RESEARCH ARTICLE

SAMRIDDHI Volume 14, Special Issue 2, 2022 Print ISSN : 2229-7111 Online ISSN : 2454-5767

COVID-19 Booster Dose Reactions on Twitter - Unveil The Possibility of Acceptance

Harshali Patil[1*], Abhijit Banubakode[2]

[1*,2] MET Institute of Computer Science, Mumbai, India; e-mail : harshalip_ics@met.edu

ABSTRACT

Entire world suffered and faced the outburst of Corona Virus in past two years. The pandemic changed almost everybody's life upside down. The scientist, researchers, medical practitioners, health workers, governing bodies worked day and night to treat and pacify the patients. Invention of vaccine to logistics and distribution was handled globally by different countries and leading authorities. The vaccination helped to control the spread of virus. After two vaccinations when the need of additional dosage (booster) is recommended, then some people come up with their findings and put forth their opinions about the requirement additional dosage where as some were against it. This paper describes in detail the need of opinion mining and sentiments of people about COVID-19 booster dose.

Keywords- Sentiment analysis, COVID-19, booster dosage, NLP, data analysis

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, (2022); DOI : 10.18090/samriddhi.v14spli02.7

INTRODUCTION

Centers for Disease Control and Prevention (CDC) launched agency wide response in January 2020 to the COVID-19 epidemic. The outbreak responses of diseases are witnessed by CDC. The largest responses to any diseases in the history are recorded. To lower the spread of COVID-19, to shield people's lives and health; CDC is actively involved.

In September 2021 CDC recommended COVID-19 vaccine additional dose for elderly and those at high risk in United States. After few months CDC recommended booster shot for all adults. The reactions of people were identified, some scientist said there is absolutely no need of extra dose where as some said extra layer of protection should be availed. Sentiment analysis is a Natural Language Processing (NLP) technique, which is used to mine the data and extract information from it. The outcome of mining determines whether data can be categorised as negative, positive or neutral. Usually the sentiment analysis is performed on textual data, where the mined data can help in business monitoring, growth, brand management, understanding feedback of users, needs and expectations.

Corresponding Author : Harshali Patil, MET Institute of Computer Science, Mumbai, India;
e-mail : harshalip_ics@met.edu

How to cite this article : Patil, H., Banubakode, A., (2022). COVID-19 Booster Dose Reactions on Twitter - Unveil The Possibility of Acceptance.
SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, Volume 14, Special Issue (2), 239-243.

Source of support : Nil

Conflict of interest : None

In recent years, the tremendous growth in number of users of micro blogging site such as Twitter has been witnessed. To get the insights about what people feel and think about the products, most of the companies prefer to mine the Twitter data.

The figure 1 shows the daily share of the population receiving COVID-19 vaccine dose worldwide. Looking at the trends in graph it is identified that the booster dose percentage is less. Figure 2 represents COVID-19 vaccine dose in India, which also leads to similar conclusion. Hence to understand what people think and feel about COVID-19 booster dose become mandatory.

Artificial Intelligence (AI) services and sentiment analysis facilitate data analysts to extract public opinions,market research in depth, brand reputation evaluation, and improve customer experiences. Customer insights are extracted by business using an intermediary sentiment analysis APIs for various platforms into their database management systems. [2]

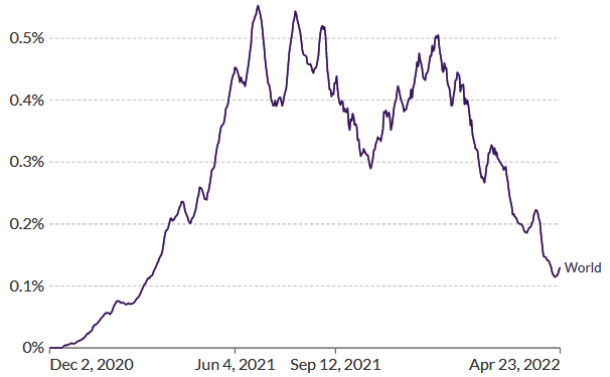

Figure 1: Daily share of (7 days rolling avg.) the population receiving a COVID-19 vaccine dose World-wide [1]

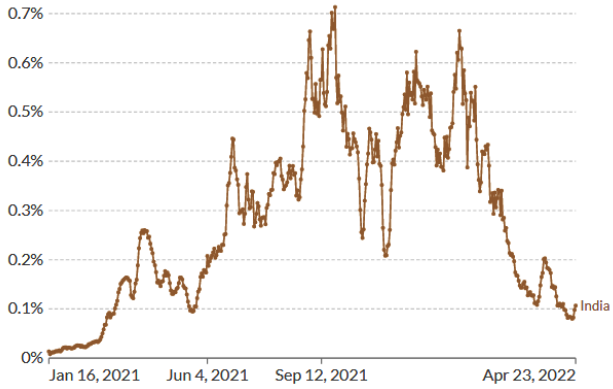

Figure 2: Daily share of (7 days rolling avg.) the population receiving a COVID-19 vaccine dose in India [1]

Related Work

At various levels of granularity, sentiment analysis has been addressed as an NLP task. Document-level classification task (Turney, 2002; Pang and Lee, 2004), sentence-level classification job (Hu and Liu, 2004; Kim and Hovy, 2004), and phrase-level classification task (Wilson et al., 2005; Agarwal et al., 2009) [3,10-15].

The collection of corpora automatically and the linguistic analysis of the corpora collected are explained by (Alexander Pak and Patrick Paroubek, 2010). A sentiment classifier has been built using the corpus. A classifier is developed that can distinguish between positive, negative, and neutral attitudes in a document.

The proposed technique can be used to any language [4].

Twitter opinion retrieval, sentiment tracking across time, irony recognition, emotion detection, and tweet sentiment quantification are just a few of the tasks that have recently gotten a lot of attention. The resources utilised for sentiment analysis and the survey conducted (Giachanou, Crestani, 2017) are discussed [5].

The word level granularity complements Twitter's context because it allows users to exchange brief pieces of information known as "tweets" (limited to 140 characters). According to a review of the literature, the methods for automatically interpreting sentiment at the word level are divided into two categories: There are two types of techniques: (1) dictionary-based and (2) corpus- based approach. Furthermore, several methodologies have been used to anticipate the sentiments of words, expressions, or documents in order to automate sentiment analysis. Natural Language Processing (NLP) and Machine Learning (ML) algorithms are used widely[6]. Multimodal sentiment analyses find out the polarity of sentiment and score for the tweets [7]. The method used to organize tweets into ordinal classes about a topic (Shihab, Yang 2018). The tweets examined were about the 2016 presidential election in the United States. Node.xl was used to collect the tweets. The Valence Aware Dictionary for Sentiment Reasoner (VADER) [8] was used to propose the classification.

SemEval-2016 Task 4 is divided into five subtasks. The first option uses a five-point scale, which gives the classification task an ordinal feel. The second option seeks to accurately estimate the prevalence of each of the classes of interest, a task known as quantification [9].

Data Collection

To identify the sentiments of people on Twitter there is no such dataset of sentiments on COVID-19 booster dose was available. Manual data collection is a prolonged process; hence the required data for the research is collected using API. For the training data, the data is collected using Twitter API.

The reactions of people are identified from tweeter posts. Tweeter API is used to pull out the tweets. To determine the sentiments of people around the globe about the third dose aka booster dose the tweets by people are collected and such recent 1603 tweets have been selected for data analysis. The sample collection of tweets is shown in figure 3.

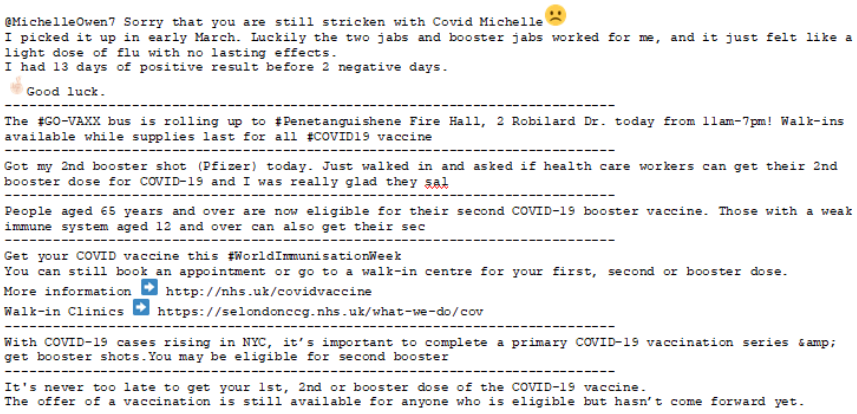

Figure 3: Sample tweets

The outcomes of data collection process give some tweets along with some of their related attributes and store them in a structured format. Collected data contains tweet text and details. The attributes are as follows Tweets: The truncated Tweet text (for Tweets more than 140 characters). To extract full tweet the tweet_mode property should be set to 'extended'.

User: represents the username.

User statuses count: Provide the tweet counts. User followers: represents number of followers.

User location: Users' location information to track the tweet

User verified: the blue colour verification badge Favourite count: the number of tweets that user has marked as favourite

Re-tweet count: users re-tweet count Create-at: represents tweet date.

The libraries used in the research paper for sentiment analysis are tweepy and Txtblob.

tweepy. Cursor(api.search, q=data, count=100, lang='en') indicates that we are searching for data, it's a piece of text. In this case the text searched is "booster dose for covid". The language English indicates that the required tweets are in English language.

FINDINGS

The Textblob library is utilised, and it has a sentiment feature built in. Sentiment property with polarity, subjectivity is returned by a named tuple. The polarity score is a float value that ranges from -1.0 to 1.0. Subjectivity is a float value between 0.0 and 1.0, with 0.0 indicating "extremely objective" and 1.0 indicating "highly subjective. The following Figure 4, shows the snapshot of tweets with sentiment analysis.

Unnamed: 0	Tweets	User	User_statuses_count	user_followers	User_location	User_verified	fav_count	rt_count	tweet_date	clean_tweet	Sentiment
0	RT @cnnphilippines: The government's vaccine e...	Philippine News One	176910	298	Manila City, National Capital Region	False	0	3	2022-04-26 15:07:20	RT The government s vaccine expert panel said ...	Neutral
1	@MichelleOwen7 Sorry that you are still strick...	John Reynolds	1192	29	NaN	False	0	0	2022-04-26 15:01:04	Sorry that you are still stricken with Covid M...	Neutral
2	RT @inquirerdotnet: Taiwan's government has ap...	Philippine News One	176910	298	Manila City, National Capital Region	False	0	4	2022-04-26 14:57:55	RT Taiwan s government has approved a second C...	Neutral
3	RT @SMDhealthunit: The #GO-VAXX bus is rolling...	Immunize Canada	35949	4858	Canada	False	0	1	2022-04-26 14:42:58	RT The GO VAXX bus is rolling up to Penetangui...	Positive
4	RT @BromleyGPs: Get your COVID vaccine this #W...	Bromley CCG	6030	6305	Bromley, London	True	0	1	2022-04-26 14:42:06	RT Get your COVID vaccine this WorldImmunisati...	Neutral
...	...	...	...	...	...	...	...	...	...	...	...
1598	More than 1 million people have now had their ...	Eden District Council	24730	7539	Penrith, Cumbria, UK	False	1	0	2022-04-18 11:01:18	More than 1 million people have now had their ...	Positive

Figure 4: Tweets with sentiment identification

The code snippets for the tweet analysis is shown in Figure 5

```
def analyze_sentiment(tweet):
  analysis = TextBlob(tweet)
  if analysis.sentiment.polarity > 0.1:
    return 'Positive'
  elif analysis.sentiment.polarity < -0.1:
    return 'Negative'
  else:
    return 'Neutral'
```

Figure 5: Sentiment Analyser

The following Figure 6 indicates the histogram of polarity of tweets/sentiments analysis.

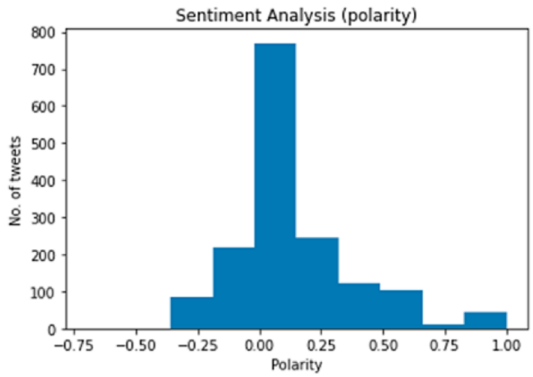

Figure 6: Polarity Histogram

The code snippets in Figure 7 indicate random tweets analysis. The original tweet, cleaned tweet and sentiment analysed are shown.

```
import random
n=random.choice(range(0,df.shape[0]))
print('n=',n)
print('Original tweet:\n'+ df['Tweets'][n])
print()
print('Clean tweet:\n'+df['clean_tweet'][n])
print()
print('Sentiment:\n'+df['Sentiment'][n])
```

```
n= 852
Original tweet:
Fighting Stigma – Delhi offers free booster dose for eligible beneficiaries in govt centres: Delhi will provide fre… https://t.co/GGg7L2l46h

Clean tweet:
Fighting Stigma Delhi offers free booster dose for eligible beneficiaries in govt centres Delhi will provide fre

Sentiment:
Positive
```

Figure 7: Sentiment analysis of random tweet sample

The data analysis of 1603 tweets results in 793 “Neutral”, 690 “Positive” and 120 “Negative” tweets. The following Figure 8 represents sentiments statistics.

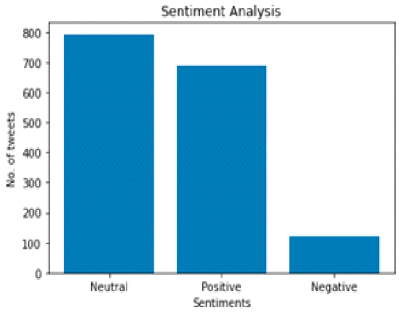

Figure 8: Statistics of Sentiments

The booster dose tweets world cloud is shown in Figure 9.

Figure 9: Booster dose tweets word cloud

Figure 10: Keywords in positive sentiment tweets

Figure 11: Keywords in negative sentiment tweets

After analysing the tweets using vaderSentiment library and considering two sentiments as positive and negative on the basis of compound values (polarity score). The result obtained is as shown in following Figure 13.

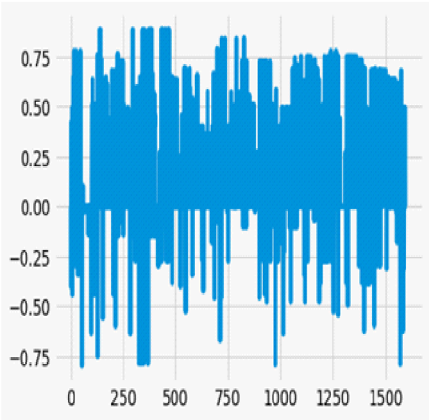

Figure 13: The compound value (polarity score) of tweets

On the basis of polarity score the sentiment classification is done. If the score is more than 0 then the tweet is classified as positive otherwise it's been considered as negative. The sentiment analyser using VADER Sentiment library is shown in following Figure 14.

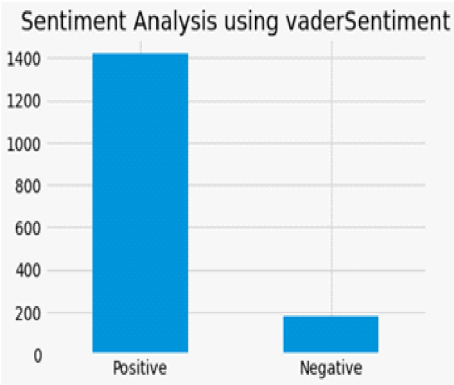

Figure 14: Sentiment Analysis - vadar Sentiment

CONCLUSION

The paper describes the sentiment analysis using Textblob and VADAR Sentiment library. The tweepy library is used for extracting tweets. The sentiment analyser in Textblob provides Negative, Positive and Neutral sentiments for booster dose of COVID-19. Using VADAR Sentiment library the polarity scores are used to find the tweets sentiments. To convert scores into positive and negative sentiments threshold value is been used. The research finding identifies that most of the population is neutral about the booster dose. In general analysis of two methods indicates that people are positive about booster dose.

REFERENCES

[1] Daily share of the population receiving a COVID-19 vaccine dose,7-day rolling average. All doses, including boosters, are counted individually, https://ourworldindata.org

[2] Naveen Mishra, "Sentiment Analysis with Machine Learning: Use Cases and Applications, Dec 2019, https:// artificialintelligence.oodles.io

[3] Apoorv Agarwal ,Boyi Xie, Ilia Vovsha, Owen Rambow, Rebecca Passonneau, "Sentiment Analysis of Twitter Data", Proceedings of the Workshop on Language in Social Media (LSM 2011), pages 30–38, Portland, Oregon, 23 June 2011.

[4] Alexander Pak and Patrick Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining", In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, European Language Resources Association (ELRA), 2010

[5] Anastasia Giachanou, Fabio Crestani, "Like It or Not: A Survey of Twitter Sentiment Analysis Methods", ACM Computing Surveys, Volume 49 Issue 2, June 2017,pp 1–41, https://doi.org/10.1145/2938640

[6] A. Kumar and T. M. Sebastian, "Sentiment Analysis on Twitter", International Journal of Computer Science Issues (IJCSI), 2012

[7] A.Kumar, G. Garg, "Sentiment analysis of multimodal twitter data", Multimed Tools Appl 78, pp 24103–24119, 2019.

[8] Elbagir, S., & Yang, J., Sentiment Analysis of Twitter Data based on Ordinal Classification, Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence.

[9] Nakov, Preslav, Alan Ritter, Sara Rosenthal, Fabrizio Sebastiani and Veselin Stoyanov. "SemEval-2016 Task 4: Sentiment Analysis in Twitter." ArXiv abs/1912.01973 (2016)

[10] Pang, L. Lee, "A sentimental education: Sentiment analysis using subjectivity analysis using subjectivity summarization based on minimum cut", ACL, 2004.

[11] P. Turne, "Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews". ACL, . 2002

[12] M Hu, B Liu, "Mining and summarizing customer reviews", KD, 2004.

[13] S M Kim, E Hovy, "Determining the sentiment of opinion". Coling, 2004

[14] T. Wilson, J. Wiebe, P. Hoffman, "Recognizing contextual polarity in phrase level sentiment analysis", ACL, 2005.

[15] Apoorv Agarwal, Fadi Biadsy, and Kathleen Mckeown, "Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams", Proceedings of the 12th Conference of the European Chapter of the ACL, pages 24–32, March, 2009.