Comparing Deep and Machine Learning Models for Sentiment and Emotion Classification from Vaccine *#sideffects*

1st Aditya Dubey
Dept of CSE
Univ. of Connecticut
Storrs, CT 06269, USA
aditya.dubey@uconn.edu

2nd Swapna S. Gokhale
Dept. of CSE
Univ. of Connecticut
Storrs, CT 06269, USA
swapna.gokhale@uconn.edu

***Abstract*—The accelerated development of Covid-19 vaccines offered tremendous promise and hope, yet stirred significant trepidation and fear. These conflicting emotions motivated many to turn to social media to share their experiences and side effects during the process of getting vaccinated. This paper analyzes sentiment and emotions from tweets collected using the hashtag *#sideffects* during the early roll out of the Covid-19 vaccine. Each tweet was labeled according to its sentiment polarity (positive vs. negative), and was assigned one of four emotion labels (joy, gratitude, apprehension, and sadness). Exploratory analysis of the tweets through word cloud visualizations revealed that the negativity of emotions intensified with the severity of side effects. Word and numerical features extracted from the text of the tweets and metadata were used to train conventional machine learning and deep learning models. These models resulted in an accuracy of 81% for binary sentiment classification, and 71% for multi-label emotion identification. The proposed framework, which yielded competitive performance, may be employed to gain insights into people's thoughts and feelings from vaccine-related conversations. These insights can be helpful in devising communication and education strategies to mitigate vaccine hesitancy.**

***Index Terms*—Covid-19 vaccine, Side effects, Sentiment, Emotions, Vaccine hesitancy.**

I. INTRODUCTION & MOTIVATION

Social media platforms such as Twitter and Facebook make it easier for people to communicate with each other, and hence, facilitate the creation of interpersonal relationships. Sentiment and emotions play an important role in these conversations, because they not only convey what people feel but also provide a valuable social function by influencing others' emotions and actions forming the basis of strong social bonds. Mining social media feeds provides new opportunities to understand individuals' opinions and thoughts around specific events and even generally. Moreover, because social media users may sense and talk about sentiment and emotions of others who are not on the platform, mining these data can also provide insights into the collective feelings of a community or a society.

The disruptions caused by the Covid-19 pandemic brought out a range of intense emotions from anxiety, fear, panic, concern, and anger [17]. Mining emotions during such challenging times can gauge the subjective well-being of the citizenry [6], and can provide insights into people's concerns about emergency measures and policies [6], [27]. An issue which offered a global sense of relief but also generated intense debate was Covid-19 vaccines. People marveled at the accelerated pace at which these vaccines were developed, however, they also expressed concern and apprehension that the side effects of the vaccine were unknown. This conflict motivated many to turn to social media platforms to share their side effects [11]. Some wanted to document their real-life experiences to reassure others while others expressed frustration about the inconvenience they faced. These feeds not only listed the various after effects but were also loaded with a rich set of emotions [18] ranging from from excitement and joy to nuisance and sadness. Mining emotions from these conversations on side effects can provide a unique glimpse into public perceptions and reactions during the process of getting vaccinated. For example, negative emotions may identify anxieties, worries, inconveniences, and disruptions that may lead to vaccine hesitancy [9]. It may also predict the onset of depressive symptoms or psychological disorders following vaccination [13], [23], [26], or determine who is vulnerable to severe side effects, and even whether someone will need a booster shot [24]. On the other hand, positive emotions may indicate an acceptance of the vaccine, and identify thoughts and arguments that resonate well with the public.

The objective of this paper is to analyze sentiment and emotions from vaccine conversations on Twitter, collected in the context of individual personal experiences just before or after receiving the Covid-19 vaccine. Tweets collected using the hashtag *#sideeffect* during the first phase of the vaccine roll out were annotated according to both sentiment/polarity (positive and negative) and four emotions (joy, gratitude, apprehension, or sadness). Tweets with emotion labels were further explored

IEEE/ACM ASONAM 2022, November 10-13, 2022
978-1-6654-5661-6/22/$31.00 © 2022 IEEE

through word cloud visualizations which revealed that the negativity of the emotions, ranging from joy to gratitude to apprehension to sadness increases with the severity of side effects. Word and numerical features extracted from the tweet data were used to train conventional machine learning and deep learning models. Our framework achieved an accuracy of 81% on binary sentiment classification and 71% on multi-label emotion identification. Conventional machine learning models, in particular Naive Bayes, outperform the sophisticated, computationally intensive deep learning models. The proposed framework can thus be used to identify both positive and negative emotions from vaccine-related conversations, and opens opportunities to use positive emotions to develop effective communication strategies to build reassurance and trust aimed at combating vaccine hesitancy.

The rest of the paper is organized as follows: Section II compares and contrasts related work. Section III reviews the steps in data preparation. Section IV discusses the interplay between emotions and severity of side effects. Section V presents the computation of features. Section VI summarizes the parameters of conventional and deep learning models. Section VII defines the performance metrics. Section VIII discusses the results. Section IX offers concluding remarks and directions for future research.

II. Related Research

During the Covid-19 crisis, sentiment and emotion analysis of social media content on topics and issues related to the pandemic has been an emerging topic of research. Kumar *et. al.* [12] propose a novel bidirectional LSTM to classify content related to Covid-19 into the emotions of sad, angry, happy and depressed. Imran *et. al.* [14] analyze the reactions of citizens from different cultures to coronavirus using deep LSTM models trained on sentiment140 data set to estimate the sentiment polarity and emotions. A geographical analysis of emotions is presented by Wang *et. al.* [28] and Chun *et. al.* [8]. Jelodar *et. al.* [15] use a LSTM recurrent neural network to classify Reddit comments from a Covid-19 group into positive and negative sentiment achieving an accuracy of 80%.

While the above work pertains to the pandemic in general, the role of mining social media feeds in understanding public attitudes, sentiment and emotions regarding Covid-19 vaccines is also recognized [5]. Nabiul *et. al.* unravel and compare the sentiment behind each vaccine candidate (Pfizer, Moderna, etc.) [1], [29]. Bustos *et. al.* [4] compare emotions before and after the vaccine roll out. Monselise *et. al.* [20] analyze the sentiment and emotion in each topic related to the vaccine including access and equity. Chandrasekaran *et. al.* [7] examine weekly patterns in the sentiment related to the vaccine from the beginning of the pandemic to early roll out. Some researchers also investigate geographical dispersion and trends in sentiment, and how they are shaped by regional cultures [2], [25], while some others examine both temporal and geographical trends [30].

Contemporary works, however, suffer from several shortcomings. Most focus on sentiment or polarity analysis, rather than the identification of specific emotions. Sentiment polarity certainly conveys useful information, but emotion detection offers richer insights. It provides reasons and rationale into why the sentiment about a particular subject is negative (or positive). For example, being negative with fear or apprehension about the vaccine is a completely different state of mind than being negative with sadness over severe side effects. Second, nearly all use automated tools such as VADER and TextBlob. While these serve as good baseline models, many characteristics that are unique to social media feeds such as misspelled words, slang, emoticons, hashtags and abbreviations, and complex linguistic constructs including irony and sarcasm can degrade their accuracy. Hence, the sentiment and emotions computed by these tools must be validated against manual annotation, but that is almost never attempted. Finally, most of these approaches study people's perceptions in the early phase of the pandemic, when vaccines were totally speculative. Although people may feel strongly even when the vaccine was hypothetical, it is expected that the most intense and difficult emotions would be displayed during a period of vulnerability just before or after having received the vaccine.

Our approach seeks to remedy many of the above shortcomings. It relies on manual annotation to identify complex emotions as human beings would interpret them, and on the tweets that are collected during the process of vaccination when people are actually dealing with the side effects. We also explore how people's emotions are related to the severity of the side effects, providing yet another perspective on people's attitudes and perceptions about the vaccine. The performance of the classification framework is consistent with the works in the literature. Thus, the proposed research can be employed to study vaccine-related conversations to improve our shared understanding of how people's experiences may shape their willingness to take additional boosters.

III. Data Preparation

Tweets were collected in late March in the second phase of the vaccine roll out. Hashtag *#sideeffect* was input to rtweet [16], which returned about 4000 tweets. These tweets were filtered to build a corpus of those describing personal, subjective pre- and post-vaccination experiences. Each tweet in the corpus was manually annotated according to its sentiment/polarity and emotion independently as shown in Figure 1. Two sentiments (positive, negative), and four emotions (joy, gratitude, apprehension and sadness) were considered. Descriptions of these four emotions are given below, and two paraphrased examples of tweets are in Table I:

- **Joy:** These tweets exude joy, excitement, and optimism. Generally, these tweets are light-hearted, either expressing happiness directly or indirectly using a heavy sense of humor and occasionally even offensive, sarcastic or ironical language.
- **Gratitude:** These tweets express gratitude and relief for many reasons and in several forms. Many are thankful for the opportunity to get vaccinated and to the scientists/science, politicians, and frontline workers for making

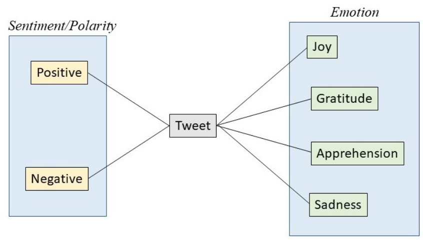

Figure 1. Tweet Annotation – Sentiment & Emotions

the experience painless, efficient and smooth. It also includes users who think that their side effects are milder than what they expected, either expressed explicitly or by comparing to other common vaccines such as shingles and the flu. Some compare side effects of the vaccine with Covid-19 itself and indicate that they would take the vaccine over again.

- **Apprehension:** These tweets express anticipation, apprehension or worry, arising from the lack of clear expectations, and hence, inability to cope with the threat of severe side effects [10]. Users are at different stages – prior to the first dose, in between the first and the second doses, or right after receiving a first/second dose. Anxiety is causing them to actively seek the experiences of their friends and family to prepare themselves mentally. Identifying such users is important because around 20% of the people express skepticism about the vaccine, and even 8-10% who took the first dose did not return for the second one. Reassuring such people that their (potential) side effects are expected and normal is necessary. At the other end, some worry whether their vaccine is working because their side effects are mild to none, and they need to be reassured as well. Few users express apprehension about the lack of availability of the vaccine, appointments, or any other logistical details.
- **Sadness:** This set of tweets express sadness, anger or frustration. Loss of contact and daily routines arising from the side effects are usually major factors in sadness. Here, side effects are never discussed with any of the positive emotions of either joy, humor, or gratitude. Some users also appear irritated about experiencing more severe side effects in comparison with friends and family members.

Each tweet was labeled by two annotators. Only those tweets where both the sentiment and emotion labels from the two annotators matched were included in the final corpus of 1792 tweets. Of these 795 (997) were labeled negative (positive). The split into different emotions is: apprehension – 350, sadness – 445, joy – 359 and gratitude – 638.

IV. Emotions & Side Effects

People's fears of the side effects often influenced their decision on whether they should take the vaccine. Against this backdrop, we explored the relationship between emotions and after effects by building word cloud visualizations from the tweets labeled as joy, gratitude, apprehension and sadness. We pre-processed the tweets to eliminate noise and other extraneous information prior to building word clouds. We tokenized the tweets and then transformed the tokens to lower case. We then normalized the tokens, removed stopwords and those containing Unicode characters. Stopwords included those imported from the nltk.corpus library as well as domain-specific words such as vaccine, side, effects, moderna, pfizer, etc. which appeared in many of the tweets. We then stemmed the words to their root form. Finally, we lemmatized the tweets, which refers to grouping together the inflected forms of a word so that they can be analyzed as a single item or lemma. All the pre-processing steps were completed using the NLTK library.

The word clouds are shown in Figures 2 through 5. Users in the joy word cloud appear to have the least side effects, as seen by the words *nothing* and *zero* in Figure 2. They use a heavy dose (pun intended) of sarcasm and humor, by bringing in *Microsoft* and *Bill Gates* alluding to Gates' role in the Covid-19 vaccine. Words such as *excited*, *happy*, *good*, *love* and *great* display positive feelings of optimism and hope, and whole-hearted acceptance. Words such as *super powers* and *microchip* may be directed at anti-vaxxerrs, who circulated a conspiracy that the government was trying to implant a microchip into our brains through the vaccine. A little sore arm may be the only side effect in this group.

Figure 2. Word Cloud – Joy

The gratitude word cloud in Figure 3 shows a greater variety of side effects, but all mostly still mild. Sore arm still seems to be common, but some instances of headache, fever, fatigue and soreness at the injection site can be seen. Words such as *happy* and *good* show subdued optimism, devoid of exuberance, joy, sarcasm, and humor found in previous word cloud. This group is more grateful rather than joyous, may be because their side effects are not as trivial and inconsequential as compared to those expressing joy. Their words of gratitude may express positive emotions as a way to reassure others, rather than share what they are actually feeling.

Table I
EXAMPLES – SENTIMENT/POLARITY & EMOTIONS

Sentiment/Polarity	
Sentiment	Example
Positive	*So very thankful to get my J&J vaccine! Except for a minor headache a few hrs later, I've had no side effects at all.*
	6hrs post 1st COVID vaccine and no side effects for me! So glad to be on my way to helping the heard immunity
Negative	*My wife got her second vaccination shot yesterday. She's pretty wiped out today and will have to take a few days off from work. Seems that women are experiencing more side effects to the vaccine than men.*
	oof i think the side effects of my covid vaccine are kicking in because i've been exhausted for the past hour and a half, even after a nap
Emotion	
Emotion	Example
Apprehension	*Is my COVID-19 vaccine working if I don't have side effects?*
	If you've had the two dose vaccine, did you have side effects from the second vaccine, but not the first? I've heard a lot of folks say vaccine 2 is a real girly dog next day. I had zero side effects from vaccine #1, but I'm feeling angsty about having my second next week.
Sadness	*Yeah forget what I said earlier about me not experiencing any side effects from the vaccine. I feel sluggish and have extreme joint pains. Send help*
	Just my luck to get hit with the side effects of the covid vaccine lol
Gratitude	*I just got my ONE AND DONE Johnson & Johnson vaccine and 6 hrs later, I can report no side effects. Not even soreness at the injection sight. Thank you President Biden!*
	Everybody hyped up these vaccine side effects because I don't feel sore or sick at all ... but I mean knock on wood still have to get the second shot
Joy (Humor,Sarcasm)	*Touring the town with Dad imagining all the places we will go when it's safe. And FYI, woke up feeling great. Side effects from vaccine were short lived.*
	Man, are the side effects of this vaccine immense joy and optimism? I'm the happiest I've felt in a year. I'm gonna see my parents this year. I'm gonna see a baseball game. I'm gonna travel. I'm so thrilled with this decision

Figure 3. Word Cloud – Gratitude

Figure 4. Word Cloud – Apprehension

The apprehension cloud in Figure 4 includes words such as *nervous*, *reactions*, *scared*, *worried*, *curious*, *question* and *wondering*. These words abundantly express the fear and anticipation of those who are yet to take the vaccine or yet to receive their second dose. They may be asking their friends about the reactions and side effects to allay their worries.

Finally, the sad word cloud includes people who have seen the harshest side effects. Fever, sore arm, sick, headache, chills nausea, tiredness, aches, and body pain list the range of symptoms. These symptoms are being compared to the flu. Superlative terms such as *pretty bad*, *kicking ass*, *worse*, *shit*, *hit*, *joke* describe the frustration resulting from the major inconvenience and disruption. Many probably refer to missing work, going to sleep earlier at night, hoping that their symptoms subside by the morning.

The word clouds considered together also lead to some interesting inferences. The clouds for joy and gratitude show the strongest overlap. The severity of the side effects appears to be intensifying from joy to gratitude to sadness. Finally,

Figure 5. Word Cloud – Sadness

the apprehension group constitutes those who are yet to be fully vaccinated (have not received even the first dose, or are in-between the first and the second doses).

V. Feature Mapping

We mapped the raw tweet data to word and numerical features. Word features are extracted from the content of the tweets, and numerical features are extracted from both the words and other metadata returned by the API.

A. Word Features

We extracted both statistical and semantic features from the words of the tweets. Statistical features are given by Term Frequency Inverse Document Frequency (TF-IDF) [31] scores computed using the sklearn library [3]. Using the the CountVectorizer, we converted the collection of pre-processed tweets into a matrix of vectors, each of which stores the frequency count of each of the words that occurs in the entire text. We then used the TfidfTransformer to transform the count matrix into a normalized TF-IDF representation matrix. The parameters are ngram(1,3) and the maximum number of features is equal to 50,000.

Semantic features were computed by GLoVe embeddings [22]. GLoVe is as an unsupervised learning algorithm that generates word embeddings based on the relationships between various words. These relationships are quantified by GLoVe's pre-defined embedding rules which assigns these numerical embeddings to each sentence according to their word relation contents. We used the GLoVe 100-dimensional vectors with 6 billion tokens trained on data from Wikipedia (the 2014 version) and Gigaword 5.

B. Numerical Features

Numerical features extracted from the content include punctuation and text-based emoticon counts. Expressions of emotion use certain punctuation characters for emphasis, and these characters may carry an emotional value that can help predict the sentiment of the tweet. Specifically, we include counts of '!' and '?', which are often used to express heightened feelings. We added counts of '@' and '#' to punctuation counts as they also provide additional information about whether a tweet is a part of a trend with the hashtag or if it is tagging someone. We also included counts of text-based emoticon counts. Unlike emojis, which are based on a Unicode value, text-based emoticons are 'smiley' or 'sad' faces drawn with simple characters, and are easier to process. We interpreted :), or ;) or =] as a happy face and :(, ;(or =[as a sad face.

The Twitter API returns additional numerical data which indicates how other users engage with a tweet and the behavior of the author of a tweet. These include the numbers of likes and retweets received by a tweet, and the number of tweets liked, status updates posted, and the numbers of friends and followers of the author.

VI. ML & DL Models

We considered three conventional machine learning models, namely, Logistic Regression, Support Vector Machines with the linear kernel, and multinomial Bayes. Implementations in the sklearn package were used, and optimal values of the hyperparameters were determined by *GridSearchCV*. We also implemented the following deep learning models, using Keras and TensorFlow 2.0.

- **Multi-Layer Perceptron:** This is a feed forward artifical neural network with an input, hidden, and several output layers. It takes multiple inputs with corresponding weights, processes this data based on the strength of each input, and returns a singular output that corresponds to all possibilities of inputs. We used ReLu activation function to minimize data loss associated with the gradient loss function. We implemented a basic MLP and also a MIMO MLP that involves multiple input and multiple output layers as shown in Figure 6. Using this architecture, we can incorporate other metadata features. One input layer is used to feed in pre-engineered GLoVe word embeddings, the second input layer feeds in numerical data. While an LSTM layer operates on language data, a simple dense neural network operates on numerical data to combine and generate final output weights. The outputs of these layers is concatenated and passed through several more dense layers to produce the final output classification.
- **Bi-directional LSTM:** This is a sequence processing model that consists of two LSTMs; one taking input in a forward direction, and the other in a backward direction. This allows the model to understand the relationships between words before and after each other, thus improving accuracy. The structure of our bi-directional LSTM is shown in Figure 7.

The basic MLP model is a simple deep learning model that was trained on the sklearn pipeline, using a combination of TF-IDF and numerical features. Thus, the basic MLP neural network was able to utilize *GridSearchCV*, similar to the

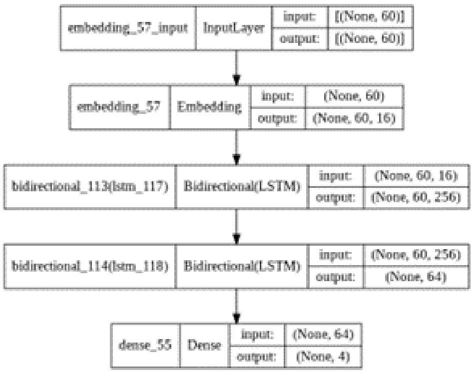

Figure 6. MIMO MLP Architecture

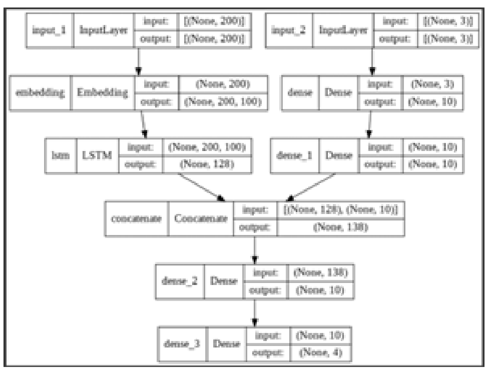

Figure 7. Bi-directional LSTM Architecture

conventional classifiers to obtain its optimal hidden layer sizes and shapes. The parameters of the models trained on the TF-IDF pipeline are summarized in Table II. On the other hand, GLoVe embeddings are often used to train deep learning models in the Tensorflow pipeline, and hence, were used to train bi-directional LSTM and MLP MIMO. These two deep learning models were run multiple times with various combinations of layer counts, neurons per layer, and validation splits.

VII. Performance Metrics

We define various metrics to assess the performance of the classifiers. Each tweet, for which the predicted label matches the ground truth (labeled assigned by the human annotators) is counted as a true positive. If there is a mismatch, it is counted as a false negative. Based on this designation, we define the following metrics:

- **Accuracy:** Accuracy is the percentage of tweets that are labeled correctly:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

- **Precision:** For a given class, precision is the percentage of tweets that actually belong to a class out of those that are predicted to belong to that class:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

- **Recall:** For a given class, recall is the percentage of tweets of a class that are predicted to be from that class:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

- **Precision-micro:** This is the sum of all true positives for classes, divided by all the positive predictions:

$$Precision\ micro = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n} (TP_i + FP_i)} \quad (4)$$

- **Recall-micro:** This is the sum of all true positives for all classes, divided by all the actual positives:

$$Recall - micro = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n} (TP_i + FN_i)} \quad (5)$$

- **Precision-weighted:** This is the weighted arithmetic mean of the binary precision scores of different classes.

$$Precision - weighted = \frac{\sum_{i=1}^{n} w_i * precision_i}{n} \quad (6)$$

- **Recall-weighted:** This is the weighted arithmetic mean of the binary recall scores of different classes.

$$Recall - weighted = \frac{\sum_{i=1}^{n} w_i * recall}{n} \quad (7)$$

- **F-score:** This metric computes the tradeoff between precision and recall. When computed using weighted (micro) precision and recall values it is designated as F1-weighted (F1-micro).

To calculate precision-micro and recall-micro scores for sentiment and emotion classification, the number of classes is 2 and 4 respectively. Per class precision and recall scores are used in the computation of weighted precision and recall. These per class scores are calculated by designating that class as positive and the remaining classes as negative. In sentiment classification, precision and recall are computed separately for positive and negative classes. In multi-label emotion detection, when calculating precision and recall for a given class, say joy, that class is labeled as positive, and the remaining three classes, namely, apprehension, sadness, and gratitude are designated as negative.

Table II
Model Hyperparameters

Model	Parameters
Logistic Regression	C = 1.00
SVM	C = 10, $\gamma = 0.0001$
Multinomial NB	$\alpha = 1$, *class_prior = None* *fit_prior = False*
NN (MLP)	solver = 'lbfgs', $\alpha = 1e^{-5}$ *max_iter* = 500, *hidden_layer_sizes = (512,512,512)*

VIII. Results & Discussion

We split the entire corpus $80\% - 20\%$ into training and test partitions using stratified sampling to preserve the relative proportions of the different sentiment and emotion labels in the two partitions. We trained the conventional machine learning models on the training data using two combinations of features. In the first experiment, we used just word features, in the second experiment we added numerical features. The results in Table III indicate that the performance of all the models for both sentiment and emotion classification improves by about 3-4% by incorporating numerical features. Therefore, we combine these numerical features with GLoVE embeddings to train deep learning models.

We observe that the Naive Bayes model offers the best performance for both the classification problems (accuracy of 0.81 for sentiment and 0.67 for emotion classification). All the machine and deep learning classifiers performed better on sentiment classification compared to emotion detection. That is, they were able to discern that the emotions of joy and gratitude were in contrast with the emotions of apprehension and sadness, but not able to differentiate as well between joy vs. gratitude and apprehension vs. sadness. The impact of numerical features in binary classification was more pronounced compared to the multi-label problem. Including these features improved the accuracy by $4 - 6\%$ for most models, with the largest improvement of about 10% seen for the Naive Bayes model. Similar to Kumar *et. al.* [12], in our research conventional machine learning models outperformed deep learning approaches. Further, we note that the overall performance of our classifiers is at par with the performance for multi-label emotion classification problems summarized in our prior work [19].

To understand how our classifiers fare with respect to individual emotions we examined 4x4 confusion matrices. These show that the classifiers were able to isolate apprehension and gratitude more easily compared to joy and sadness. Joy was the most challenging, since these tweets combine pure joy, humor as well as sarcasm. Sarcasm often blends positive and negative elements, and most readily confuses the classifiers [21]. All in all, given that we sought to identify complex emotions such as apprehension and gratitude which do not appear on the main axes of any emotion models and groups, we consider our models which achieve an accuracy of around 65%, similar to other contemporary research, to be quite successful.

IX. Conclusions and Future Research

This paper mines sentiment and emotions from social media conversations on the side effects of the Covid-19 vaccine. Although the Covid-19 vaccine offered a promising path toward the end of the pandemic, many were nervous and uncertain about its side effects. The research reported in this paper unconvers positive emotions of joy and gratitude which signal an acceptance of the vaccine, and negative emotions of apprehension and sadness which may later turn to vaccine hesitancy. Therefore, this paper contributes towards our collective understanding of the perceptions of the Covid-19 vaccine, and how it is received by the public. These results can thus be useful for developing plans to disseminate authoritative information about the vaccine aimed at building understanding and trust.

Our future research is concerned with incorporating features related to sarcasm and irony in improving classification accuracy. Mining emotions from other topical conversations related to the Covid-19 pandemic is also a concern of the future.

References

[1] K. N. Alam, S. Md Khan, A. Rab Dhruba, M. Khan, J. Al-Amri, M. Masud, and M. Rawashdeh. Deep learning-based sentiment analysis of COVID-19 vaccination responses from Twitter data. *Computational and Mathematical Methods in Medicine*, 2021.

[2] J. Md Tarique Ansari and N. A. Khan. Worldwide COVID-19 vaccines sentiment analysis through Twitter content. *Electronic Journal of General Medicine*, 18(6):2516–3507, 2021.

[3] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and Gaël Varoquaux. "API Design for Machine Learning Software: Experiences from the Scikit-learn project". In *Proc. of ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.

[4] V. P. Bustos, C. D. Comer, S. M. Manstein, E. Laikhter, E. Shiah, H. Xun, B. T. Lee, and S. J. Lin. Twitter voices: Twitter users' sentiments and emotions about COVID-19 vaccination within the United States. *European Journal of Enviornment and Public Health*, 6(1), 2022.

[5] F. Cascini, A. Pantovic, Y. A. Al-Ajlouni, G. Failla, V. Puleo, and A. Melnyk. Social media and attitudes towards a COVID-19 vaccination: A systematic review of the literature. *eClinicalMedicine*, 48, June 2022.

[6] R. Chandra and A. Krishna. COVID-19 sentiment analysis via deep learning during the rise of novel cases. *PLOS ONE*, August 2021.

[7] R. Chandrasekaran, R. Desai, H. Shah, V. Kumar, and E. Moustakas. Examining public sentiments and attitudes toward COVID-19 vaccination: Infoveillance study using Twitter posts. *JMIR Infodemiology*, 22(1), Jan-Jun 2022.

[8] S. AE Chun, A. Chih-Yuan Li, A. Toliyat, and J. Geller. Tracking citizens' concerns during COVID-19 pandemic. In *The 21st Annual International Conference on Digital Government Research*, June 2020.

[9] R. Durante. No-vaxxers: The emotional drivers and how to navigate barrires to the COVID-19 vaccine. https://www.marugroup.net/insights/blog/no-vaxxers-and-emotional-barriers, January 2022.

[10] N. H. Frijda. *The Emotions*. Cambridge University Press, 1986.

[11] M. Harris. My personal experience with the Pfizer vaccine. https://www.psychologytoday.com/us/blog/letters-your-therapist/202102/my-personal-experience-the-pfizer-vaccine, February 2021.

[12] H. S Hemanth Kumar, Y. P Gowramma, S. H Manjula, D. Anil, and N. Smitha. Comparison of various ml and dl models for emotion recognition using twitter. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, pages 1332–1337, 2021.

[13] Y. Hoffman, Y. Palgi, R. Goodwin, M. Ben-Ezra, and L. Greenbalt-Kimron. Severe Covid-19 vaccine side-effects are rare in older adults yet are linked with depressive symptoms. *The American Journal of Geriatric Psychiatry*, 30(1):115–116, January 2022.

[14] A. Shariq Imran, S. Muhammad Daudpota, Z. Kastrati, and R. Batra. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets. *IEEE Access*, 8:181074–181090, 2020.

[15] H. Jelodar, Y. Wang, and R. Orji. Deep Sentiment Classification and topic discussion on novel coronavirus or COVID-19 online discussions: NLP using LSTM Recurrent Neural Network approach. *IEEE J Biomed Health Forum*, 10:2733–2742, October 2020.

[16] M. W. Kearney. Collecting Twitter data. https://cran.r-project.org/web/packages/rtweet/rtweet.pdf, January 2020.

[17] Y. Li, S. Luan, Y. Li, and R. Hertwig. Changing emotions in the COVID-19 pandemic: A four-wave longitudinal study in United States and China. *Soc Sci Med*, September 2021.

Table III
PERFORMANCE METRICS OF ML & DL MODELS

		Emotion (Multi-class)		Sentiment (Binary)	
ML Models					
Classifier	Metric	TF-IDF	TF-IDF Numerical	TF-IDF	TF-IDF Numerical
Logistic Regression	Accuracy	0.52	0.61	0.69	0.75
	F1-score (micro)	0.52	0.61	0.69	0.75
	F1-score (weight)	0.50	0.61	0.69	0.74
SVM	Accuracy	0.54	0.61	0.62	0.68
	F1-score (micro)	0.54	0.61	0.62	0.69
	F1-score (weight)	0.53	0.60	0.62	0.68
Naive Bayes	Accuracy	0.64	0.65	0.70	0.81
	F1-score (micro)	0.52	0.65	0.7	0.81
	F1-score (weight)	0.51	0.65	0.70	0.80
Basic MLP	Accuracy	0.55	0.59	0.72	0.74
	F1-score (micro)	0.55	0.59	0.72	0.74
	F1-score (weight)	0.54	0.57	0.72	0.73
DL Models					
Classifier	Metric	GLoVE	GLoVE Numerical	GLoVE	GLoVE Numerical
Bi-LSTM	Accuracy		0.64		0.71
	F1-score (micro)		0.63		0.70
	F1-score (weight)		0.64		0.71
MIMO MLP	Accuracy		0.62		0.67
	F1-score (micro)		0.62		0.67
	F1-score (weight)		0.60		0.67

[18] L. McClurh. COVID-19 vaccination has been conjuring up emotions and memories. https://www.npr.org/2021/03/21/978727427/covid-19-vaccination-has-been-conjuring-up-emotions-and-memories, March 2021.

[19] A. Mondal and S. S. Gokhale. Multi-label classification of Parrott's emotions. In *International Conference on Software Engineering and Knowledge Engineering*, pages 359–364, July 2021.

[20] M. Monselise, C. Chang, G. Ferreira, R. Yang, and C. Yang. Topic and sentiments of public concerns regarding COVID-19 vaccines: Social media trend analysis. *J. of Med Internet Res*, October 2021.

[21] H. Nguyen, J. Moon, N. Paul, and S. S. Gokhale. Sarcasm detection in politically motivated social media content. In *2021 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, pages 1538–1545, 2021.

[22] J. Pennington, R. Socher, and C. Manning. "GloVe: Global Vectors for Word Representation". https://nlp.stanford.edu/projects/glove/, August 2014. Accessed: 2022-01-21.

[23] S. Reinfeld, R. Caceda, R. Gil, H. Strom, and M. Chacko. Can new onset psychosis occur after mRNA based COVID-19 vaccine: A case report. *Psychiatry Res*, October 2021.

[24] J. Ries. People with anxiety and depression may need a COVID-19 booster shot. https://www.healthline.com/health-news/people-with-anxiety-and-depression-may-need-a-covid-19-booster-shot, October 2021.

[25] F. M. J. M. Shamrat, S. Chakraborty, M. M. Imran, J. N. Muna, Md M. Billah, P. Das, and Md. O. Rahman. Sentiment analysis on Twitter tweets about COVID-19 vaccines using NLP and kNN classification algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*, 23(1):463– 470, 2021.

[26] N. V. Uvais. New-onset panic disorder following Pfizer-BioNTech COVID-19 vaccination. *Prim Care Companion CNS Disord*, 24(1), 2022.

[27] N. Vemprala, P. Bhatt, R. Valecha, and H. R. Rao. Emotions during the COVID-19 crisis: A health versus economy analysis of public responses. *American Behavioral Scientist*, October 2021.

[28] J. Wang, Y. Fan, J. Palacios, Y. Chai, N. Guetta-Jeanrenaud, N. Obradovich, C. Zhou, and S. Zheng. Global evidence of expressed sentiment alterations during the COVID-19 pandemic. *Nature Human Behavior*, pages 349–358, March 2022.

[29] X. Yang and V. Sornlertlamvanich. Public perception of covid-19 vaccine by tweet sentiment analysis. In *2021 International Electronics Symposium*, pages 151–155, 2021.

[30] S. Yousefinaghani, R. Dara, S. Mubareka, S. Papadopoulos, and S. Sharif. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. *International Journal on Infectious Diseases*, 108:256–262, July 2021.

[31] R. Zafarani, M. Ali Abbasi, and H. Liu. *Social media mining: an introduction*. Cambridge University Press, 2014.