A Classification Approach to Detect Public Sentiments towards COVID-19 Vaccines

Karabo Johannes Ntwaagae, Nkwebi Peace Motlogelwa, Edwin Thuma, Tebo Leburu-Dingalo and Gontlafetse Mosweunyane

Department of Computer Science, University of Botswana

Abstract
In this paper, team University of Botswana Computer Science (UBCS) investigate the opinions of Twitter users towards vaccine uptake. In particular, we build three different text classifiers to detect people's opinions and classify them as *provax*-for opinions that are for vaccination, *antivax* for opinions against vaccination and *neutral*-for opinions that are neither for or against vaccination. Two different datasets obtained from Twitter, 1 by Cotfas and the other by Fire2022 Organizing team were merged to and used for this study. The dataset contained 4392 tweets. Our first classifier was based on the basic BERT model and the other 2 were machine learning models, Random Forest and Multinomial Naive Bayes models. Naive Bayes classifier outperformed other classifiers with a macro-F1 score of 0.319.

Keywords
Vaccination, BERT classification, Sentiment Analysis, COVID-19

1. Introduction

In recent years, there has been rapid increase in internet access and usage all over the world. This increase in internet access has resulted in large volumes of structured and unstructured data being deposited on online repositories [1]. The increase in media sharing across multiple social media platforms has sparked interest from researchers and policy makers across the globe on mining and analyzing this data [2]. Twitter has been the mostly used platform for getting dataset used to provide informative conclusions on whether people share similar views and behavior on certain topics of interest. The increase in volume of this structured and unstructured data comes with a challenge of mining and analyzing the mined data [1]. To address this challenge, researchers have developed and employed different techniques to deal with both mining the data from social media platform and analyzing the data, of which sentiment analysis is one. According to Catapang et al. [3], sentiment analysis, a common natural language processing task, has been done on several major topics such as politics, health, economy, vaccinations, and sports to name but a few. Of recent, sentiment analysis has been done on studies that aimed at analyzing people's opinion on the outbreak of COVID-19. COVID-19 is a virus that causes

Forum for Information Retrieval Evaluation, December 9-13, 2022, India
ntwaagaek.ac.bw (K. J. Ntwaagae); motlogel@ub.ac.ub.bw (N. P. Motlogelwa); thumae@ub.ac.bw (E. Thuma); leburut@ub.ac.bw (T. Leburu-Dingalo); mosweuny@ub.ac.bw (G. Mosweunyane)
https://www.ub.bw/connect/staff/830 (N. P. Motlogelwa); https://www.ub.bw/connect/staff/1966 (E. Thuma); https://www.ub.bw/connect/staff/202 (T. Leburu-Dingalo); https://www.ub.bw/connect/staff/1379 (G. Mosweunyane)

severe respiratory problems and was first detected in China, Wuhan Province in December 2019. China was later hit hard by the virus while the rest of the world felt the immense impact of the virus in early 2020. This outbreak was later declared a global pandemic by the World Health Organization (WHO) in 2020. Since its outbreak, COVID-19 has claimed over 6.5 million lives to date and the virus was detected in over 613 million people (positive cases) worldwide [4]. As COVID-19 presented a greater threat to the entire population, countries all over the world implemented some measures and restrictions on their people as a way of trying to reduce the spread of the virus [5]. Some of the restrictions which included lockdowns and working from home, resulted in a lot of internet users spending most of their time on their gadgets as a way of reducing boredom, getting updates on the virus as well as sharing their views on different social media platforms on issues relating to COVID-19. This availed lots of media on social media platforms which was used by researchers to analyse the impact of the restriction on the society as well as making recommendations on what could work to curb the spread of the virus. Although the restrictions worked to some extent to reduce the spread, scientists and doctors believed the most effective and fast way of reducing the pandemic is vaccination [6]. In their study carried out in 2020, [5] argued that vaccinations are a way of developing immunity against pandemics such as COVID-19 if the world reach herd immunity.

For the world to attain herd immunity or break the spread of the COVID-19 pandemic, Trapman [7] and Frontanet et al. [8] highlighted that a vaccination rate of at least 65% to 67% must be reached. According to the currents statistics reported in [9], over 4.9 million people have been fully vaccinated around the world, which makes up to 62% vaccination rate, 3% short of the 65% of attaining the herd immunity. The society has since had different views on vaccinations which resulted in the slow rollout due to different reason ranging from politics to the safety of the vaccines. The hesitancy to take vaccines may lead to the pandemic taking years to end as countries are continuing to record new cases daily. This paper aims to analyze the opinions of the public on COVID-19 vaccines by considering the data collected from Twitter (provided by the FIRE2022 organizing team). The tweets were then classified using the basic BERT model with Hugging face, random forest technique and Naïve Bayes technique.

2. Related Work

Sentiments Analysis is becoming a widely used natural language processing technique to mine people's opinions from different media sharing platforms. It has been used in many social issues in the past to cope with mining people's opinions from text [1]. Kitchat et al. [10] defines sentiments analysis as a natural language processing technique used to mine the people's opinions or the underlying tone of the given data. Due to its wide usage, various sentiment analysis methods have been developed to analyse the data mined from different social media platforms. In previous studies, sentiment analysis has been used to get people's opinions on different social issues such as the perception of economic weakening, measuring the impact of movement bans on society, measuring the desire to stock products under the pandemic and the vaccination of children to mention a few [3, 6].

A study by Machucha et al. [11] used sentiment analysis to get the people's general opinions on corona virus using binary logistic regression. They classified the sentiments as positive and negative only. After extracting the data, they preprocessed it by removing all punctuation marks, transforming each tweet to lowercase and tokenizing tweets to facilitate the removal of non-English words and stopwords. The tweets were stemmed and rejoined. TF-IDF vectorizer was used for vectorization in this study. When the evaluated the performance of their classifier, they obtained a classification accuracy of 78.5%.

Delizo et al. [12] conducted a study that uses sentiment analysis to examine the polarity of COVID-19 related opinions on Twitter using Multinomial Naïve Bayes algorithm, targeting tweets from Philippines. Their preprocessing stage included removing special characters, hyperlinks, hashtags and mentions, and removing tweets containing less than 2 words and less than 3 characters from each tweet. 10% of the cleaned tweets was manually labelled as either positive or negative Words were then converted to their lowercase equivalents. TF-IDF vectorizer was used for feature extraction from the training data. Multinomial Naïve Bayes algorithm was employed to test the effectiveness of the proposed it achieved an overall accuracy score of 72%.

The study of [5] focused on using aspect-base twitter sentiment analysis to classify the society's views on the vaccination and determine which vaccine type is preferred by the public from 8 countries being USA, UK, Canada, Turkey, France, Germany, Spain and Italy. They used TF-IDF and word2Vec applications to determine the aspects from the datasets and used 4 different BERT models (i.e mBERT-base, BioBERT, ClinicalBERT and BERTurk) for classification. In their evaluation, they reported F1 scores of between 84% and 88% with a classification accuracy of 87%. The vaccine type that received less negative sentiment was Novavax, developed by Pfizer/BioNTech.

In another study that aims to use sentiment analysis to analyses people's opinions on COVID-19 vaccine, Trapman [6] focused on using sentiment analysis to map the determinants of COVID-19 vaccine uptake from mining Twitter data. They focused on determining whether the 5As (based on a study by Thompson et al. [13]) being Affordability, Awareness, Acceptance, Activation and Access together with the addition of the sixth A, Assurance, could directly cover and organize all the determinants identified from tweets regarding COVID-19 vaccine uptake using the bottom-up approach. The dataset used in this study was collected from people's tweets using the Qualitative Data Analysis (QDA) software. The dataset put together by choosing keywords related to COVID-19 vaccines and combining them into an OR query. After preprocessing data, topic modelling was performed using WordStat software and a 33-topic model was seen as optimal, described by top-weighted keywords. In their final step, the authors linked the 17 determinants with the 6As. Their results indicate that the 6A taxonomy was successful in capturing all the determinants of COVID-19 uptake.

The study by Kitchat et al. [10] also used sentiment analysis to investigate people's opinions from tweets in New York City. Keywords relating to COVID-19 vaccines were used to filter out relevant tweets. To further filter out relevant tweets, the location from which the tweets were from was also specified and in this study, New York City was specified as the location of tweets. After collecting the data, they cleaned the data by removing hashtags, punctuation marks, URLs, emojis stopwords (through tokenization) and converting all letters to lowercase. The tweets were stemmed and rejoined after tokenization. The tweets were classified as positive, negative and neutral. To test the performance of the proposed system, they used 6 machine learning models, namely, Random Forest, Decision Trees, Support Vector Machine (SVM), Naïve Bayes, Logistic Regression and Multi-Layer Perceptron (MLP). According to their results, all machine learning models used performed exceptionally well, giving an accuracy of over 90% except for Naïve Bayes which registered an accuracy of 82.13%. Multi-Layer Perceptron model performed better than all other models, with an accuracy of 93.63

3. Methodology

In this Section, we present our BERT and machine learning approaches for classifying tweets into three classes, namely: ProVax, AntiVax and Neutral. AntiVax indicates that there is hesitancy in the tweet (of the user who posted the tweet) towards the use of vaccines. Neutral indicates that the tweet does not have any discernible sentiment expressed towards vaccines or is not related to vaccines. We propose using the basic BERT classifier and 2 machine learning classifiers provided by scikit-learn library [1], namely, Multinomial Naive Bayes classifier, which is a Nave Bayes variant used in text classification in python as well as the Random forest classifier.

3.1. Training Dataset

The training dataset was pre-processed to make it compatible with machine learning models and BERT model by converting words to numbers which are easily understood by the models. In addition, the following pre-processing steps were applied to the main dataset:

- For the BERT experiment, BERTokenizer provided by the Transformer library was used to convert the tweets into tokens and tokenIDs.
- For the machine learning algorithms, the TfidfVectorizer from scikit-learn library was employed to perform feature extraction from the training as well as stopwords removal.
- The training data was stemmed and rejoined.
- Hashtags and mentions not removed, as well as punctuations not removed.

The training dataset contains 4392 tweets. The dataset was obtained from Twitter by different teams as follows:

- **Cotfas dataset**: Crawled from Twitter between November-December 2020. 2792 tweet-texts were crawled along with tweetIDs and labels (proVax, neutral, antiVax). This crawled data formed the first part of the training data.
- **IRMiDis Fire2022 organizing team dataset**: Tweets crawled from March to December 2020. Tweets were annotated with three labels by crowdworkers. A total of 1600 tweets were crawled, and this formed the second part of the training data.

Of this 4392 tweets, 1676 are provax, 1081 are antivax and 1635 are neutral. During training, the training dataset was subdivided such that 3952 tweets train our classification models and 440 tweets are used for validation.

[1] https://scikit-learn.org

3.2. Testing Dataset

The same pre-processing done in the training dataset was performed on the test data set, except for pre-processing that deals with labelling the tweets as provax, neutral and antivax.

4. Description of Runs

We submit 3 runs for: Task 1: COVID-19 vaccine stance classification from tweets. Below is a brief description of each run:

4.1. Run 1 - UBCS

This is our baseline run. We used BERT model to build a BERT classifier for the identification of sentiments. Since BERT works with fixed-length sequences, we set a maximum length for the sequences, guided by the number of tokens in the tweets. Most tweets seemed to contain less than 140 tokens and we then set our maximum length to 150 to allow for tweets that may be longer than the ones we used for training during testing. We also set the batch size to be 16 because 16 and 32 are recommended by BERT authors as the best sizes for fine tuning the model. We built our sentiment classifier on top of the basic BertModel. In our classifier, we used a dropout layer for some regularization and a fully connected layer for our output. . Both the training and test dataset underwent the same pre-processing steps as described in Section 3.1.

4.2. Run 2 - UBCS and Run 3 - UBCS

In these runs, we used the Random Forest and Multinomial Naïve Bayes models respectively. We vectorized our tweets into numbers using TFIDFVectorizer. We defined variable **X**, to hold all the features to be used in predictions while variable **y** holds the labels. We then imported the Machine Learning models, Random Forest and Multinomial Naives Bayes (an implementation of Naïve Bayes algorithm in scikit-learn), from scikit-learn library using their default parameters. For all the three runs, we divided the data into training data(90% of the data), validation data(10%). Data loaders for the 3 sets of data were created for each set of data.

5. Results and Analysis

In this paper, we employ different text classifiers and compare their performance to determine the best classifier among them. Table 1 presents the results of our investigation. Run 3 - UBCS is our best run as it performed better than all the other runs in all the evaluation metrics, with a Macro F1 score and Accuracy of **0.319** and **0.337** respetively.

6. Discussion and Conclusion

The results of our investigation suggests no significant difference in the performance of the different classifiers used in this study. The results might have been affected by not preprocessing data further and not converting the emojis to text. A study by [14], suggest that BERT based

Table 1
Task 1 - Evaluation Results

Run	Accuracy	macro-F1 Score
Run 1 - UBCS	0.311	0.304
Run 2 - UBCS	0.324	**0.319**
Run 3 - UBCS	0.327	**0.319**

models produce better performance as BERT is empirical and a powerful tool. This is also evidenced by the overall performance of teams that used BERT models in the previous studies. Further work needs to be carried out in order to determine whether emojis can significantly improve the classification accuracy when building a classifier to detect the opinions of Twitter users towards vaccine uptake. A further study with more focus on the effects of stopword removal and semantic enrichment of tweets when build classifiers to detect people's opinions is also recommended.

References

[1] Y. Woldemariam, Sentiment analysis in a cross-media analysis framework, in: 2016 IEEE International Conference on Big Data Analysis (ICBDA), 2016, pp. 1–5. doi:10.1109/ICBDA.2016.7509790.

[2] D. Yeung, Social media as a catalyst for policy action and social change for health and well-being: Viewpoint, Journal of Medical Internet Research 20 (2018) e94. doi:10.2196/jmir.8508.

[3] J. K. Catapang, J. V. Cleofas, Topic modeling, clade-assisted sentiment analysis, and vaccine brand reputation analysis of covid-19 vaccine-related facebook comments in the philippines, in: 2022 IEEE 16th International Conference on Semantic Computing (ICSC), 2022, pp. 123–130. doi:10.1109/ICSC52841.2022.00026.

[4] WorldoMeter, Covid live - coronavirus statistics - worldometer 2022, 2022. URL: https://www.worldometers.info/coronavirus.

[5] I. Aygun, B. Kaya, M. Kaya, Aspect based twitter sentiment analysis on vaccination and vaccine types in covid-19 pandemic with deep learning, IEEE Journal of Biomedical and Health Informatics 26 (2022) 2360–2369. doi:10.1109/JBHI.2021.3133103.

[6] A. Baj-Rogowska, Mapping of the covid-19 vaccine uptake determinants from mining twitter data, IEEE Access 9 (2021) 134929–134944. doi:10.1109/ACCESS.2021.3115554.

[7] P. Trapman, Review 2: "Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics", Rapid Reviews COVID-19 (2020). Https://rapidreviewscovid19.mitpress.mit.edu/pub/bat0d0fq.

[8] A. Fontanet, S. Cauchemez, Covid-19 herd immunity: where are we?, Nature Reviews Immunology 20 (2020) 583–584.

[9] H. Ritchie, E. Mathieu, L. Rodés-Guirao, C. Appel, C. Giattino, Coronavirus (covid-19) vaccinations, 2020. URL: https://ourworldindata.org/covid-vaccinations?country=OWID_WRL.

[10] K. Kitchat, K. Limjumroonrat, T. Surasak, Twitter sentiment analysis towards covid-19 vaccines: A case study in new york city, in: 2021 IEEE 2nd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), IEEE, 2021, pp. 1–6.
[11] C. R. Machuca, C. Gallardo, R. M. Toasa, Twitter sentiment analysis on coronavirus: Machine learning approach, in: Journal of Physics: Conference Series, volume 1828, IOP Publishing, 2021, p. 012104.
[12] J. P. D. Delizo, M. B. Abisado, M. I. P. De Los Trinos, Philippine twitter sentiments during covid-19 pandemic using multinomial naïve-bayes, International Journal 9 (2020).
[13] A. Thomson, K. Moon, G. Vallee-Tourangeau, The 5as: A practical taxonomy for the determinants of vaccine uptake, Vaccine 34 (2015). doi:10.1016/j.vaccine.2015.11.065.
[14] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers), Association for Computational Linguistics, 2019, pp. 4171–4186. URL: https://doi.org/10.18653/v1/n19-1423. doi:10.18653/v1/n19-1423.