Analysing sentiment on BioNTech/Pfizer COVID-19 vaccine from tweets

Fabian Afatsawo

Abstract

In this report, we take a look at the sentiment on the BioNTech/Pfizer COVID-19 vaccine. By means of sentiment analysis of tweets relating to the vaccine. Theretofore, two pre-trained language processing models have been used. These models are: NLTK VADER and HugginFace Transformers. In addition to that, tweet engagement is taken into account. As a result of insufficient and unlabeled data, no clear verdict is found.

1 Credits

Griffin Leow has done great work for scraping tweets [1]. His code has been used for scraping. The Twitter API is not accessible to everyone. A Twitter Developer account is necessary to get access to the API. This can be requested to Twitter.

2 Introduction

To this day the COVID-19 virus had a huge impact on the whole world. There are a lot of people working on the battle against the virus. One way is to create vaccines. Vaccines teach the immune system to create specific antibodies that protect the body from diseases. Therefore, many organizations have been trying to make vaccines that will help us beat COVID. In the meantime, effective vaccines have been found. Some of which are approved and others are under review.

2.1 Vaccination demand

There are first and foremost two crucial factors about the vaccine. At first supply and secondly demand. In order to vaccinate the majority of people, vaccines need to be produced in large quantities. On the other hand, vaccines only are effective when people are willing to take them. If vaccines help in preventing people to get seriously ill, why won't they want them? There is scepticism among the vaccines. Some argue that the vaccines are too quickly developed. Others think that the long term effects are unknown. There are even people that do not believe in vaccines in general. Because of this scepticism, it is important to assess the sentiment about vaccines. We focus on one vaccine, namely BioNTech/Pfizer. This RNA vaccine teaches the immune system to create the necessary antibodies to tackle the COVID-19 virus. The vaccine is approved in The Netherlands and will be used to start the vaccine program [2].

2.2 Data

Nowadays, people love to express their opinions on social media. This information can be used as data to derive sentiment. Twitter has 340 million users [3], with over 500 million tweets that are being posted daily [4]. In this report, N tweets have been scraped in relation to the vaccine. This scraping is done through the use of Tweepy. Tweepy is a Python library for accessing the Twitter API [5].

2.3 Sentiment Analysis

On the obtained data, sentiment analysis can be performed. Sentiment Analysis is a branch of Natural language processing. With this, you aim to interpret and classify emotions based and text. Sentiment analysis is done by training a model with a large amount of labeled data. The labels are emotions. There are pre-trained models available that can be used for analysis. Two pre-trained models used in this report, namely NLP VADER [6] and HuggingFace Transformers [7].

2.4 Research question

In this report, the central question is: 'What is the general sentiment around the BioNTech/Pfizer COVID-19 vaccine?'.

3 Methods

3.1 Scraping

The Tweepy library is used to retrieve the data. The library has functions at which you can search for different keywords. Multiple hashtags relating to the vaccine have been used. Examples are: #pfizer, #biontech, #pfizerbiontech and #pfizervaccin.

3.2 Sentiment analysis

3.2.1 NLTK VADER

NLTK has a sentiment analyzer called VADER (Valence Aware Dictionary and sEntiment Reasoner. The analyzer uses a lexical approach. This means it uses words or vocabularies that have been assigned predetermined scores.

The VADER analyzer is broadly used on social media. The accuracy is high on short text, but decreases with larger texts.

Pros:

- Fast
- No pre-processing
- Easy to use
- Use of 3 classes
- Short text accuracy

Cons:

- Does not take context into account
- Long text accuracy

3.2.2 HuggingFace Transformers

Transformers provides pre-trained models for Natural Language Understanding (NLU) and Natural Language Generation (NLG) tasks. Transformers has interoperability between PyTorch and TensorFlow.

Pros:

- No pre-processing
- Easy to use
- Takes context into account

Cons:

- Relatively slow
- Limited to 2 classes

4 Results

Figure 1 shows that the VADER model has an evenly distribution with a small majority (41%) towards the positive sentiment. Interesting!!!

4.1 Sentiments

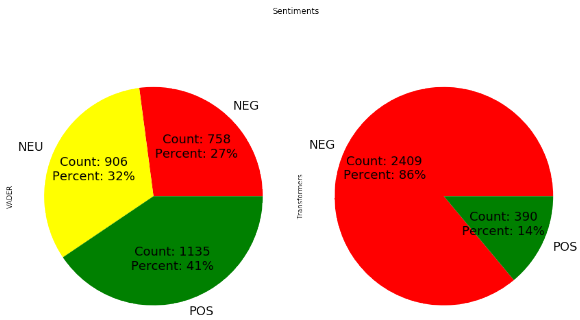

Figure 1: Sentiment results of both models.

Both models show different results. The Transformers model is limited to two classes namely, positive (POS) and negative (NEG). While the VADER model has three classes adding by a neutral class. In absence of the neutral class, the tweets assigned to this class are removed.

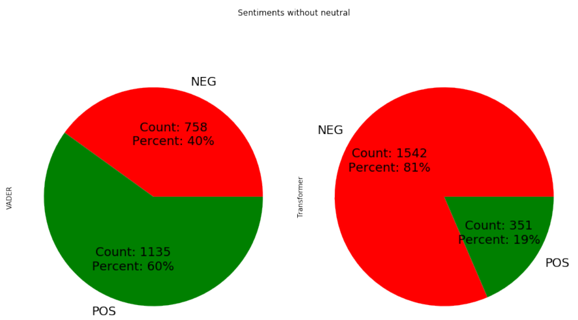

Figure 2: Sentiment results discarding neutral labels.

Figure 2 shows that both models have different views. The VADER model gives the majority of the tweets (60%) a positive sentiment. On the other hand, the Transformers model tends to give an even bigger majority (81%) a negative label. However, the models have a consensus on (37.5%) of the data. Here, both models assign equal labels to the tweets. Judging from figure 3 the consensus shows that the majority of tweets (69%) have a negative sentiment.

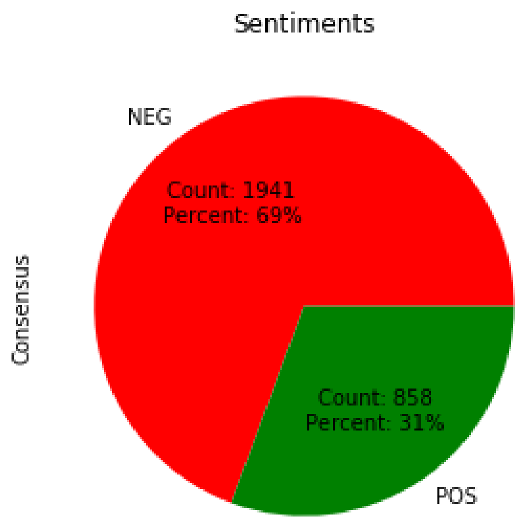

Figure 3: Sentiment results from consensus of both models.

4.2 Engagement

Engagement captures the number of interactions a tweet has. For a tweet, this will be views, favorites and retweets. We only have data for favorites and retweets, these will be used to define engagement. A lot of people resonate with a tweet when it has a high engagement. Therefore engagement can help in finding the sentiment about the virus.

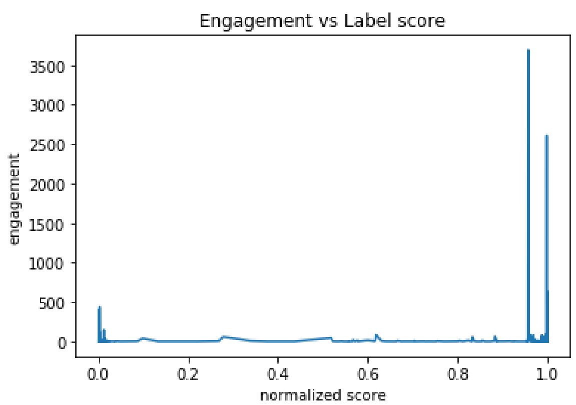

Figure 4: Engagement vs Transformers label score.

In fig N Engagement is mapped against Transformers label score. The raw score ranges from $0.5 - 1.0$ for both 'NEG' and 'POS' labels. This score is normalized by mapping it onto a scale of $0.0 - 1.0$. Whereby, 0.0 is negative and 1.0 is positive.

It is clear to see that the most engagement is located at the extremes. However, it also shows that positive tweets have more engagement. In terms of engagement, positive side has a ratio of 0.68 against 0.32 on the negative side. Let's look at some examples of tweets with high engagement.

```
------------------------------
#Pfizer's second dose didn't seem to help

https://t.co/87SmXYRq3A
Retweets: 119 | Favorites: 162
------------------------------
```

Figure 5: Negative tweet

```
------------------------------
my mom has been 48 hours since receiving the #Pfizer vaccine.
She had a little nausea; that's it! Now feels perfect.
Retweets: 13 | Favorites: 621
------------------------------
```

Figure 6: Positive tweet

By assessing these two examples, the correct labels are given to these tweets.

5 Discussion

Judging from the consensus between the models, the general sentiment around the vaccine is negative by two-thirds. Although judging from the engagement, the sentiment tends to be positive by two-thirds. These results are completely the opposite. What can be the reason for this? Is the data not significant? Or are the models not performing well?

The data may be insufficient. There are 2799 tweets scraped, this is not much. The Twitter API has a time frame restriction. With this, it allows scraping for tweets as much as 7 days back. More data means more information to analyse. In addition, the data is not labelled. The correct labels are unknown because no human assessment has taken place.

The models show very different results, with one migrating to positive and the other towards negative sentiment. What can be the cause of this difference? At first, the models are fundamentally different. NLTK VADER uses a lexical approach which does not include context. Negation can change the polarity of a sentence. For example the sentence: 'This is not so cool'. Here in this example, 'so cool' is positive. In essence is the sentence negative because of the word 'not'. Other than that, sarcasm and irony may be misinterpreted.

6 Conclusion

What is the general sentiment around the BioNTech/Pfizer COVID-19 vaccine? According to the consensus between the models, the general sentiment around the vaccine is two-thirds negative. Although judging engagement, the sentiment is two-thirds positive. This outcome is not clear enough to make a clear judgment on the question. There is insufficient data, more information might lead to a clearer result. Human assessment on the correct labels of the data can help in determining the performance of the used models.

References

[1] *Scraping Tweets with Tweepy Python, Medium https://medium.com/python-in-plain-english/scraping-tweets-with-tweepy-python-59413046e788*

[1] *BioNTech/Pfizer information https://www.rijksoverheid.nl/onderwerpen/coronavirus-vaccinatie/soorten-coronavaccins/biontech-pfizer-comirnaty*

[3] *Twitter users https://blog.hootsuite.com/twitter-demographics/*

[4] *Twitter daily amount of tweets https://www.dsayce.com/social-media/tweets-day/*

[5] *Tweepy https://www.tweepy.org/*

[6] Hutto, C.J. Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

[7] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... Rush, A. (2020). Transformers: State-of-the-Art Natural Language Processing. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 1. https://doi.org/10.18653/v1/2020.emnlp-demos.6