

## Online Guard: Identifying the misinformation in social media and its impact on COVID-19 vaccination progress in different countries

Sanjida Akter Sharna (Student), Dr. William Eberle (Advisor)  
Department of Computer Science, Tennessee Tech University

### Background and Major Contributions

- The emergence of the novel coronavirus pandemic has caused a myriad of problems worldwide. One such problem is misinformation, which in itself should be considered a risk.
- Since the outbreak of the COVID-19 pandemic, popular social media platforms are flooded by exaggerated phony news which is affecting our society.
- Although the pertinent tools and existing techniques can support fact-checking and identification of conspiracy, misinformation, and negative sentiment at various stages, a complete end-to-end solution is complicated.
- We propose a thorough analysis and identification system named Online Guard using natural language processing tools and supervised learning techniques to identify the relationship between misinformation from the negative sentiment of COVID19 vaccine-related tweets and vaccination progress rate and its impact in different countries for a particular time period.

### Data Collection and Pre-processing for the Experimentations

we have collected two data set from the Kaggle site. One is **COVID-19 all vaccines tweet** dataset and another is **country vaccination** dataset.

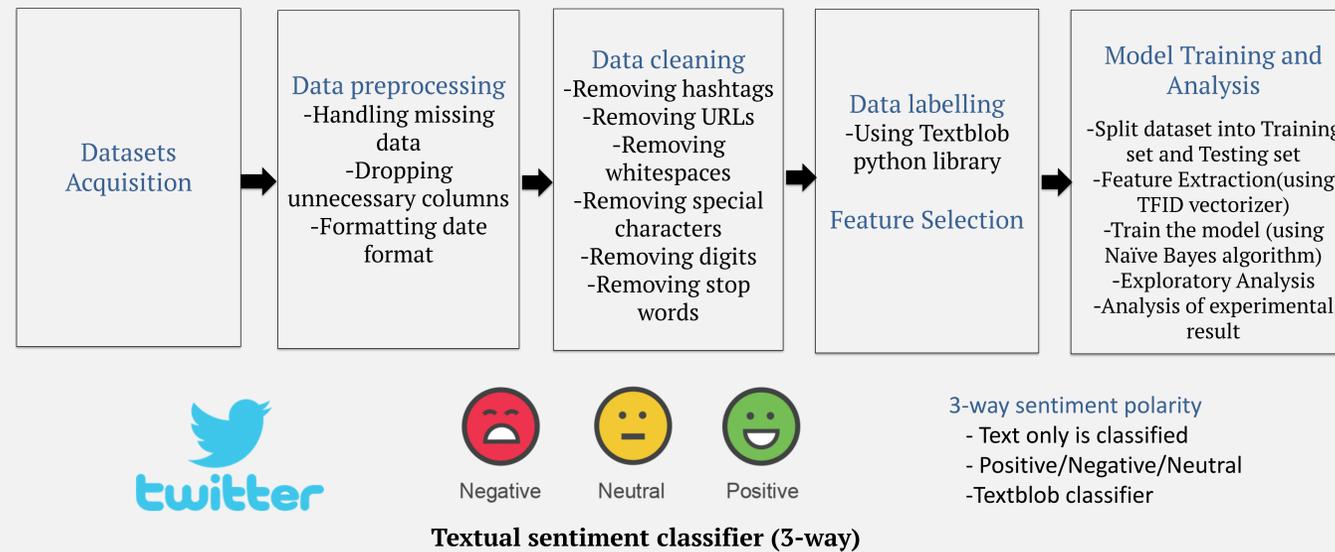
**Table 1:** Attribute description For COVID-19 all vaccines tweet dataset

Attribute Name	Attribute Description
User location	The location the twitter user
text	Text content of the tweet
date	Date of the tweet

**Table 2:** Attribute description For Country Vaccination dataset

Attribute Name	Attribute Description
country	The country for which the vaccination information is provided
date	Date for the data entry
total_vaccinations	The absolute number of immunizations in the country
people_vaccinated	The number of people who receive at least one vaccine shot
people_fully_vaccinated	The number of people received entire set of vaccine immunization scheme

### Understanding Experimental Analysis using Flow diagram



### Experimental Result Analysis

	precision	recall	f1-score	support
Negative	0.87	0.33	0.48	6755
Neutral	0.77	0.98	0.87	40992
Positive	0.90	0.61	0.73	20716
accuracy			0.80	68463
macro avg	0.85	0.64	0.69	68463
weighted avg	0.82	0.80	0.79	68463

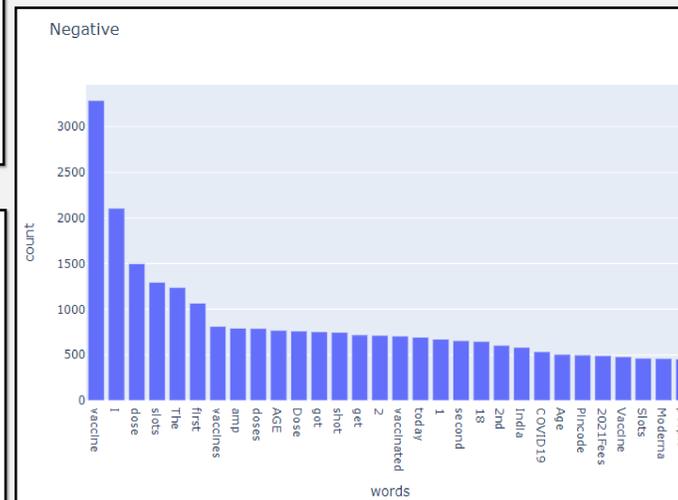
**Figure 1:** Performance of the ML model (Naïve Bayes)

	pred:negative	pred:neutral	pred:positive
actual:negative	2232	3862	661
actual:neutral	63	40253	676
actual:positive	273	7856	12587

**Figure 2:** Confusion Matrix

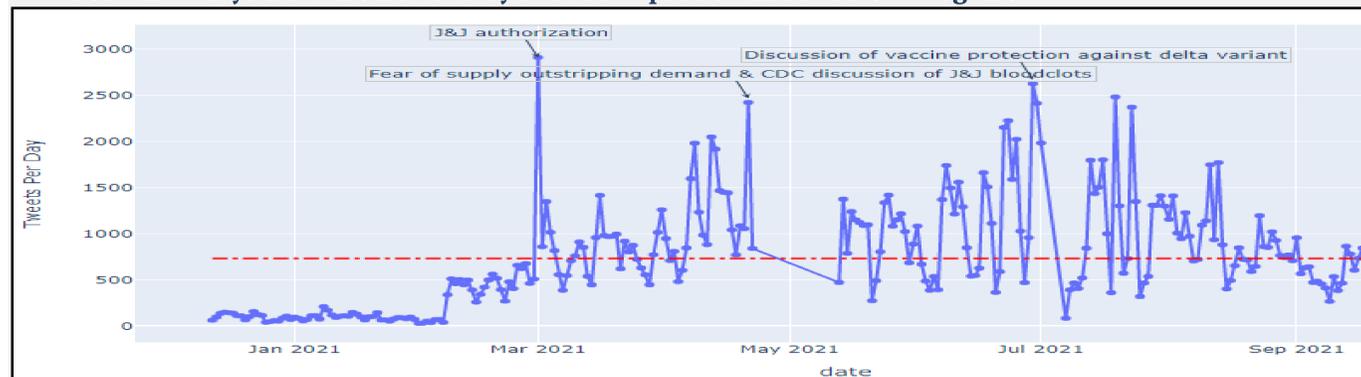
**Training Accuracy: 80.721%**

**Test Accuracy: 80.4405%**



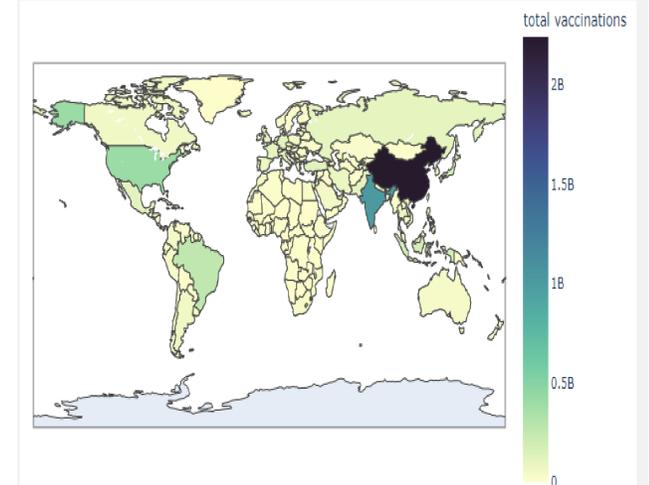
**Figure 3:** Bar chart of negative word count

### Time Series Analysis of Tweets Per Day and its Impact on Vaccination Progress

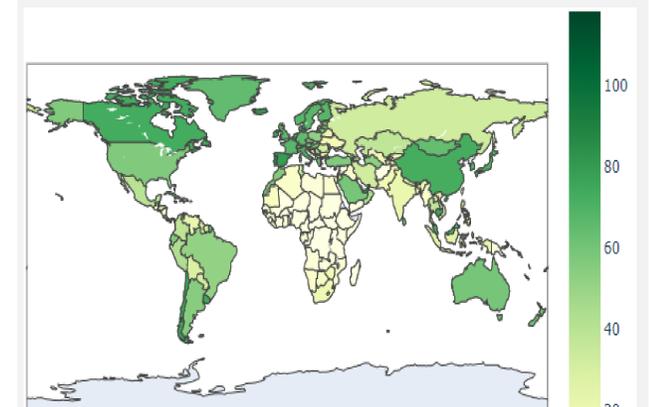


### Exploratory Analysis of Vaccination Progress in Different Country

From the below figure we can see that in most of the countries total vaccination progress is near to zero



**Figure 4:** Total vaccination in different countries (January 2021- September 2021)



**Figure 5:** Fully vaccinated percentage in different countries (January 2021- September 2021)

### Conclusion: Empirical Findings Observations

- Identifying misinformation and negative emotions are important to provide an intimation of terrors lurking inside social media platforms (Twitter) during the pandemic crisis
- Propagation and acceptance of covid vaccine activities are negatively co-related (hypothesis) with social media platforms' increasing misinformation

### Reference

- <https://www.kaggle.com/gpreda/all-covid19-vaccines-tweets>
- <https://www.kaggle.com/gpreda/covid-world-vaccination-progress>