty-five headlines which would allow for easier model training and evaluation which can be seen below in **Figure 3**.

Date	Top 1	Top 2	Top 3	…	Top 25
2018-05-01	North Korea to open its sky, South Korean medi...	The Mueller probe ain't ending anytime soon	BRAND NEW: 2018 – Renault Alpine A110 – Start-...	…	For first time health ministry will regulate d...
…	…	…	…	…	…
2023-04-30	Israel will raze Khan al-Ahmar, no timetable f...	China’s Ding Liren becomes world chess champio...	Russian forces suffer radiation sickness after...		UN food body warns Sudan violence could fuel r...

Figure 1: 'WorldNewsData.csv' from *Kaggle*'s "Top 25 World News (2018-2023)" dataset

Date	Open	High	Low	Close	Volume	Adj Close
2018-05-01	93.209999	95.290001	92.790001	95.000000	31408900	89.093185
…	…	…	…	…	…	…
2023-04-28	304.010010	308.929993	303.309998	307.260010	36446700	305.322327

Figure 2: *Yahoo Finance* API retrieval of *MSFT* stock ticker price sheet in USD

Date	Label	Combined Text
2018-05-01	0	North Korea to open its sky, South Korean medi… The Mueller probe ain't ending anytime soon…BRAND NEW: 2018 – Renault Alpine A110 – Start-...For first time health ministry will regulate d...
…	…	…
2023-04-28	1	Israel will raze Khan al-Ahmar, no timetable f… China’s Ding Liren becomes

		world chess champio… Russian forces suffer radiation sickness after… UN food body warns Sudan violence could fuel r...

Figure 3: Resulting *Pandas* dataframe for combined 'WorldNewsData.csv' headlines and *MSFT* ADJ Close price change label

4.2) Pre-processing

The next step in the Natural Language Processing pipeline is to pre-process the unstructured textual data. The value and importance of pre-processing vast amounts of text means the difference of effectiveness in the subsequent sentiment analysis and models to be created (Zhang et al., 2022). The transformation process involves first tokenizing all of the text and then cleaning the resulting token arrays by essentially normalizing all pieces of text on a standardized basis. This means that irrelevant characters, emojis, unnecessary stop-words, and punctuation are removed. The remaining text can be further pre-processed through the use of stemming and or lemmatizing which means reducing words to their base forms. This entire process is extremely valuable as not only does it normalize all of the text to a common structured lowercase format, it also eliminates noise and unfamiliar data (Harshith, 2019). The overall importance is to limit the skew on the sentiment, classification, and time series forecasting analysis. The entire process can be seen below in the table in **Figure 4**.

Original Text	Aides to Donald Trump, the US president, hired an Israeli private intelligence agency to orchestrate a “dirty ops” campaign against key individuals from the Obama administration who helped negotiate the Iran nuclear deal, the Observer can reveal.
Tokenized	['aides', 'to', 'donald', 'trump', ',', 'the', 'us', 'president', ',', 'hired', 'an', 'israeli', 'private', 'intelligence', 'agency', 'to', 'orchestrate', 'a', '“', 'dirty', 'ops', '”', 'campaign', 'against', 'key', 'individuals', 'from', 'the', 'obama', 'administration', 'who', 'helped', 'negotiate', 'the', 'iran', 'nuclear', 'deal', ',', 'the', 'observer', 'can', 'reveal', '.']
Lowercase and Punctuation Removed	['aides', 'to', 'donald', 'trump', 'the', 'us', 'president', 'hired', 'an', 'israeli', 'private',

	'intelligence', 'agency', 'to', 'orchestrate', 'a', 'dirty', 'ops', 'campaign', 'against', 'key', 'individuals', 'from', 'the', 'obama', 'administration', 'who', 'helped', 'negotiate', 'the', 'iran', 'nuclear', 'deal', 'the', 'observer', 'can', 'reveal']
Stopwords Removed	['aides', 'donald', 'trump', 'us', 'president', 'hired', 'israeli', 'private', 'intelligence', 'agency', 'orchestrate', 'dirty', 'ops', 'campaign', 'key', 'individuals', 'obama', 'administration', 'helped', 'negotiate', 'iran', 'nuclear', 'deal', 'observer', 'reveal']
Lemmatized	['aide', 'donald', 'trump', 'u', 'president', 'hired', 'israeli', 'private', 'intelligence', 'agency', 'orchestrate', 'dirty', 'ops', 'campaign', 'key', 'individual', 'obama', 'administration', 'helped', 'negotiate', 'iran', 'nuclear', 'deal', 'observer', 'reveal']

Figure 4: Table displaying the pre-processing pipeline for an example new's headline input

4.3) Feature Extraction

In order to begin the analysis of the text, it is extremely beneficial to construct a visual analysis through the use of textual feature extraction. This allows for the interoperability of the results to be better understood. The types of features that will be extracted for the data are relevant values such as overall combined text length and most frequently occurring words. The importance in extracting these two features can aid the analysis in the underlying structure and overall repeating patterns in the data (Usmani et al., 2021). Similar to the pre-processing stage, feature extraction can help to remove unwanted noise and filter out (un)common words. This allows for the key themes and context to be derived from the overall new's headline corpus of text. In the WordMap seen below in **Figure 5**, the most frequently occurring words are displayed. The color and orientation of the words have no particular meaning; however, the font-size itself indicates the words that appear the most in the text corpus. Similarly as seen in the bar chart below in **Figure 6**, words with the highest occurring frequency are directly related to recent traumatic global events that have taken place in the last five years (Eskandar, 2023).

Figure 5: Python package, *WordCloud,* and its corresponding produced WordMap for the dataset

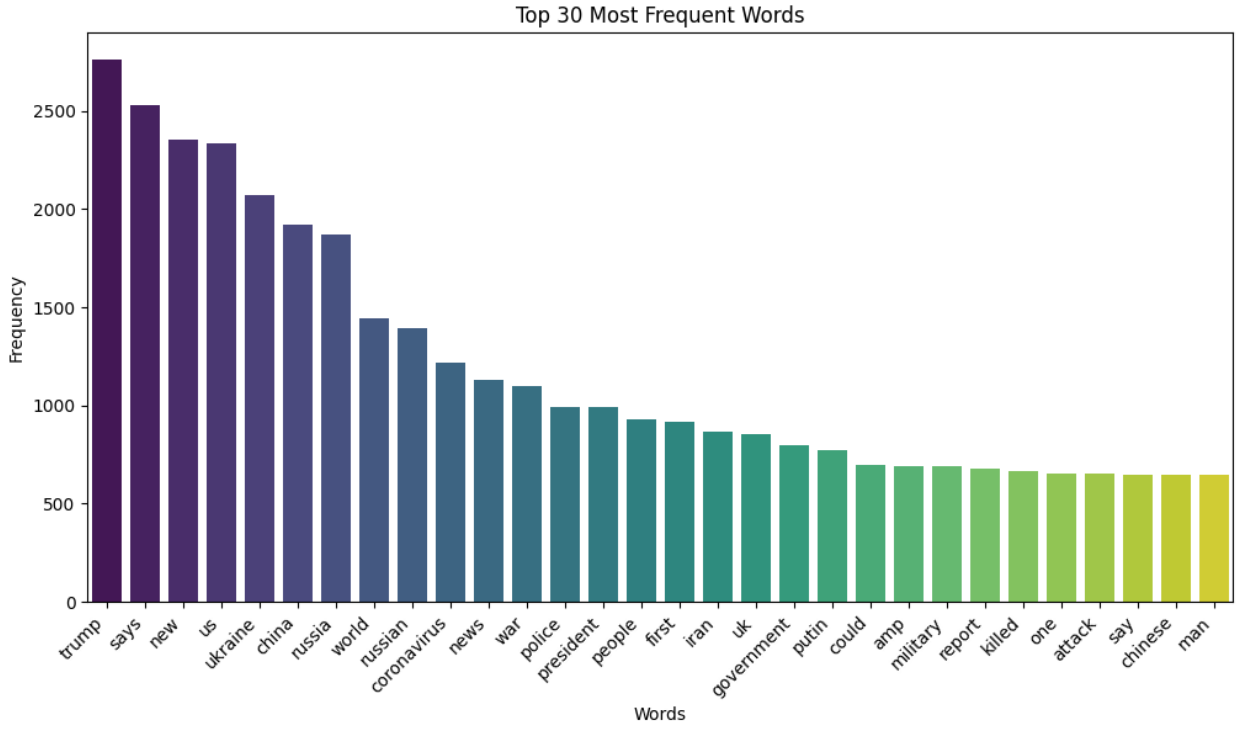

Figure 6: Top 30 most frequently occurring words from the combined new's headline text corpus

4.4) Sentiment Analysis ("VADER", "finBERT")

Now that the text has been properly pre-processed and the unwanted noise and outliers have been filtered out, it is now properly structured as input for pre-built sentiment analysis tools. The two models that will be used are "VADER" from Python's *Natural Language Toolkit* and a modified version of the well known language model "BERT", coined "finBERT" as it was trained on financial texts (Huang et al., 2020). The value of performing sentiment analysis with these two very different models will reveal how the difference in training a model on social media text versus financial pieces of text, can impact the calculated sentiment scores for each new's headline (Ashtiani et al., 2023). "VADER" is more suited for social media text and information as it was trained particularly on the sentiments expressed in social media and through emoticons. This is because it is a "Valence Aware Dictionary and sEntiment Reasoner " meaning that unlike "finBERT", it is better attuned toward social sentiment rather than financial sentiment (Hutto et al., 2015). In the following two charts below in **Figure 7** and **Figure 8**, the distribution of sentiment scores from the two analysis tools can be seen. These scores are a compounded average score of the sentiment score for each of the twenty five top headlines for each date listed in the dataset. The results reveal that "VADER" returns a mean sentiment score with what appears to be a binomial distribution centered around -0.1 whilst "finBERT" produces a binomial distribution centered around -0.75. This very clear distinction comes as a result of the model's pre-training history (Genc, 2020).

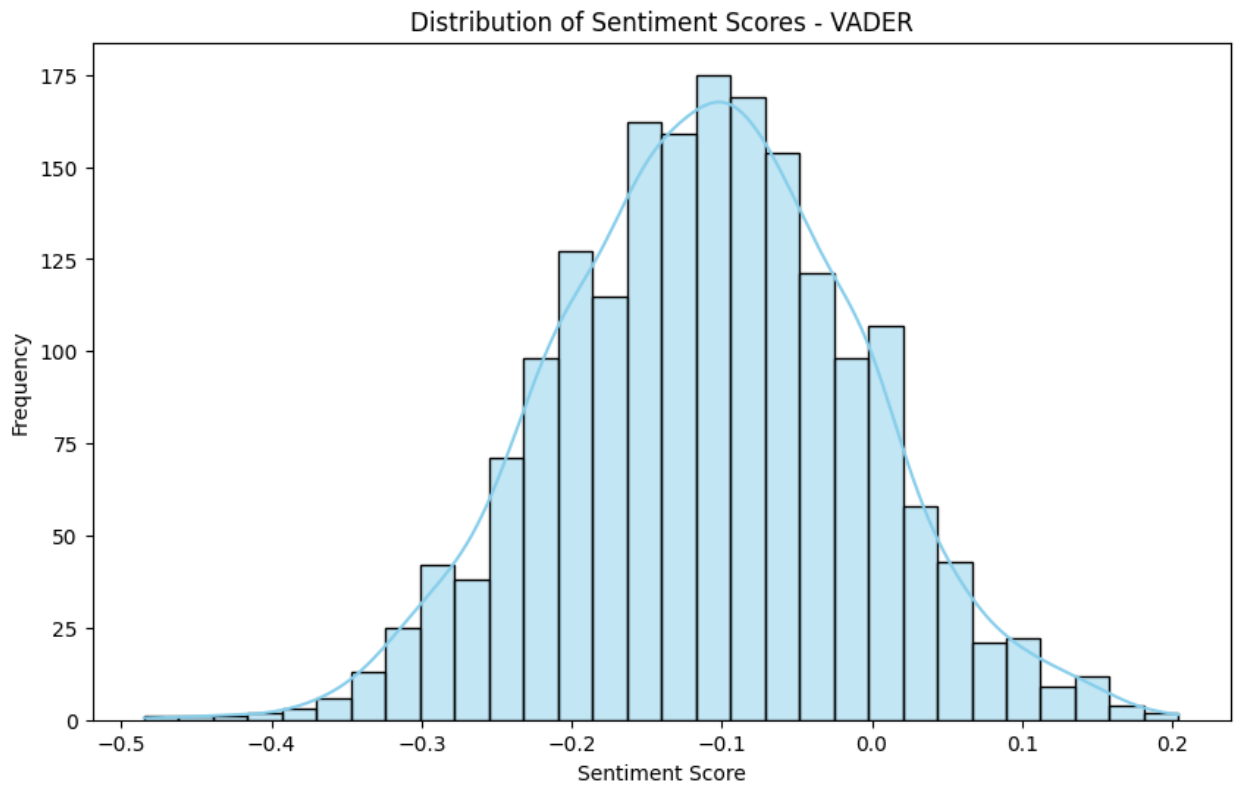

Figure 7: Distribution of "VADER" Lexicon generated sentiment scores

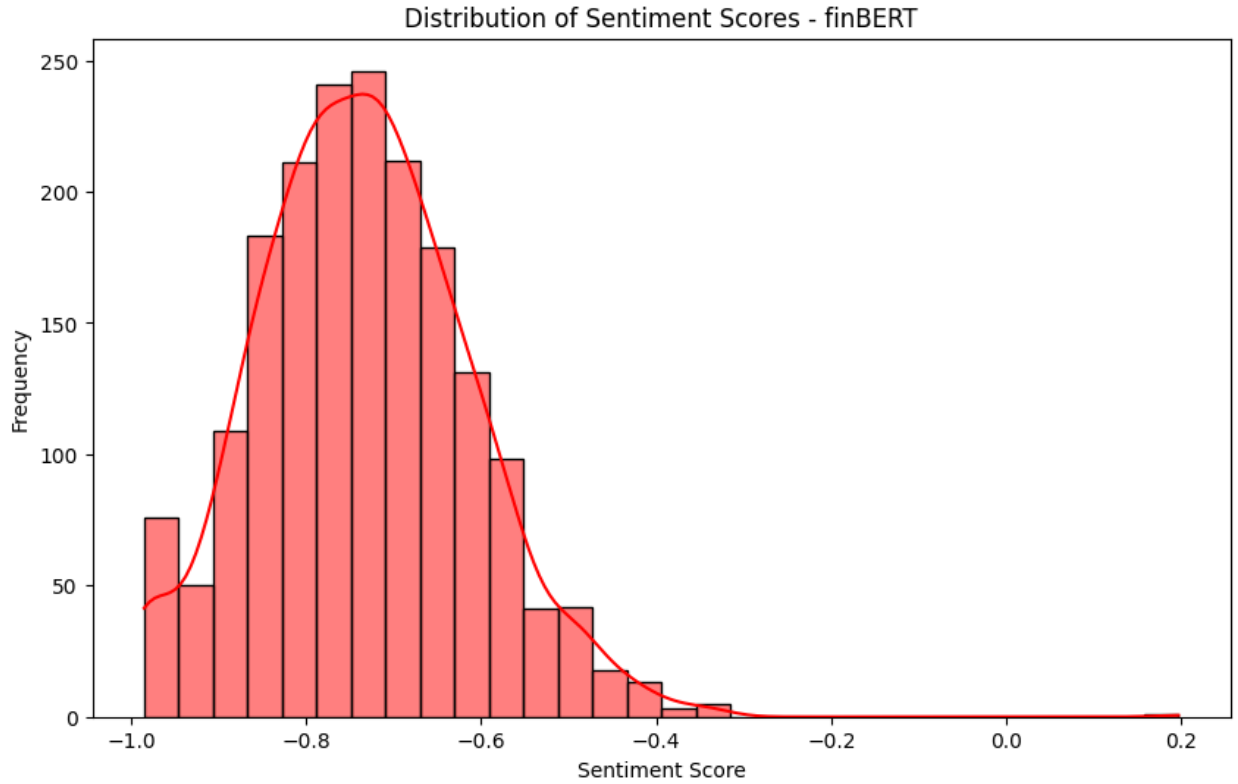

Figure 8: Distribution of financial literature pre-trained "finBERT" generated sentiment scores

5) Results

Logistic Regression Classification Report			
	F1-Score (0)	**F1-Score (1)**	**Accuracy**
MSFT	0.46	0.42	0.44
^DJT	0.49	0.46	0.48
JNJ	0.49	0.54	0.52
VISA	0.44	0.52	0.48

Figure 9: Classification report for Logistic Regression and price prediction

2-Gram Logistic Regression Classification Report			
	F1-Score (0)	**F1-Score (1)**	**Accuracy**
MSFT	0.31	0.54	0.45
^DJT	0.51	0.47	0.49
JNJ	0.42	0.56	0.50
VISA	0.17	0.58	0.45

Figure 10: Classification report for 2-Gram Logistic Regression and price prediction

LSTM RNN Classification Report			
	F1-Score (0)	**F1-Score (1)**	**Accuracy**
MSFT	0.50	0.57	0.54
^DJT	0.50	0.52	0.51
JNJ	0.45	0.58	0.52
VISA	0.50	0.48	0.49

Figure 11: Classification report for LSTM recurrent neural network and price prediction

Figure 9 reveals the Logistic Regression classification report for all four stock tickers, their corresponding "F1-Scores" and "Accuracy" scores. In a binary classification problem, it would appear that the accuracy metric would be the most vital piece of information; however, the calculation of the "F1-Scores" for each class label, reveals how accurate a model's predictive capabilities are for each distinct class (Kundu, 2021). This score calculation can mathematically defined as the following set of equations seen below in **Figure 12**:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}$$

Figure 12: Mathematical equations for calculating Precision, Recall, and F1-Score

The "F1-Score" equation can be broken down into two smaller calculations which reveal the "Precision" and "Recall" statistics. The calculation of "Precision" can be seen as the total number of *True Positive* examples divided by the sum of *True Positive* and *False Positive* examples. The calculation of the "Recall" statistic is very similar except the total *True Positive* examples are divided by the combined sum of *True Positive* and *False Negative* examples (Wood, n.d.). The mathematical importance of these three statistical calculations reveal the model's ability to precisely predict and its overall sensitivity when making a classification.

The overall accuracy classification report findings can be seen below in **Figure 13** and as one would assume, the Recurrent Neural Network model overall outperforms both 1-Gram Logistic Regression and 2-Gram Logistic Regression for all four stock sectors tested. The LSTM neural network model achieves higher accuracy due to several factors which include being able to handle sequential data, the ability to store previously known information, and its adaptability to both short and long term memory dependencies (Brownlee, 2021). Although the Logistic Regression models make for a simplistic baseline model, their ability to adapt to new patterns seen through the text, is non-existent. Predicting the binary price movement of a stock based on new's headlines is a dynamically changing environment which is why the LSTM and its unique ability to handle supervised sequential learning, surpasses that of the Logistic Regression models (Dobilas, 2022).

Accuracy Classification Report				
	Logistic Regression	**2-Gram Logistic Regression**	**LSTM RNN**	**Average**
MSFT	0.44	0.45	0.54	0.48
^DJT	0.48	0.49	0.51	0.49
JNJ	0.52	0.50	0.52	0.51
VISA	0.47	0.45	0.49	0.47

Figure 13: Classification report comparing all three models and their price prediction accuracies

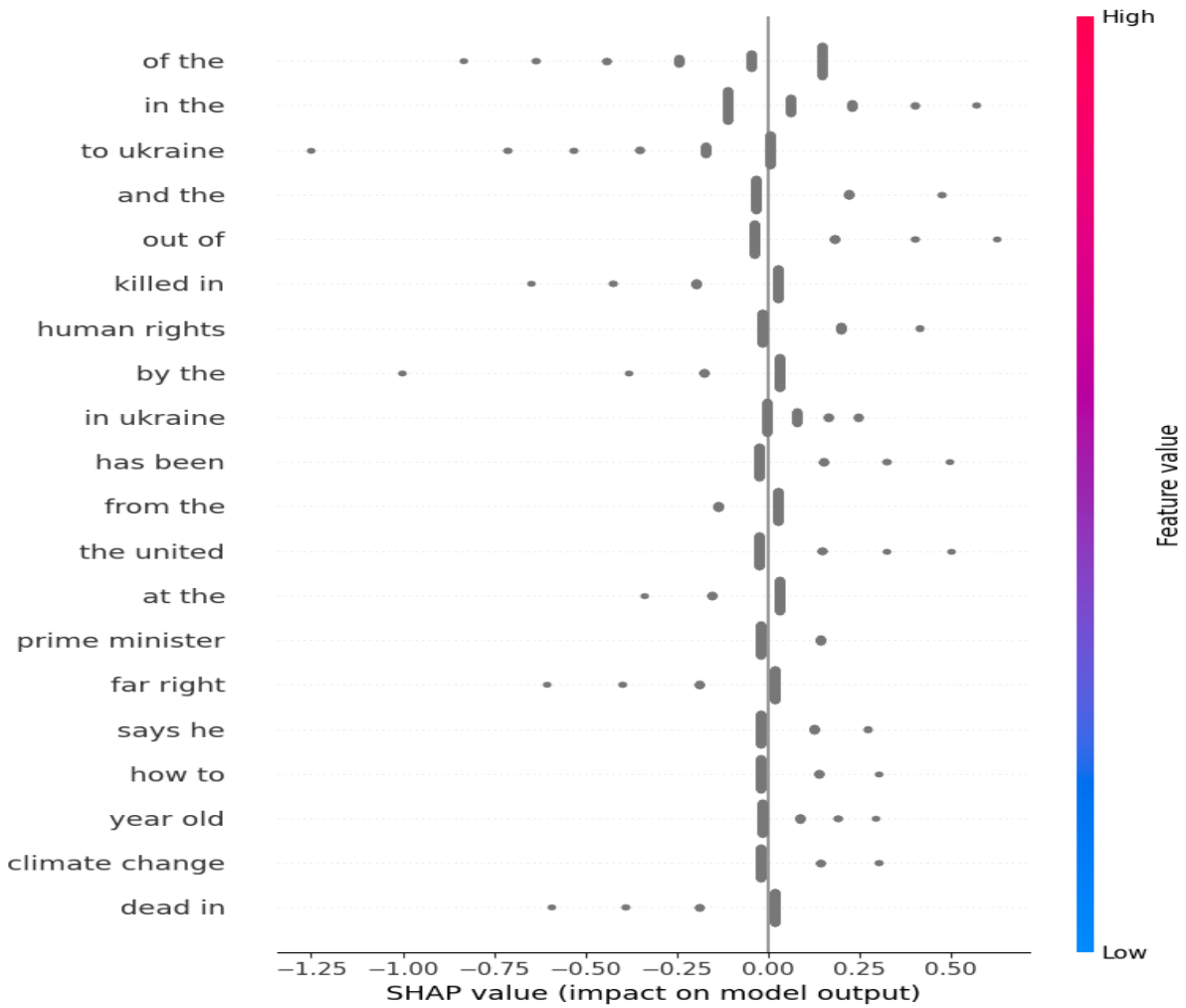

Figure 14: SHapley Additive exPlanations for 2-Gram Logistic Regression

In the above graphic displayed in **Figure 14**, the SHapley Additive exPlanations, "SHAP", are used to reveal the extracted bi-grams from the text corpus and their overall impact on the model output, whether it is positive or negative (Cohen, 2021). On the left hand side, a total of twenty bi-grams are listed and their overall impacts on model output. On the X-axis, the SHAP value is displayed along a number line which contains values ranging from -1.25 on the left to 0.50 on the right. A value being further on the negative side implies that it has a stronger negative effect on the prediction while a value being further on the positive side implies the opposite. Looking on the right hand side, along the Y-axis, a gradient describing the frequency occurrence of each feature helps underline the importance of each bi-gram (Molnar, 2023). For example, the bi-gram "to ukraine" has a very negative SHAP value and is placed as the third most valuable feature. This means that this piece of text that is composed of simply two words, has an extremely negative sentiment and pushes the prediction of the model in the negative direction.

6) Discussion

In conclusion, this study re-examines and provides a more in depth analysis into the intricate relationship between the media and the stock market. When comparing the three different types of models to each other, it becomes increasingly important to note that the accuracy differences most likely appear as a result of correlation between the concurrent media headlines and the specific stock sector which further emphasizes the importance of sectorized sentiment stock analysis. The baseline Logistic Regression model presented moderate predictive power with *MSFT* and *VISA* showing the lowest accuracies which suggests that this simplistic model struggled to capture the complexities of the text and the relationship it shared with a specific stock market. Upgrading the baseline model to incorporate bi-gram textual relationships, resulted in slightly improved results over the basic Logistic Regression, for all four stock tickers. The use of bi-grams emphasized the importance of both context and sequential relationships found in the new's headlines (Jones, 2018).

The most effective model for all four stock tickers was the LSTM Recurrent Neural Network, most particularly for *MSFT* and *JNJ*. These results demonstrate the true power and capabilities of a Long Short Term Memory model to capture the sequential relationships in the text and temporal dependencies (Shah et al., 2022). The combined results reveal the impact of the news and how it is highly stock specific, and in the context of this experiment, it most influenced the medical market. The overall findings through this research highlights the importance of context, sequence, and sentiment in textual analysis. This becomes more evident when comparing the baseline Logistic Regression model to the 2-Gram Logistic Regression model where the only existing difference is the use of bi-grams that introduce sequential context. In the advanced and ever evolving world of Natural Language Processing, this study merely is the "tip of the iceberg", as it leaves room for more possible model improvements and new types of machine learning techniques to be applied and evaluated.

References

Arora, S. (n.d.). *Top 25 World News (2018-2023)*. Kaggle. https://www.kaggle.com/datasets/suruchiarora/top-25-world-news-2018-2023/code

Ashtiani, M. N., & Raahemi, B. (2023). News-based intelligent prediction of financial markets using text mining and Machine Learning: A systematic literature review. *Expert Systems with Applications*, *217*, 119509. https://doi.org/10.1016/j.eswa.2023.119509

Bollen, J., Mao, H., & Zeng, X. (2010). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

Brownlee, J. (2021, July 6). *A gentle introduction to long short-term memory networks by the experts*. MachineLearningMastery.com. https://machinelearningmastery.com/gentle-introduction-long-short-term-memory-networks-experts/

Chandola, D., Mehta, A., Singh, S., Tikkiwal, V. A., & Agrawal, H. (2022). Forecasting directional movement of stock prices using Deep Learning. *Annals of Data Science*, *10*(5), 1361–1378. https://doi.org/10.1007/s40745-022-00432-6

Cohen, I. (2021, May 23). *Explainable AI (XAI) with SHAP - regression problem*. Medium. https://towardsdatascience.com/explainable-ai-xai-with-shap-regression-problem-b2d63fdca670

Coursesteach. (2023, September 25). *Deep Learning (Part 7)-Logistic Regression Cost Function*. Medium. https://medium.com/@Coursesteach/deep-learning-part-7-6e78057a9ca6#:~:text=The%20logistic%20regression%20cost%20function%2C%20also%20known%20as%20the%20log,or%20the%20log%20loss%20function.

Dobilas, S. (2022, March 5). *LSTM Recurrent Neural Networks — How to Teach a Network to Remember the Past*. Medium. https://towardsdatascience.com/lstm-recurrent-neural-networks-how-to-teach-a-network-to-remember-the-past-55e54c2ff22e

Dolphin, R. (2020, October 21). *LSTM Networks | A Detailed Explanation*. Medium. https://towardsdatascience.com/lstm-networks-a-detailed-explanation-8fae6aefc7f9

Dong, H., & Gil-Bazo, J. (2020). Sentiment stocks. *International Review of Financial Analysis*, *72*, 101573. https://doi.org/10.1016/j.irfa.2020.101573

Eskandar, S. (2023, April 25). *Exploring Feature Extraction Techniques for Natural Language Processing*. Medium. https://medium.com/@eskandar.sahel/exploring-feature-extraction-techniques-for-natural-language-processing-46052ee6514

Genc, Z. (2020, July 31). *Finbert: Financial sentiment analysis with bert*. Medium. https://medium.com/prosus-ai-tech-blog/finbert-financial-sentiment-analysis-with-bert-b277a3607101

Harshith. (2019, November 20). *Text Preprocessing in Natural Language Processing*. Medium. https://towardsdatascience.com/text-preprocessing-in-natural-language-processing-using-python-6113ff5decd8#:~:text=Significance%20of%20text%20preprocessing%20in%20the%20performance%20of%20models.&text=Data%20preprocessing%20is%20an%20essential,process%20of%20building%20a%20model.

Huang, A. H., Wang, H., & Yang, Y. (2020). FinBERT - A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research*. https://doi.org/10.1111/1911-3846.12832

Hutto, C., & Gilbert, E. (2015). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Jones, A. B. (2018, December 2). *Sentiment analysis on reviews: Feature Extraction and Logistic Regression*. Medium. https://medium.com/@annabiancajones/sentiment-analysis-on-reviews-feature-extraction-and-logistic-regression-43a29635cc81

Kundu, R. (2022, December 16). *F1 Score in Machine Learning: Intro & Calculation*. V7Labs. https://www.v7labs.com/blog/f1-score-guide

Mehta, Y., Malhar, A., & Shankarmani, R. (2021). Stock price prediction using machine learning and sentiment analysis. *2021 2nd International Conference for Emerging Technology (INCET)*. https://doi.org/10.1109/incet51464.2021.9456376

Molnar, C. (2023, August 21). *Interpretable Machine Learning*. 9.6 SHAP (SHapley Additive exPlanations). https://christophm.github.io/interpretable-ml-book/shap.html

Pagolu, V. S., Reddy, K. N., Panda, G., & Majhi, B. (2016). Sentiment analysis of Twitter data for predicting stock market movements. *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*. https://doi.org/10.1109/scopes.2016.7955659

Pant, A. (2019, January 22). *Introduction to Logistic Regression*. Towards Data Science. https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148

Puh, K., & Bagić Babac, M. (2023). Predicting stock market using natural language processing. *American Journal of Business*, *38*(2), 41–61. https://doi.org/10.1108/ajb-08-2022-0124

Saxena, S. (2023, October 25). *What is LSTM? Introduction to Long Short-Term Memory*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/

Shah, D. V., Dashora, M., Churamani, N., & Prasad, B. (2022). Stock price prediction using LSTM-Arima hybrid neural network model with sentiment analysis of news headlines. *2022 International Conference on Futuristic Technologies (INCOFT)*. https://doi.org/10.1109/incoft55651.2022.10094422

Souma, W., Vodenska, I., & Aoyama, H. (2019). Enhanced news sentiment analysis using Deep Learning Methods. *Journal of Computational Social Science*, *2*(1), 33–46. https://doi.org/10.1007/s42001-019-00035-x

Usmani, S., & Shamsi, J. A. (2021). News sensitive stock market prediction: Literature review and suggestions. *PeerJ Computer Science*, *7*. https://doi.org/10.7717/peerj-cs.490

Vicari, M., & Gaspari, M. (2020). Analysis of news sentiments using natural language processing and deep learning. *AI & SOCIETY*, *36*(3), 931–937. https://doi.org/10.1007/s00146-020-01111-x

Wen, Y., & Arbogast, I. (2021, March 21). *How COVID-19 Has Impacted Stock Performance by Industry*. Saint Louis Fed Eagle. https://www.stlouisfed.org/on-the-economy/2021/march/covid19-impacted-stock-performance-industry

Wood, T. (n.d.). *F-Score*. DeepAI. https://deepai.org/machine-learning-glossary-and-terms/f-score

Xing, F. Z., Cambria, E., & Welsch, R. E. (2017). Natural language based financial forecasting: A survey. *Artificial Intelligence Review*, *50*(1), 49–73. https://doi.org/10.1007/s10462-017-9588-9

Zhang, Q., Qin, C., Zhang, Y., Bao, F., Zhang, C., & Liu, P. (2022). Transformer-based attention network for stock movement prediction. *Expert Systems with Applications*, *202*, 117239. https://doi.org/10.1016/j.eswa.2022.117239