**PAPER DETAILS**

**Paper Title :** Sentiment Analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm.

**Authors** : FMJM Shamrat, Sovon Chakraborty, MM Imran, Jannatul Naeem Muna, Md Masum Billah, Protiva Das, Md. Obaidur Rahman.

**Publish Date :** 2021/7

**Paper Type :** Journal

**Journal Name** : Indonesian Journal of Electrical Engineering and Computer Science [Volume-23, Issue-1 ,Page-436-470]

**Review By** : Md. Shaiful Islam, Department of Computer Science, European University of Bangladesh.

**WHY THEY HAVE CONDUCTED THIS RESEARCH?**

In this paper the authors of this paper have to try the find out the vaccine, which is more preferable for people based

On the twitter tweets about COVID-19 in 2019,2020 and 2021 times. We all are known that the last of the year 2019,

the full year of 2020 and the first some months of 2022 were a very critical time for the people of world. Because

that was the when the people was effecting by COVID-19 much. Almost the entire world into lockdown. For

decreasing the of infection of COVID-19 the scientists of the worlds tried to discover many vaccine. Some were be

succeed. But people afraid that time to be vaccinated. They thought that there has a lot of side effects of taking

these vaccines. But some of them a little number of vaccine worked well. At that time people spent their time at

home much. People loved to share their opinion and emotions on social media. And the twitter is one of the best

common and genuine social media to share opinion. So here the author collect tweet from twitter to analysis which

vaccine is the best according to people’s opinion. And this is the reason that is why the author had conducted this

research.

**PROPOSED SYSTEM**

They proposed a model to find out the best vaccine for COVID-19. They worked to make the base three words

Moderna, Pfizer and AstraZeneca. They collect data from twitter and analyzed these data to find out the best

vaccine. General People however have been expressing their feelings about the safety and effectiveness of the

vaccines on social media like Twitter. In this study, such tweets are being extracted from Twitter using a Twitter API

authentication token.

**ARCHITECTURE/METHODOLOGY**

The authors of this paper actually used Natural Language Processing (NLP) as main methodology of this

research paper. Actually here used different many algorithm of NLP. They collected data from twitter, at

first they collected data as word according to some hashtags i.e. #Pfizer, #Moderna, and #AstraZeneca. But

we all are known that we need data for machine learning project in CSV format. Here they used Tweepy library

converted data word to CSV file. They also used other some NLP algorithm. One of them one is KNN. In this research,

an analysis based on public sentiments about approved covid-19 vaccines based on Twitter data is showed using

natural language processing and supervised KNN classification algorithm. Supervised KNN used here as classifier for

the classification of the polarity data.

They have used RNN for sentiment classification. RNN is a machine learning algorithm which used for classifying

Sentiment . And here used an another machine learning algorithm LDA for sentiment analysis.

And to determine the positive, negative, and neutral sentiment behind the tweets, valence aware dictionary and

sentiment reasoned (VADER) is implemented.

**EXPERIMENTED OUTCOME**

Outcome of a research paper or result of a research paper is the main and very important part of a research paper.

Because we judge a paper by it’s outcome result. For better outcome the better data preprocessing is important.

They proceed visualization of data using world cloud. They dividet it into four part mean,amax,amin,median. And

The polarity calculation of the Pfizer was mean=0.133202, amax=1.0, amin=-1.0, median=0.0. For the moderna

mean=0.114308, amax=1.0, amin=-1.0, median=0.0 and for the AstraZeneca was mean=0.078346, amax=1.0,

amin=-1.0, median=0.0. And the subjectivity calculation was the for three vaccine: For Pfizer was mean=0.34001,

amax=1.0, amin=-1.0, median=0.333333, for the Moderna was mean=0.348156, amax=1.0, amin=-1.0,

median=0.375 and for the AstraZeneca was mean=0.304983, amax=1.0, amin=-1.0, median=0.3 ,

They have visualized also data as plots and scatter plots. And at last they find out tweet the positive , Negative and

Neutral Percentage of Pfizer, Moderna and AstraZeneca . The percentage of Pfizer’s positive, negative and neutral

was 47.2, 37.5 and 15.12 . For the moderna 46.16, 40.71 and 13.13 And for the AstraZeneca 40.08, 40.06 and 13.86.

After that they sentiment analysis and visualized the data as comparison bar. And from here they can see that the

Vaccine Moderna and Pfizer’s tweets were more positive than Negative and Neutral. But AstraZeneca ‘s percentage

of positivity was lower than negativity.