Comparing sentiment analysis models and topics on Elon Musk tweets and its correlation with stock price

Anonymous ACL submission

Text Mining Project (732A81)
Author: Dinuke Jayaweera (dinja628)

Abstract

We will be using sentiment analysis of all the tweets of Elon Musk during a given time frame using three models of sentiment analyzers. By doing this we get the general emotion of Elon Musk. From this, we quantify his emotions and compare them to the performance of the company Tesla which he is the CEO of. Since the data we use is time series based, we have to quantify correlation using appropriate methods. We choose Euclidean distance. Since the sentiment and stock price could have some lagging or leading correlation, we will also use Dynamic Time Warping [1] and Pearson correlation. We get that in general there is a slight lagging correlation. To investigate this further, we do sentiment analysis on specific topics of the tweets to see if some topics such as "Twitter" or "Tesla" gives valuable information on the correlation, which we see that they do.

1 Introduction

We will be using sentiment analysis on Twitter data of Elon Musk's tweets to see a correlation. This is of interest since a positive correlation could contribute to financial rewards for stock market traders if we can by analyzing the current sentiment, choose winning buy-sell targets for the stock trader. In this case, we will be doing the correlation of the sentiment with the Tesla stock price. This since Elon Musk is the CEO of Tesla and also happens to tweet quite a lot. We will use a sentiment analyzer from vaderSentiment to get polarity scores to classify tweets as positive, negative or neutral. [2] Furthermore, we will also use BERT which also outperforms other more sophisticated models [3] and Logistics Regression models

2 Theory

2.1 VADER

As our main baseline sentiment analysis tool we use VADER (Valence Aware Dictionary and sEntiment Reasoner). We use the analysis tool vaderSentiment for python [4] This is a lexicon-based approach.

2.2 BERT

BERT stands for "Bidirectional Encoder Representation from Transformers". We use pre-trained BERT which is already trained on a large text corpus. This combined with fine-tuning it for the specific task at hand can give us a powerful model to use. [3] This is a machine-learning approach. We follow Tensorflow's official tutorial to implement the fine-tuning. [5] The fine-tuning code is in the repository with the name "Final-BERT-FuneTuning"

2.3 Logistic Regression

Logistic regression is commonly used for dichotomous cases, we in this task will use it with its multinomial case with 3 categories [6] in our case is positive, negative and neutral. This is a machine-learning approach.

2.4 Euclidean Distance

We will use the Euclidean distance to see for similarities between the two time-series. [7]

2.5 Pearson Correlation

The time series of interest can be lagging or leading if there is a correlation, in this case, we will use the Pearson correlation coefficient to get in what direction and strength the correlation is tilted. [7]

3 Data

The data used for sentiment analysis is tweets from Elon Musk on Twitter. To fine-tune the pre-trained BERT model, we use pre-labeled data, in this case, we found pre-labeled sentiment data that is based on Reddit data. Both of these data are downloaded from Kaggle [8] and [9] respectively. Twitter data is very useful in this task since Elon Musk usually tweets a lot about a variety of subjects and aspects of life. This will let us do a useful sentiment and emotion analysis of Elon Musk. We will only use the date and text columns of the Twitter data. The stock data has a lot of information about the stock but we are interested in only the stock price and date so we filter it to either the opening or closing stock price and the date of that trading day.

The data used to see if there is a correlation between sentiment and stock price is the stock price of the company Tesla the data is from MarketWatch [10]

4 Method

4.1 data

The data is handled differently depending on the model being used.

For the Vader sentiment, we choose to go through each row of the tweets data and give in a score of 1, -1 or 0 for positive negative and neutral respectively.

For the BERT model, we use pre-labeled Twitter data and fine-tune it using Keras Tensorflow. After that, we load the finished model and use it to get the sentiment of the Elon Musk Twitter data.

For the Logistic Regression, we use the pre-labeled Twitter data and split it into training and testing. We train the data on the model and Logistic Regression will suit the nature of this type of classification. [11]

4.2 normalization

When the sentiment data for all the models are accumulated and in series. We will see a clear trend usually. To deal with this fact, we use linear de-trending, for both stock price and also sentiment. This will give us a better view to analyze the correlations to scale. This is a unique method used in this project which has not been seen widely used for many sentiment-to-stock analysis projects seen online. This fact greatly differentiates the findings of this project from others that do not do the accumulate and de-trend part.

5 Results

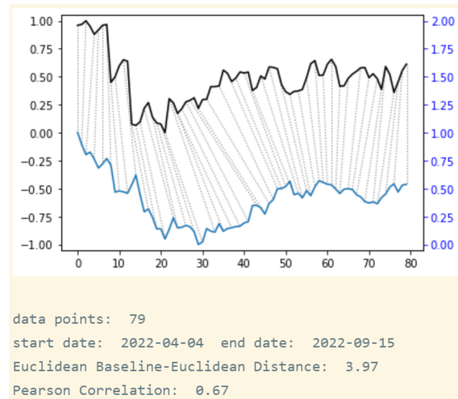

Figure 1: Using BERT sentiment analysis on all tweets. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

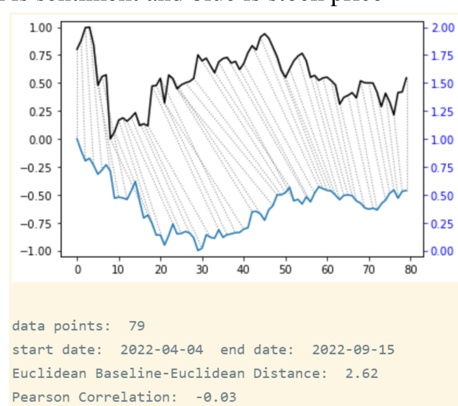

Figure 2: Using logistic regression model, sentiment analysis on all tweets. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

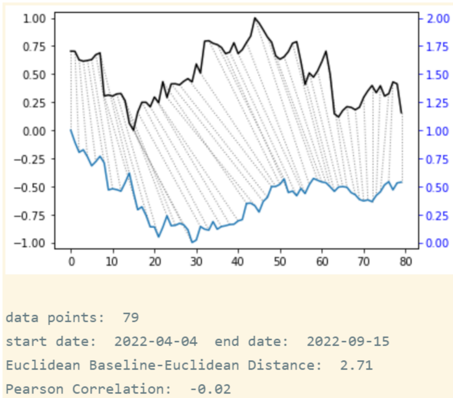

Figure 3: Using Vader sentiment analysis on all tweets. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

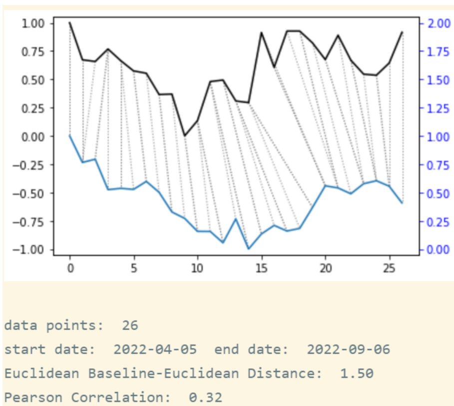

Figure 4: Using BERT sentiment analysis on tweets mentioning Twitter. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

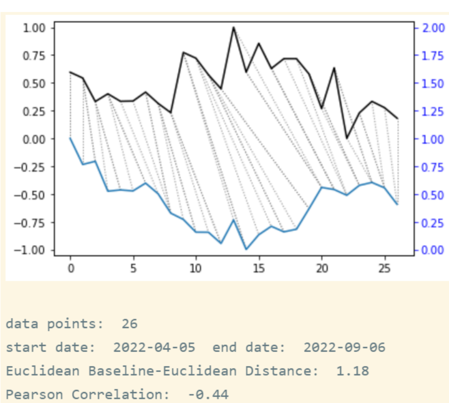

Figure 5: Using logistic regression model, sentiment analysis on tweets mentioning Twitter. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

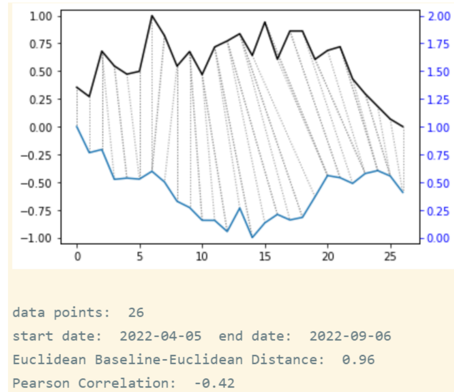

Figure 6: Using Vader sentiment analysis on tweets mentioning Twitter. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

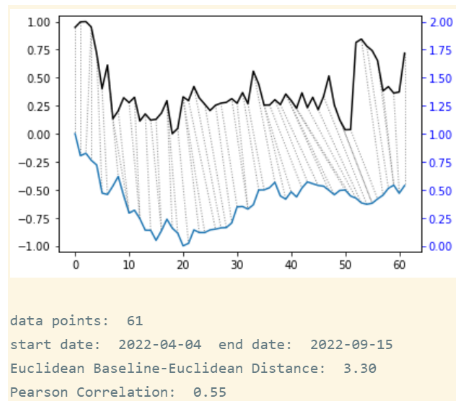

Figure 7: Using BERT sentiment analysis on tweets mentioning Tesla. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

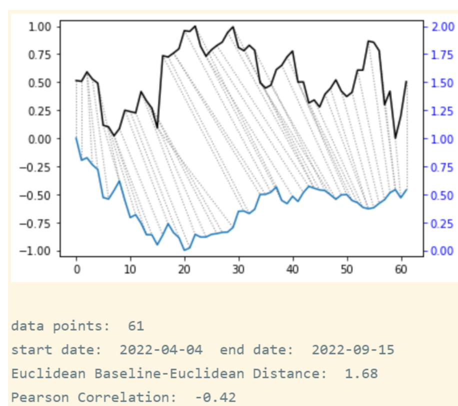

Figure 8: Using logistic regression model, sentiment analysis on tweets mentioning Tesla. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

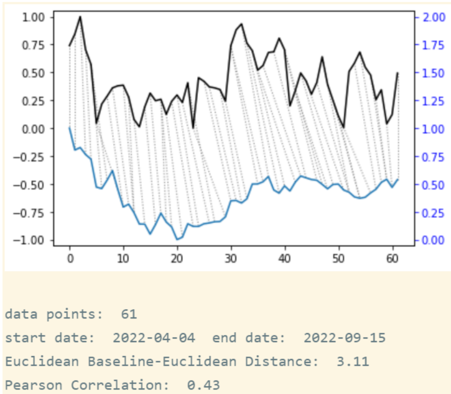

Figure 9: Using Vader sentiment analysis on tweets mentioning Tesla. By using dynamic time warping, we get a plot where black is sentiment and blue is stock price

6 Discussion

To make this discussion clear, we will be assuming that the Vader Sentiment is our baseline. Another important clarification to be made is that when we refer to "sentiment" in the case where we use all the tweets of Elon Musk, we mean that it is emotion analysis since using all the tweets implies that we are measuring Elon Musk's general mood and emotional state. These results are presented in figures 1 to 3. However, for the results of figures 4 to 9, we are doing a sentiment analysis of Elon Musk on the given topic. First, we look at the results looking at figures 1,2 and 3 which are using all the tweets of Elon Musk and the stock price of Tesla during 2022-04-04 and 2022-09-15. Since we have the most data points on these results we will use them to compare Vader Sentiment which is a lexicon-based approach, logistic regression and BERT where the latter two are machine learning based. Starting our observations of the resulting dynamic time-warping plot, we see a possible correlation where the stock price seems to be following Elon Musk's tweets sentiment somewhat. This makes us hope full that we can predict the stock price by using sentiment or emotion analysis. Since the plots seem quite similar between the three models, we will look deeper into some actual correlation values to get a better understanding of them. We see that the Vader and Logistic models have more similar Pearson correlation coefficient and Euclidean score. Whereas the BERT model seems to find a better correlation. Since the BERT is deviating from the other two models, it might be observing wrong correlations. Now that we do observe a difference between the models where we have a higher trust for the Vader and Logistic models, we will investigate if a certain topic of Elon Musk's tweets has a higher or lower correlation with the stock price of Tesla. We choose two topics to compare between, "Twitter" and "Tesla". The choice came down to the fact that Elon Musk during this period was talking actively about acquiring Twitter the company and this could have caused his Tesla company's stock value to fluctuate depending on that. Furthermore, the choice of using Tesla is a natural decision since this topic is also the name of the company of this analysis. Comparing the sentiment analysis vs stock price of these two topics, we see that the Pearson correlation coefficient which indicates the lag shift of the correlation seems to be quite similar between the two topics, around 0.4 in either direction. However, when looking at the Euclidean score, we see that the topic "Tesla" has a range of 1.68 to 3.30 whilst the topic "Twitter" has a significantly lower range between 0.96 to 1.50.

7 Conclusion

After investigating the correlation between sentiment and stock price using three different models for both specific topics and the general emotion analysis of Elon Musk's tweets without any specific topic. We can conclude that there seems to be some sort of correlation in the desired direction which is that the stock price follows the sentiment of Elon Musk's tweets. Furthermore, investigating if specifying certain topics have more correlation than others, we found that the topic of Tesla had a greater correlation than the topic of Twitter. However the topic of Tesla did not have a significantly higher correlation to the stock price than the general emotion analysis of Elon Musk's tweets without a given topic.

It is hard to say from the specific values of the Pearson correlation and Euclidean distance itself if it is worth the risk to do any buy-sell stock trades based on Elon Musk's emotion or sentiment on specific topics. This is out of the scope of this project but will be interesting for further studies to do further investigation on how this information can be used to do reliable stock trades.

References

[1] Toni Giorgino. Computing and visualizing dynamic time warping alignments in r: the dtw package. *Journal of statistical Software*, 31:1–24, 2009.

[2] Shihab Elbagir and Jing Yang. Twitter sentiment analysis using natural language toolkit and vader sentiment. In *Proceedings of the international multiconference of engineers and computer scientists*, volume 122, page 16, 2019.

[3] Manish Munikar, Sushil Shakya, and Aakash Shrestha. Fine-grained sentiment classification using bert. In *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, volume 1, pages 1–5, 2019.

[4] Clayton Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 216–225, 2014.

[5] Classify text with BERT. https://www.tensorflow.org/text/tutorials/classify_text_with_bert. Accessed: 2023-01-15.

[6] Raymond E Wright. Logistic regression. 1995.

[7] Michael R. Berthold and Frank Höppner. On clustering time series using euclidean distance and pearson correlation, 2016.

[8] Twitter sentiment analysis. https://www.kaggle.com/code/paoloripamonti/twitter-sentiment-analysis/data. Accessed: 2023-01-15.

[9] Reddit data sentimental analysis. https://www.kaggle.com/code/deepakvedantam/reddit-data-sentimental-analysis/data?select=Reddit_Data.csv. Accessed: 2023-01-15.

[10] Tesla. https://www.marketwatch.com/investing/stock/tsla/download-data. Accessed: 2022-12-22.

[11] Ravinder Ahuja, Aakarsha Chug, Shruti Kohli, Shaurya Gupta, and Pratyush Ahuja. The impact of features extraction on the sentiment analysis. *Procedia Computer Science*, 152:341–348, 2019. International Conference on Pervasive Computing Advances and Applications- PerCAA 2019.