Original Article

Opinions for Receiving COVID-19 Vaccines Based on Sentiment Analysis

Jafar Majidpour, Kamal Al-Barznji
[1,2] Department of Computer Science, University of Raparin, Ranya, Kurdistan Region, Iraq.
Email: jafar.majidpoor@uor.edu.krd[1] , kamal.barznji@uor.edu.krd[2]
DOI: 10.47750/pnr.2022.13.S01.197

Abstract

Due to vaccination hesitancy/refusal, a call to action is urgently needed for the COVID-19 vaccine deployment. Find out what the public thinks about the COVID-19 vaccine. Sentiment analysis is a technique for getting text sentiment scores. Therefore, we proposed architecture to analyze the textual data collection of people's opinions on COVID-19 vaccines using two of the best sentiment analysis techniques, the Bidirectional Encoder Representations from Transformers (BERT) technique and the Valence Aware Dictionary for sEntiment Reasoning (VADER) technique of Natural Language Processing (NLP). A questionnaire survey of corona vaccines recipients who recommend COVID-19 collected the data. Finally, recommendations for the corona vaccine were investigated, and various studies were done to determine its efficacy. Accuracy, precision, recall, and f1-score are standard evaluation criteria. The data shows the proposed model's excellent sentiment analysis performance, indicating that most vaccine users prefer to recommend others to get the vaccines.

Keywords: BERT technique, COVID-19 vaccines, Sentiment analysis, VADER technique.

INTRODUCTION

A worldwide health emergency brought on by the coronavirus disease 2019 (COVID-19) pandemic has led to unending attempts to lower infection rates, prevent fatalities, and develop treatments to lessen its aftereffects. The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is the virus that causes COVID-19, according to the World Health Organization (WHO), and the first case was reported in China in December 2019. The highly contagious illness quickly spread worldwide, and the WHO declared a pandemic in March 2020 [1]-[3]. Globally, the COVID-19 pandemic has had an impact on daily life. Over 260 million individuals have been infected with the virus, and COVID-19 is responsible for over 5 million fatalities, according to a recent assessment by the WHO. All nations now share the shared objective of eradicating COVID-19 and resuming daily operations. Since the pandemic's start, several research teams at significant pharmaceutical corporations and academic institutions throughout the globe have been creating vaccines [4]. WHO has granted emergency use authorizations for eight COVID-19 vaccines, including Sinovac, Sinopharm/BIBP, Moderna, AstraZeneca/Oxford, Johnson & Johnson, Covishield, and Pfizer/BioNTech. In several nations, additional vaccinations are also given under the national regulatory agencies, according to the WHO [5], vaccine hesitancy is one of the top 10 global health issues]. In addition, there has been an increase in conspiracy theories about the efficiency and safety of the COVID-19 vaccine, which may make people less likely to get vaccinated. Therefore, it is essential to keep an eye on the dissemination of false information about the COVID-19 vaccine and to lessen its effects in order to boost public trust and effectively combat vaccination reluctance. A large-scale real-time tracking tool for disease outbreaks and government epidemic control measures was routinely used on social media throughout the worldwide pandemic [6].

The use of machine learning, deep learning, and transformer-based algorithms in sentiment analysis research and work has significantly increased in recent years. For this reason, two different sentiment analysis techniques have been used in this study and compared: first, the Valence Aware Dictionary for Sentiment Reasoning (VADER) method, and second, the Bidirectional Encoder Representations from Transformers (BERT) method of Natural Language Processing (NLP) techniques.

Sentiment analysis, commonly referred to as opinion mining, is a method of natural language processing. It is utilized in various contexts, such as evaluating positive or negative product evaluations, determining if a political party's campaign was successful,

evaluating movie reviews, and evaluating tweets and other social media material. Sentimental analysis is used to understand how individuals genuinely feel about many things, including products, services, organizations, movies, news stories, events, problems, and their characteristics. Sentiment analysis and machine learning are used by all social media monitoring tools and businesses to help them learn more about mentions, brands, and products [7]. However, medical sentiment analysis is a field of study that examines the opinions, sentiments, attitudes, and emotions of patients and doctors toward various clinical contexts (treatment side-effects, medical diagnosis concerns, emotional consequences of illness, emotional context during the onset or evolution of a specific disease, patient-physician relationship, physician attitudes in clinical notes) expressed in a written text [8]. A sentiment analysis tool with a rule-based vocabulary designed exclusively for social media sentiments is called VADER. A sentiment lexicon, or set of lexical characteristics commonly categorized as positive or negative depending on their semantic orientation, is used by VADER. In order to obtain a reliable point estimate of the sentiment valence (intensity) of each lexical attribute, VADER is based on a wisdom of crowds (WotC) technique [9]. With English-language sentiment lexicons, VADER uses gold-standard algorithms. Humans evaluate and approve the lexicons. They use qualitative techniques to enhance the emotion analyzer's effectiveness [2].

BERT is a transformer architectural model that Google Research researchers introduced in 2018 [10]. It may be used to do sentiment analysis, question answering, and natural language processing. BERT makes use of numerous NLP (Natural Language Processing) algorithms and architectures from the past. BERT is a transformer architecture encoding stack. Transformer architecture refers to an encoder-decoder network that uses self-attention on the encoder side and attention on the decoder side. Pre-training and fine-tuning are the first two phases of BERT [11]. BERT receives input in the form of a language format sequence $X = (I_0,..., I_n)$ and produces contextualized vector representations $H = (h_0,..., h_n)$ for each of the sequence's components [12].

The following are the primary contributions (research highlights) we made to this study paper:

1. We collected a questionnaire survey from those who received the corona vaccination and they would recommend it to others.
2. We have evaluated the level of discomfort felt following each vaccination dose.
3. We looked at how consuming vitamins affected those who had received doses of the corona vaccination in pain reduction.
4. We have looked at recommendations for obtaining the corona vaccination based on sentiment analysis techniques.

There are five sections to the work that is discussed in this essay. The related work is in section two. The dataset utilized for experiments and the proposed methods of machine and deep learning is described in section three. The results and discussion are found in section four. Section five includes a summary of the conclusion.

RELATED WORK

Companies employed a variety of approaches to understanding customer feedback and sentiments. In several attempts to monitor social media postings or medical questionnaires queried regarding vaccination and illness incidence, NLP approaches have been employed effectively. Authors of [13] kept an eye on tweets about the virus and saw the beginning of an outbreak. In a report [14], links were shown between effective public health initiatives and favorable attitudes as well as between rising measles cases and unfavorable social media posts about the measles vaccine. More recently, the study [15] employed sentiment analysis using the Nave Bayes algorithm by scanning Twitter data with the phrase "Vaccine COVID-19" to gauge the opinions of the Indonesian people through a social network analysis of the COVID-19 vaccine.

Additionally, many sentiment analysis studies on the COVID-19 vaccination have been carried out. Research [16] utilized articles from Google News and Twitter to gauge the polarity of the media's coverage of the COVID-19 vaccination in Africa. Additional research looked at the drivers of COVID-19 vaccine adoption in China and public opinion analysis of the COVID-19 vaccination in India [17], [18]. The authors of the research [19] examined multiple deep models for sentiment analysis of drug reviews using various embeddings. They used clinical BERT that had already been trained for embedding using LSTM, and the results were subpar. Additionally, a potent Deep Learning Model called BERT Base Uncased is offered in the study effort [20] to clarify the problem of sentiment analysis. In the experimental assessment, the BERT model outperformed the other machine learning techniques with a good prediction and high accuracy.

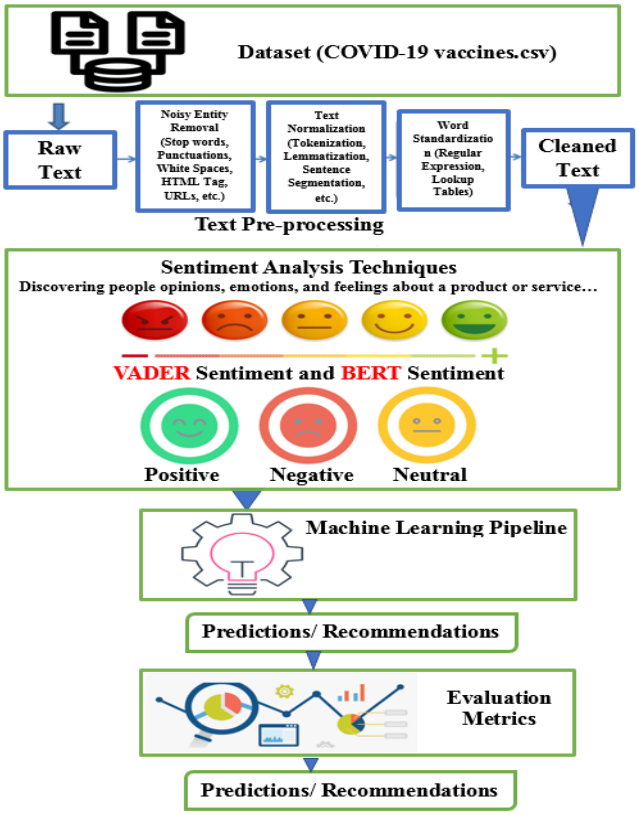
In saying that it is touchier to presume articulations in online media contexts, particularly when supervising web-based media messages and film reviews [21], VADER sets itself apart from competitors. Another benefit of VADER is that it tells you how

positive, neutral, or negative an evaluation is as well as the motivation and cynicism score. The sum of the odds for the positive, neutral, and negative outcomes is one. The compound score is also a very helpful statistic if we need one way to quantify emotion. The majority of threshold values are shown as positive: The most helpful measures for multidimensional [22] are compound score>=0.05 neutral: compound score between -0.05 and 0.05 negative: compound score=-0.05. In addition to assigning a score to each word, VADER also looks at other linguistic and grammatical elements, including punctuation, capitalization, and the use of emoticons [23]. A Twitter dataset has been used using the VADER sentiment analysis tool, as in the paper [24]. The study's objective was to evaluate how well VADER performed when fourteen punctuation marks in the English language were concatenated. Exclamation (!) and question (?) marks were shown to raise Positive and Negative Polarity scores and reduce Neutral Polarity scores, according to the study.

MATERIAL AND METHODS

This study's objective is to determine the percentage of persons who have gotten the corona vaccination and have advised others to do so. The fundamental components of this design are the VADER and BERT methods. The research's proposed architecture is shown in Fig. 1.

Figure 1. Proposed architecture for suggestions of COVID-19 vaccines based on sentiment analysis techniques



All of the steps of the methods proposed are explicitly described in Algorithm I. This work is based on characteristics such as the impact of taking multivitamins before receiving the vaccine on the degree of discomfort and pain reduction or increase, as well as the level of pain following each dose of the vaccine, including those who took multivitamins or did not take them, we investigated; and the effect of these cases in recommending these people to others to receive the vaccine, we have done using data analysis. Our suggested model has three steps: A questionnaire survey was first conducted, after which some of its properties were chosen and examined using state-of-the-art data analysis tools, such as Sentiment Analysis approaches (VADER and BERT techniques).

The remaining four sections of the study are organized as follows: dataset, text pre-processing, sentiment analysis techniques (VADER and BERT techniques), and evaluation metrics.

Algorithm I. Proposed Method:

1	***Inputs**: **D:** Raw Comments Dataset, **comment:** a comment which is selected from the **D**;*
2	*For **comment**=1:**D** // Preprocessing steps*
3	***comment**=Delete stop-words, sparse terms, specific words of **comment***
4	***comment**= Accent marks before lemmatizing the text data of **comment***
5	***comment**=Remove accent marks, punctuation, and diacritics of **comment***
6	*End for*
7	***VM**= Build a **VADER** model which sentiment **comment** based on **VADER** technique*
8	***BM**= Build a **BERT** model which sentiment **comment** based on **BERT** technique*
9	***[accuracy$_{VM}$, precision$_{VM}$, recall$_{VM}$, f1-score$_{VM}$]** = Calculate average metrics [Accuracy, Precision, Recall, F1-score] based on five-fold CVs from **VM** model*
10	***[accuracy$_{BM}$, precision$_{BM}$, recall$_{BM}$, f1-score$_{BM}$]** = Calculate average metrics [Accuracy, Precision, Recall, F1-score] based on five-fold CVs from **BM** model*
11	***AAT$_{VM}$**= Calculate **Average-Time-Test** from **VM** model*
12	***AATBM= Calculate Average-Time-Test from BM model***
13	***End***

A. Dataset

Our dataset for this study is based on a questionnaire survey that was completed at the University of Raparin, and it includes both the design and data gathering processes.

a. Designing the Questionnaire

Department of Computer Sciences and Department of Biology in the University of Raparin, Ranya, Iraq's Kurdistan Region, designed the questionnaire survey (in Kurdish). Four sessions were organized to discuss issues with various occupational and internist experts who commonly interact with COVID-19 patients. Every piece of evidence pertaining to the impact of multivitamins and the degree of suffering experienced by patients following each dosage of the corona vaccination has been

addressed and examined throughout each meeting. These sessions resulted in the formulation of 22 items. The majority of inquiries concern the quantity, kind, and frequency of multivitamin use as well as the degree of discomfort experienced following each dosage of the corona vaccination. All of the questionnaire's questions are listed in Table I.

Table I: All Questions in the Questionnaire

No.	Questions
1	Informational Question (gender, age, marital status, and level of education)
2	Have you ever been infected with the coronavirus?
3	Have you received the coronavirus vaccine?
4	Which of these vaccines have you received? (Type of vaccine)
5	Your vaccine dose:
6	Did you experience pain or symptoms after receiving the vaccine?
7	After which dose of the vaccine did you experience pain or symptoms?
8	Your pain level:
9	The duration of your pain:
10	Which of the following types of pain and illness did you experience after receiving the vaccine?
11	Which of the following treatments have you used since receiving the vaccine?
12	Which of these vitamins have you used so far?
13	How long have you been taking vitamins?
14	After receiving the vaccine, have you followed the guidelines of the Ministry of Health (such as wearing masks and keeping a distance of two meters between people, etc.)?
15	After receiving the vaccine, have you visited any places where people have been infected with the virus and had close contact with them?
16	Did you know about the effect of the vaccine on protecting the body from the virus?
17	How did you get information about the coronavirus vaccine?

18	In general, what is your opinion on the reasons for getting or not getting the vaccine?
19	In general, what is your opinion on the reasons for getting or not getting the vaccine? Would you recommend that others get the vaccine?
20	Do you have confidence in the effectiveness of the vaccine in protecting the body from the virus?
21	What are the reasons for not getting the vaccine?
22	Do you know anyone or anyone who has suffered from pain or other illnesses as a result of receiving the vaccine?

The questionnaire also gathered demographic data, including age and gender, and details regarding the individuals' COVID-19 status.

In a pilot study, the first iteration of the questionnaire survey was presented to 50 participants (27 of whom had received the corona vaccination and 23 had not) who were chosen using intentional sampling [25]. The final questionnaire survey form was developed to be used in the survey after examining the findings, correcting any issues, and taking into account participant opinions on the questions.

b. Data Collection

The sample period ran from December 28, 2021, to February 27, 2022. In addition to printed copies, the questionnaire survey was completed online using Google Forms' free services (https://docs.google.com/forms/d/1pn1AlXWVkBvMh7tpz5uta4goiE3TqDQ6VXSKSe-zclY/edit). Participants were given the online link to complete the survey. The Facebook social media platform was used to disseminate 95% of the surveys. Through social networks, the questionnaire survey was disseminated throughout Iraq to cover the majority of ethnic groups. The link to the poll questionnaire was also promoted on a Messenger channel that broadcasts the daily Hajiawa Post and has over 75,000 subscribers from the general public. Twenty years of age or older Iraqi residents were requested to participate in the poll.

The invitation's language was as follows: "The University of Raparin created this questionnaire survey to find out how much pain, discomfort, and side effects there were after receiving the corona vaccination. We appreciate your cooperation." After reading a page detailing the study, participants had to tick a box stating that they wanted to take part. The option to leave the study at any time was also disclosed to the participants. The questionnaire was filled out by participants once they gave their permission. 725 participants completed the questionnaire throughout the research period. The questionnaire's printed version, which was delivered at random to University of Raparin students, has been answered by an additional 39 participants.

B. Text Pre-processing

An approach for cleaning and preparing text data for usage in a particular environment is called text preprocessing. It is employed by developers in nearly all pipelines for natural language processing (NLP). In our study, we delete stop-words, sparse terms, specific words, and accent marks before lemmatizing the text data. We also remove accent marks, punctuation, and diacritics.

C. Sentiment Analysis Techniques

Our primary goal in this article is to use opinion analysis techniques, which include VADER and BERT techniques. It analyzes the opinions of those who have received the corona vaccine by using data analysis techniques, analyzes how many of the opinions are positive, and suggests others to receive the vaccine. Or how many negative comments are there that do not suggest others receive the vaccine.

a. VADER Sentiment

Using the lexical sentiment classifier VADER sentiment, each email is initially given a sentiment label [26]. A set of words with meaning annotations, typically between -1 and 1, is known as a sentiment lexicon. VADER Sentiment can also translate sentiment ratings for individual words into ratings for whole phrases. When establishing the support for a sentence's mood, negation words and booster words (such as very in extremely joyful) are also taken into account (e.g., not in not happy). We employed the VADER approach to examine user perceptions in our suggested model.

b. BERT Technique

Words that had previously learned the occurrence and representations of words from unannotated training data now have context thanks to pre-trained language models. The BERT is a pre-trained language model that is made to take into account a word's context from both the left and right sides at the same time [10]. Although the idea is straightforward, it improves performance on several NLP tasks, including sentiment analysis and question-and-answer systems. When left and right are trained individually, as in other models like ELMo, BERT can extract more context characteristics from a sequence [27]. There are two BERT configurations that call for several parameters:

1. $BERT_{base}$: The pre-trained model contains a hidden layer with a size of 768, 12 Transformer blocks, 110 M parameters, and 12 self-attention heads.
2. $BERT_{large}$: The pre-trained model includes 340M parameters, 24 Transformer blocks, 16 self-attention heads, and a hidden layer size of 1024. It also has 24 Transformer blocks and 24 Transformer blocks.

We improved our pre-trained $BERT_{base}$ model using the sentiment analysis data we collected.

D. Evaluation Metrics

Accuracy is a straightforward model evaluation metric, and the results are compared using well-known metrics like F1-Measure, Precision, Recall, and Accuracy. The proportion of labels that a model correctly predicts is utilized as an assessment metric for the various methods. A model's accuracy is 85 percent, for instance, if it correctly predicts the labels for 85 out of 100 observations in a test dataset [7]. Recall (R) is the percentage of related things that are also included in the set of suggested items, whereas precision (P) is the percentage of recommended items that are linked to the user. The F1-measure aids in combining precision (P) and recall (R) into a single metric [28]. The metrics are specified as follows, and Table III presents the results.

$$Precision(P) = \frac{Correctly\ Recommended\ Items}{Total\ recommended\ Items} \quad (1)$$

$$Recall(R) = \frac{Correctly\ Recommended\ Items}{Total\ Useful\ Recommended\ Items} \quad (2)$$

$$F1 - Measure = \frac{2*P*R}{P+R} \quad (3)$$

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (4)$$

There are four possible outcomes in accuracy: True Positive (TP) denotes that the model's result prediction and the actual observation both came true. When a model predicts a correct result but the actual observation is erroneous, this is known as a false positive (FP). False Negative (FN) shows that although the observation was accurate, the model projected a false result.

The True Negative (TN), which shows that the model predicted a false result while the actual result was likewise wrong, is the last.

RESULT AND DISCUSSION

The framework of the suggested model is depicted in Fig.1 of this article, which suggests a way to assess users' opinions of those who have received the corona vaccination and advise others to do the same.

Preprocessing, training, and testing are all included in this section, after which the outcomes are assessed using both models (VADER and BERT techniques). The following parts go into further information on it.

A. Preprocessing

A CSV file format is utilized to transfer all of the data obtained from the questionnaire survey into the pre-processing phases and the training of models. In the context of the data. Our model may now be trained using the normalized data.

B. Train and Test phases

At this stage, the models are given all the information collected from the individuals who got the corona vaccination, which includes data from all 725 survey participants, of whom 504 were administered the corona vaccines as part of the training process. The characteristics of the questionnaire survey questions that were used as input to the $BERT_{base}$ and VADER models are shown in Table II, which includes the following: having pain and discomfort after the injection of the Corona vaccine, the duration of pain and discomfort, the level of pain, the location of the pain, the use of antibiotics to reduce pain, and the type of antibiotics.

Table II: Features used as input for the $BERT_{base}$ and VADER models

Questions	Attribute
Did you experience pain or symptoms after receiving the vaccine?	☐ yes ☐ no
After which dose of the vaccine did you experience pain or symptoms?	☐ first dose ☐ second dose ☐ both dose
Your pain level:	☐ low ☐ high ☐ medium
The duration of your pain:	☐ less than a day ☐ one to two days

	☐ three to five days ☐ ☐ a week more than a week
Which of the following types of pain and illness did you experience after receiving the vaccine?	☐ I didn't feel any pain ☐ I felt a little weak ☐ ☐ Chills and fever Cough ☐ Shortness of breath ☐ ☐ Sore throat Stomach pain ☐ Pain in the legs and arms ☐ Shoulder and neck pain
Which of the following treatments have you used since receiving the vaccine?	☐ Nothing ☐ Panadol ☐ Acetaminophen ☐ ☐ Ibuprofen Other types
Which of these vitamins have you used so far?	☐ ☐ Vitamin B Vitamin C ☐ ☐ Vitamin D Multivitamin ☐ I haven't taken any vitamins
How long have you been taking vitamins?	☐ ☐ 1 month 3 months ☐ ☐ 6 months 1 year ☐ more than 1 years

In this study, we used two scenarios for sentiment analysis; in the first scenario, we used the VADER technique to analyze user comments. VADER is a lexicon and rule-based sentiment analysis tool, as was already described. It combines a sentiment lexicon, a collection of lexical elements that are often classified as either positive or negative depending on their semantic orientation. The VADER approach was chosen for this study for a number of reasons, including its relative effectiveness in handling social media texts, movie reviews, and product reviews. This is so that VADER may convey both the positivity and negativity score as well as the degree of a sentiment's positivity or negativity [29]. VADER has many advantages, which are listed as follows:

1. It's excellent for posting on social media because it's easy to type.
2. No training data is needed since it is built from a human-curated gold standard vocabulary that is generalizable, valence-based, and based on human judgments.
3. VADER uses emoji to classify sentiment.
4. When used online, it's quick enough.
5. It doesn't have a significant speed-performance tradeoff, which is good news.

The following lists the standard VADER thresholds:

1. Negative Sentiment: a score of 0.05 or more
2. Neutral Sentiment: a composite score of -0.05 and 0.05 indicates that
3. Negative Sentiment: you must have a composite score of less than -0.05

In the second scenario, the BERT technique was used to analyze user comments. This pre-trained $BERT_{base}$ model was adjusted using our sentiment analysis dataset. Some characteristics set $BERT_{base}$ apart from other language representations. Despite the fact that words might have several meanings depending on the context, context-free word-embedding methods like Word2Vec and GloVe only provide a single representation for each word in the lexicon [30]. By producing word representations depending on their context, $BERT_{base}$ offers contextualized representation.

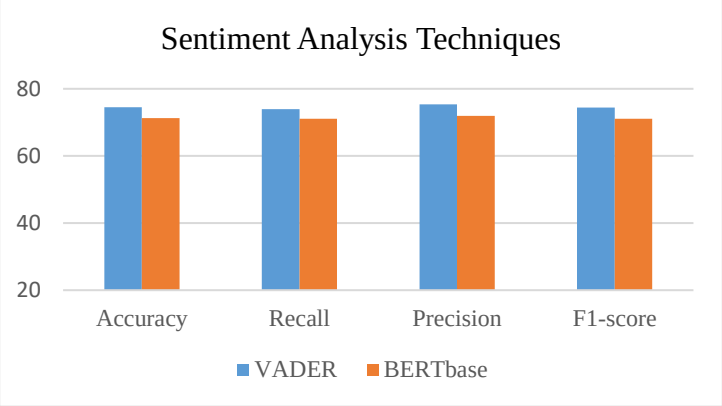
Test purposes include five-fold cross-validation (CV), which means we trained and tested our system five times with different data, utilizing 80% of the data for training and just 20% for testing, resulting in 403 and 101 user comments, respectively, for VADER and $BERT_{base}$ models training and testing. Once all five-fold CVs had been determined, we averaged them. Table III, and Fig.2 shown the results of all sentiment analyses.

Table III: Results obtained using our proposed model for the VADER and $BERT_{base}$ models.

Sentiment Analysis Techniques	Accuracy	Recall	Precision	F1-score	ATT
VADER	74.52	73.93	75.32	74.46	0.046
$BERT_{base}$	71.27	71.04	71.89	71.12	0.039

ATT=Average-Time-Test

Figure 2. The obtained results for both techniques VADER and $BERT_{base}$



C. Results

Using VADER and $BERT_{base}$ models, we analyze the opinions of the participants in the questionnaire survey and then examine the correctness of the models based on their declared measurements in this section.

According to the data, both models' performance accuracy for analyzing the opinions of questionnaire survey respondents based on five-fold CVs is shown in Table III. The VADER analysis was better than the $BERT_{base}$ model, which was able to correctly analyze the comments with an analysis accuracy of 74.5, so the results can be clearly seen in Fig. 2. The low number of comments is one of the factors contributing to $BERT_{base}$ model's poor performance in comment analysis; as was already mentioned, $BERT_{base}$ model relies on deep learning techniques, and generally speaking, accuracy increases when the data is adequate. The Average-Time-Test for model testing is short for both models, as shown in Table 3, and the $BERT_{base}$

model uses less time than the VADER model. The VADER and $BERT_{base}$ models were implemented using the Python programming language. An NVIDIA GeForce GTX 1080 Ti GPU was used for all training procedures.

CONCLUSION

The article's principal purpose is to utilize opinion analysis techniques (VADER and $BERT_{base}$) to study the opinions of persons who have gotten the corona vaccination using data analysis techniques for opinions that are positive and suggest others to obtain the vaccine and for the negative opinions, we studied and were able to determine that most individuals who received the corona vaccination wish to promote the vaccine to others based on their pain and discomfort thresholds and the methodologies of opinion analysis.

Additionally, we were able to examine the effect of multivitamins on people's decisions to recommend vaccinations to others using both BERT and VADER techniques. In the future, we'll employ creative methods for examining customer feedback on datasets with adequate data.

REFERENCES

1. C. A. Melton, O. A. Olusanya, N. Ammar, and A. Shaban-Nejad, "Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence," J. Infect. Public Health, vol. 14, no. 10, pp. 1505–1512, 2021.
2. A. A. Reshi et al., "COVID-19 Vaccination-Related Sentiments Analysis: A Case Study Using Worldwide Twitter Dataset," Healthc., vol. 10, no. 3, pp. 1–28, 2022.
3. S. Ghosh and L. C. Das, "Using Data Mining Techniques for COVID-19 : A Systematic Review," Int. J. Data Sci. Technol., vol. 8, no. 2, pp. 36–42, 2022.
4. K. T. Shahriar, M. N. Islam, M. M. Anwar, and I. H. Sarker, "COVID-19 analytics: Towards the effect of vaccine brands through analyzing public sentiment of tweets," Informatics Med. Unlocked, vol. 31, no. 100969, pp. 1–9, 2022.
5. S. Yousefinaghani, R. Dara, S. Mubareka, A. Papadopoulos, and S. Sharif, "An analysis of COVID-19 vaccine sentiments and opinions on Twitter," Int. J. Infect. Dis., vol. 108, pp. 256–262, 2021.
6. A. R. Rahmanti et al., "Social media sentiment analysis to monitor the performance of vaccination coverage during the early phase of the national COVID-19 vaccine rollout," Comput. Methods Programs Biomed., vol. 221, no. 106838, pp. 1–7, 2022.
7. K. Al-Barznji and A. Atanassov, "Big Data Sentiment Analysis Using Machine Learning Algorithms," in Proceedings of 26th International Symposium "Control of Energy, Industrial and Ecological Systems", Bankia, Bulgaria., 2018, pp. 53–58.
8. A. Brezulianu, A. Burlacu, I. V. Popa, M. Arif, and O. Geman, "'Not by Our Feeling, But by Other's Seeing': Sentiment Analysis Technique in Cardiology—An Exploratory Review," Front. Public Heal., vol. 10, no. April, pp. 1–9, 2022.
9. M. Chiny, M. Chihab, Y. Chihab, and O. Bencharef, "LSTM, VADER and TF-IDF based Hybrid Sentiment Analysis Model," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 7, pp. 265–275, 2021.
10. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 2019, vol. 1, no. Mlm, pp. 4171–4186.
11. P. Sudhir and V. D. Suresh, "Comparative study of various approaches, applications and classifiers for sentiment analysis," Glob. Transitions Proc., vol. 2, no. 2, pp. 205–211, 2021.
12. S. Tabinda Kokab, S. Asghar, and S. Naz, "Transformer-based deep learning models for the sentiment analysis of social media data," Array, vol. 14, no. February, p. 100157, 2022.
13. A. Alessa and M. Faezipour, "Preliminary flu outbreak prediction using twitter posts classification and linear regression with historical centers for disease control and prevention reports: Prediction framework study," JMIR Public Heal. Surveill., vol. 5, no. 2, pp. 1–18, 2019.
14. V. Raghupathi, J. Ren, and W. Raghupathi, "Studying public perception about vaccination: A sentiment analysis of tweets," Int. J. Environ. Res. Public Health, vol. 17, no. 10, 2020.
15. Pristiyono, M. Ritonga, M. A. Al Ihsan, A. Anjar, and F. H. Rambe, "Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm," IOP Conf. Ser. Mater. Sci. Eng., vol. 1088, no. 1, p. 012045, 2021.
16. S. Gbashi, O. A. Adebo, W. Doorsamy, and P. B. Njobeh, "Systematic delineation of media polarity on COVID-19 vaccines in Africa: Computational linguistic modeling study," JMIR Med. Informatics, vol. 9, no. 3, 2021.
17. P. Sv, R. Ittamalla, and G. Deepak, "Analyzing the attitude of Indian citizens towards COVID-19 vaccine e A text analytics study," Diabetes Metab. Syndr. Clin. Res. Rev., vol. 15, pp. 595–599, 2021.
18. F. Yin, Z. Wu, X. Xia, M. Ji, Y. Wang, and Z. Hu, "Unfolding the Determinants of COVID-19 Vaccine Acceptance in China," J. Med. Internet Res., vol. 23, no. 1, 2021.
19. C. Colón-Ruiz and I. Segura-Bedmar, "Comparing deep learning architectures for sentiment analysis on drug reviews," J. Biomed. Inform., vol. 110, no. August, p. 103539, 2020.
20. M. P. Geetha and D. Karthika Renuka, "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model," Int. J. Intell. Networks, vol. 2, no. June, pp. 64–69, 2021.
21. S. Sagar, "Twitter Sentiment Analysis Using Vader," IJARIIT (Volume 4), vol. 4, no. 1, pp. 485–489, 2018.

22. V. Bonta, N. Kumaresh, and N. Janardhan, “A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis,” Asian J. Comput. Sci. Technol., vol. 8, no. S2, pp. 1–6, 2019.
23. A. Alsaeedi and M. Z. Khan, “A study on sentiment analysis techniques of Twitter data,” Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 2, pp. 361–374, 2019.
24. A. Oad, I. H. Koondhar, P. K. Butt, A. Oad, M. Ahmed, and S. R. Bhutto, “VADER Sentiment Analysis without and with English Punctuation Marks,” Int. J. Adv. Trends Comput. Sci. Eng., vol. 10, no. 2, pp. 1483–1488, 2021.
25. A. B. and S. Ebrahim, "Handbook of Health research methods: Investigation, measurement and analysis," McGraw-Hill Education (UK), 2005.
26. E. Hutto, C.J. and Gilbert, “VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text,” Eighth Int. AAAI Conf. Weblogs Soc. Media, vol. 8, no. 1, pp. 216–225, 2014.
27. M. E. Peters et al., “Deep contextualized word representations,” in Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 2227–2237, 2018.
28. K. Al-Barznji and A. Atanassov, “Collaborative Filtering Techniques for Generating Recommendations on Big Data,” in International Conference Automatics and Informatics, CAN, Sofia, Bulgaria, 2017, pp. 225–228.
29. Bonta, V., & Janardhan, N. K. N., "A comprehensive study on lexicon based approaches for sentiment analysis," Asian Journal of Computer Science and Technology, 8(S2), 1-6, 2019.
30. Biswas, E., Karabulut, M. E., Pollock, L., & Vijay-Shanker, K., "Achieving reliable sentiment analysis in the software engineering domain using bert," International Conference on Software Maintenance and Evolution (ICSME), 2020 IEEE, pp. 162-173.