Covid-19 vaccines : how do we feel about it?
Machine Learning for Natural Language Processing 2020

Ghiorghita Eliza
Ensae Paris
eliza.ghiorghita@ensae.fr

Lesauvage Romain
Ensae Paris
romain.lesauvage@ensae.fr

Abstract

We analyze here the global sentiment of the public discourse on Twitter concerning the coronavirus vaccines, as well as its declinations by types of vaccine. The evolution of the public sentiment seems related to the main world news on Covid-19.

1 Problem Framing

In this project[1], we are working on the analysis of sentiment about coronavirus vaccines. We based our work on social media data - nearly 40,000 observations collected from Twitter and available on Kaggle[2]. This daily updated database contains tweets concerning Covid-19 vaccines. Our project unfolds in a two-steps task, using NLP techniques, applied to our corpus : firstly, a sentiment is associated to each tweet; then we split the corpus in 4 classes, according to the type of vaccine they refer to. The final purpose of these methods is analysing the time evolution of the global attitude towards vaccines in our corpus -and how this attitude declines by category of vaccine. The notebook linked to the project can be found here. here.

2 Experiments Protocol

2.1 Sentiment analysis

We start our study by labelling a sentiment analysis to each tweet. Since our database is not annotated, we cannot use it to train our model, so we choose a second annotated tweets database also available on Kaggle[3]. We divide this dataset into one train and one test dataset, on which we will calculate all our indicators.

We start with a first baseline model: for each word, we calculate the percentage of positive tweets in which it appears, this gives us an average sentiment per word. Then we predict, for each tweet, its sentiment by averaging the sentiment of the words. This model leads to an accuracy of **77%** on the annotated dataset.

We then study an actual NLP model using the BERT model for the purpose of sentiment analysis. We start by formatting our dataset and proceed to the fine-tuning of the BERT model in our case[4]. We add a linear layer to the BERT model in order to perform a classification, and we obtain an average accuracy of **92%** on the annotated dataset, much higher than the baseline model.

We then apply this trained model on our dataset of tweets about the vaccines. We decide to look at two versions: one with a positive or negative sentiment, another one adding a category with neutral tweets. We prefer to focus on the second version, as it provides additional information.

2.2 Typology by type of vaccine

We then tried to classify tweets by types of vaccines they refer to : either a particular type of vaccine ("MessengerARN", "viral vectors", "inactive" and "Protein subunit" vaccines) or vaccines in general ("none of the above" or "more than one").

The first basic typology, containing 2-3 of the more common terms associated with each type gave quite unsatisfying results when applied on the tokenized corpus : only 57,0% of tweets were identified as mentioning one type of vaccines or more.

We then refined our initial typology by looking for other possible missing terms (misspellings, online adaptation of the original terminology a.s.o). Applying word-embedding methods to our

[1] You can find our github here and our Google Colab here.
[2] All Covid-19 vaccines dataset from Kaggle.
[3] Sentiment analysis dataset from Kaggle.
[4] We also test the best parameters for the model, by using a grid-search for the learning rate for instance. We decide to use the BCE loss and the Adam optimizer

corpus (word2vec algorithm) allowed us to identify the most similar words to the terms used in the first typology . We then took case-by-case decisions on which terms to keep and which ones to let go (for example take 'jandj' but not 'j' as too vague). This second, enriched typology gave better results : **74,8%** of tweets were classified as speaking about one or more types of vaccines.

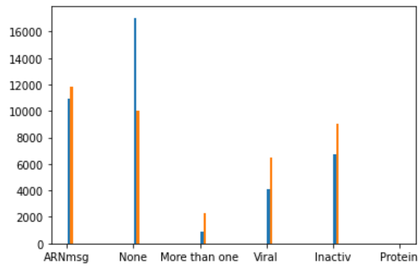

Figure 1: Tweets classification by using basic and enriched typology

Finally, we tried a third typology, by computing an average embedding on which tweet, based on the similarity of each one of its words to the 4 different classes.This gave less satisfying results : only 26,3% of tweets are classified out of "General". So we'll further use our second typology, combining word-embedding algorithm and user's decision, as it seems to be the most appropriate one (figure 1).

3 Results

Firstly, we take a look at the sentiment evolution during the period when messages were posted (12/20-03/21), by focusing on the proportion of positive tweets across the time[5] (figure 2).

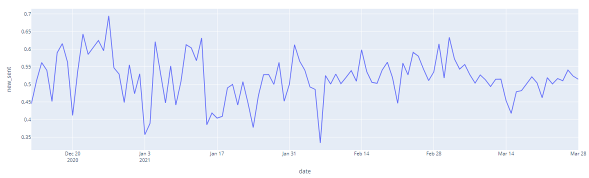

Figure 2: Timeline of positive tweets about COVID-19 vaccines

We notice an unsteady evolution of the proportion of tweets expressing a positive attitude towards vaccines. Highs and lows mainly occur at moments coinciding with public news and decisions -like the following negative peaks :

- **December 20th** : several countries in EU decide to close their boarders with the UK, being suspicious about the English variant;
- **January 3rd** : UK announced a new lockdown due to the variant;
- **February 6th** : worries about variants in NYC.

Since February, there is an improvement of the global attitude (positive trend), related to the massive access to vaccines and procedures put in place all around the world.

In order to specify our analysis, we decide to look at the differences of time evolution in the sentiment for MRNA, Inactiv and Viral vaccines (figure 3).

Figure 3: Timeline of positive tweets about COVID-19 vaccines

As a reminder, the MRNA vaccines are those from PfizerBiontech and Moderna, the viral ones are from AstraZeneca and Johnson&Johnson, and the inactive ones are Sinopharm/Sinovac etc. We observe that the trend for MRNA vaccines seems to be constant whereas the trend is very fluctuating for the two others and, in particular, for the viral vaccines we have a decrease of popularity. It can be explained by the recent news on thrombosis cases, in reaction to injections of AstraZeneca or Johnson&Johnson vaccine.

4 Discussion/Conclusion

In conclusion, after assigning to each tweet a sentiment and a type of vaccine to which it referred, we were able to analyze on the one hand the evolution of the sentiment of adherence of the population to vaccines, as well as to detail this analysis by type of vaccine. The results show that the sentiment created by the tweets reflects major events related to Covid-19 and that the recent trend of improving sentiment follows good news from countries where vaccination has worked. Different trends are identified by type of vaccines : AstraZeneca types vaccines meet a growing public distrust while MRNA are steadily more popular.

[5] We do not look at the evolution of neutral and negative tweets because there are few neutral tweets so the positive tweets contain almost all the information.