IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 5, 2021, accepted September 2, 2021, date of publication September 7, 2021, date of current version September 23, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3111035*

# Biden vs Trump: Modeling US General Elections Using BERT Language Model

**ROHITASH CHANDRA[1], (Senior Member, IEEE), AND RITIJ SAINI[2]**
[1]School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia
[2]Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India

Corresponding author: Rohitash Chandra (rohitash.chandra@unsw.edu.au)

**ABSTRACT** Social media plays a crucial role in shaping the worldview during election campaigns. Social media has been used as a medium for political campaigns and a tool for organizing protests; some of which have been peaceful, while others have led to riots. Previous research indicates that understanding user behaviour, particularly in terms of sentiments expressed during elections can give an indication of the election outcome. Recently, there has been tremendous progress in language modelling with deep learning via *long short-term memory* (LSTM) models and variants known as *bidirectional encoder representations from transformers* (BERT). Motivated by these innovations, we develop a framework to model the US general elections. We investigate if sentiment analysis can provide a means to predict election outcomes. We use the LSTM and BERT language models for Twitter sentiment analysis leading to the US 2020 presidential elections. Our results indicate that sentiment analysis can provide a general basis for modelling election outcomes where the BERT model indicates Biden winning the elections.

**INDEX TERMS** Language models, deep learning, election modelling, sentiment analysis, BERT, US elections.

## I. INTRODUCTION

Political forecasting is an area where analytical and statistical methods predict election outcomes mainly using surveys and qualitative approaches [1]. This also includes analysis of manifesto of political parties while looking at the trend of the popular news media, which is also known as political analysis [2], [3]. The forecasting of elections became more difficult with growing opposition in government, especially in countries such as the USA, where the elections take place between two dominant parties [4]. There are major challenges in getting a good representation of opposing political viewpoints when it comes to data collection [5]–[8]. Social networks such as Facebook and Twitter have somewhat addressed limitations of representation in sampling via surveys. Social networks have been at the forefront of political campaigns and activism during elections [9]–[11].

Over the last decade, there has been some interest in using social media to forecast the outcome of elections. This has been mainly through artificial intelligence via natural language processing (NLP) methods [12], [13]. These methods range from basic statistical methods to complex models that include deep learning [14], [15]. Election modelling include strategies such as topic modeling and sentiment analysis [16]–[18] and some of the relevant studies are reviewed as follows. Agarwal *et al.* [19] used long short-term memory (LSTM) networks and prominent word embedding for 41 million tweets for the 2019 Indian general elections where the predictions showed a close correlation with the actual results. Suciati *et al.* [20] used machine learning to detect buzzer accounts that disseminate information deliberately for the 2019 Indonesian Presidential elections. Mohbey [21] analyzed user opinion for topic modeling for the 2019 Indian general elections and gathered information that could assist the government and businesses to revise strategic policies. Vijayaraghavan *et al.* [22] presented a framework using deep learning for analyzing election-related conversation on Twitter on a continuous and longitudinal basis for the 2016 US Presidential elections. Li *et al.* [23] used deep hierarchical graph convolution for election prediction from geospatial data taken from the 2016 Australian census.

Sentiment analysis applies NLP methods [24] to provide an understanding of affective states and emotions [25]–[27]. Sentiment analysis has been prominent in understanding

The associate editor coordinating the review of this manuscript and approving it for publication was Khoa Luu.

customer behaviour [28], health and medicine [29], stock market predictions [30], and modelling election prediction such as the 2012 US Presidential elections [15]. Data from social media with deep learning provides a powerful tool in sentiment analysis [16]–[18]. Language models are continuously updated with innovative methods in deep learning. Attention-based mechanism has been used to improve long short-term memory (LSTM) models for language modelling [31]. The transformer model is an enhanced LSTM model that incorporate attention mechanism into encoder-decoder LSTM models [32], [33]. Moreover, *bidirectional encoder representations from transformers* (BERT) [34] developed by Google, has been at the forefront of language models. BERT has been trained with a large data corpus with more than 300 million models parameters useful for tasks such as topic modeling, language translation, and sentiment analysis. Recently, BERT has been applied in China for COVID-19 topic modelling and sentiment analysis [35]. BERT has also been used for time-dependent sentiment analysis [36] and document retrieval [37]. Apart from other language modelling applications, we believe that BERT can be very useful for modelling election outcome via sentiment analysis.

The 2020 US Presidential election featured an intense competition between Democrat party candidate Joe Biden and Republican party candidate Donald Trump. Due to an intense campaign prior to the elections, there has been political unrest and fierce online activities during the first wave of COVID-19 [38]. The political conflict between the two presidential candidates reflected in dispute and abusive debates on social networks such as Twitter which led to Capital riots just after the elections [39]. President Donald Trump was banned by Twitter as it was alleged that his comments led the Capitol riots. Social media plays a crucial role in political campaigns, activism and unrest [40]. This has been shown by analysis of tweets before and during the Capitol riots [39]. Although there has been some work done using tweets in predicting election outcomes [41], our paper focuses on sentimental analysis via deep learning using tweets during US presidential elections.

In this paper, we present a framework that uses sentiment analysis via state-of-art language models to understand public behavior during elections. We employ BERT and LSTM-based language models for sentiment analysis. We use the internet movie database (IMDB) as training dataset that provides polarity scores indicating positive and negative sentiments. We investigate if sentiment analysis from social media can help in modelling and understanding voter behaviour during the elections.

The rest of the paper is organized as follows. In Section 2, we present a framework that uses sentiment analysis to predict election outcomes. Section 3 presents a visualization of the data and prediction results, and Section 4 provides a discussion with focus on the implications of the results. Section 5 concludes the paper with directions for future research.

## II. METHODOLOGY

### A. TWITTER DATA EXTRACTION AND PROCESSING

We extract the raw US-2020 elections dataset [42], [43] that features tweets from October 15th 2020 to November 8th 2020. We consider tweets that have geo-location within USA from the dataset of 1.72 million tweets. The dataset follows a similar extraction process implemented for the study of 2017 UK elections [44].

We implement language-based processing by classifying individual tweets based on user identification and filter the English language origin tweets. After language-based processing, further cleaning is done to remove links and special characters in tweets. We process the US-2020 elections tweets using a software application known as *tweepy* [43]. We consider tweets only in English for the sentiments that relate to the US elections using the *langdetect*[1] python library. We consider tweets only in English for the sentiments that relate to the US elections using the langdetect python library. We process the special phrases and expressions such as hashtags (#), emotion symbols (emojis), stop words (eg. ''the'', ''an'', ''you'' ), https links, and abbreviations and translate them into known English words as shown in Table 1. We also convert all tweets to lowercase strings. We do not correct the misspellings, this bias is present both in the training and the test dataset. Hence, we allow our model to learn representations of tweets that feature misspelling to make predictions. While metadata of each tweet contains a multitude of attributes, we focus in extracting only specific variables such as tweet location, retweet count, and time of tweet created. In case if the user state location is missing, non-applicable (NA) is given. We also create a data dictionary to map major states such as Kentucky, Wyoming, New Hampshire, and others for the final polarity mapping. We do not remove the stop words (eg. ''a,'' ''and,'' ''but,'' ''how,'' ''or,'' and ''what.''), since it can eliminate information regarding the sentiment. Similar approach has been used in our previous work [45]. We also remove text case sensitivity (i.e., lower or upper case) in the tweets.

### B. FRAMEWORK

A tweet with a political viewpoint could feature sentiments for or against a subject, such as a political party or candidate. The sentiments expressed in such tweets are not easy to classify since the way emotions are expressed with words are often complex by different users with different regional and cultural backgrounds. Sentiment analysis is challenging given various features in tweets, number of character limit in Twitter, semantics, and context. The tweets that feature sentiments that are for or against the subject can have a score which is known as polarity which can be highly susceptible to inconsistency and redundancy [46]–[48]. Moreover, some users change their stand about a matter with time. Often in Twitter debates, people express comments with serious

[1] https://pypi.org/project/langdetect/

**TABLE 1.** Examples of tweet processing, before and after removing special characters, web-links, and newline.

| Original Tweet | Transformed word/tweet |
|---|---|
| Count how many times #Trump says "if you look at..." #LPTVDebateTakes | Count how many times Trump says if you look at... LPTVDebateTakes |
| @IslandGirlPRV @BradBeauregardJ @MeidasTouch This is how #Biden made his ! #TrumpIsNotAmerica! \nhttps://t.co/uBqAFU86Ip | IslandGirlPRV BradBeauregardJ MeidasTouch This is how Biden made his ! TrumpIsNotAmerica ! |
| FBI Allegedly Obtained Hunter Biden Computer, Data on Ukraine Dealings, Report Claims #JoeBiden #HunterBiden https://t.co/pDNmB0NqRU | FBI Allegedly Obtained Hunter Biden Computer, Data on Ukraine Dealings, Report Claims JoeBiden HunterBiden |
| @DeeviousDenise @realDonaldTrump @nypost There won't be many of them. Unless you all have been voting more than once again. But God prevails. BO was the most corrupt President ever. Dark to light. Your lies are all coming through. They wouldn't last forever. #Trump | DeeviousDenise realDonaldTrump nypost There wont be many of them. Unless you all have been voting more than once again. But God prevails. BO was the most corrupt President ever. Dark to light. Your lies are all coming through. They wouldnt last forever. Trump |

opposition of political views that leads to hate speech and online abuse [49], [50].

Our overall goal is to review sentiments expressed in the tweets few months prior to the elections and find if they can provide insights regarding the election outcome. We use sentiment analysis via deep leaning (LSTM or BERT model) for understanding the nature of the tweets in terms of polarity (i.e., intensity of sentiments indicating support for either Biden or Trump). Figure 1 presents the framework for sentiment analysis which provides an indication of US election outcome. The predicted sentiment polarity score is a real number which includes negative and positive values in a range [−1,1]. At first, the US election tweets are collected by software applications and then processed as described earlier. The BERT and LSTM language models are then trained using a labelled dataset (IMDB dataset) to predict the polarity of processed-tweets. The polarity score is then used to map the overall nature of voters in the electoral states, solely on the basis of the total sentiments for individual candidates (i.e., Biden and Trump). The framework provides a prediction for all the respective US electoral states, which also includes the swing states. Finally, we analyse the predictions and compare with the actual electoral results for all US states with emphasis on the swing states i.e., the states that are highly unpredictable.

### C. WORD EMBEDDING

Word embedding is a technique that maps textual tokens, e.g., words, into dense and low-dimensional vector representations which are generated by large unlabelled corpus. Mikolov *et al.* [51] proposed *word2vec* word embedding which uses a simple neural network to learn word associations which can be used to find synonymous words and provide additional words given an incomplete sentence. The word2vec model features two training approaches, which includes the skip-gram model and the common bag of words (CBOW) model. CBOW embeds a word on the basis of the words within the surrounding context, while skip-gram embeds the word within the surrounding context starting from the current word. These methods have been used for measuring semantic similarity [52] between texts and topics [53]. They have been used in conjunction with deep neural networks for language modelling tasks such as topic modeling and semantic analysis [54]. We note that we use word2vec embedding in our framework (Figure 1) for the LSTM language model. Although word embedded models can be trained, our framework employs a pre-trained word2vec word embedding for the LSTM model from the natural language toolkit (NLTK) library[2] where vector embedding size is set to 100, and the maximum length of input text is limited to 140.

### D. TECHNICAL SETUP AND MODEL TRAINING

In the case of the BERT model, we use inbuilt word embedding based on BERT-base uncased [3] where the English language is used for masked language modeling (MLM) [34]. The BERT and LSTM models are trained further using the IMDB dataset [55] with training and test dataset using 80/20% split and a batch-size of 100. The batch size defines the number of training instances before updating the internal model parameters which can play an important role in improving model performance. The IMDB dataset classifies the data into either "positive" or "negative" class based on movie reviews. Note that our model prediction gives a sentiment polarity score. We note that the polarity score is in the range [−1,1]; hence, our predictions are transformed since we use sigmoid activation function in the output layer of the respective models.

The adaptive moment estimation (Adam) [56] optimizer is used for training with a learning rate $lr = 3e - 05$, numerical stability constant $c = 1e - 08$, and maximum gradient norm $n = 1$ which clips the gradient. We use the limit of 140 characters for input sequences in our BERT model. Since BERT is a pre-trained model, we refine it with a training time of 4 epochs which gave good performance in trial experiments. Note that data processing is slightly

[2]https://www.nltk.org/
[3]https://huggingface.co/bert-base-uncased

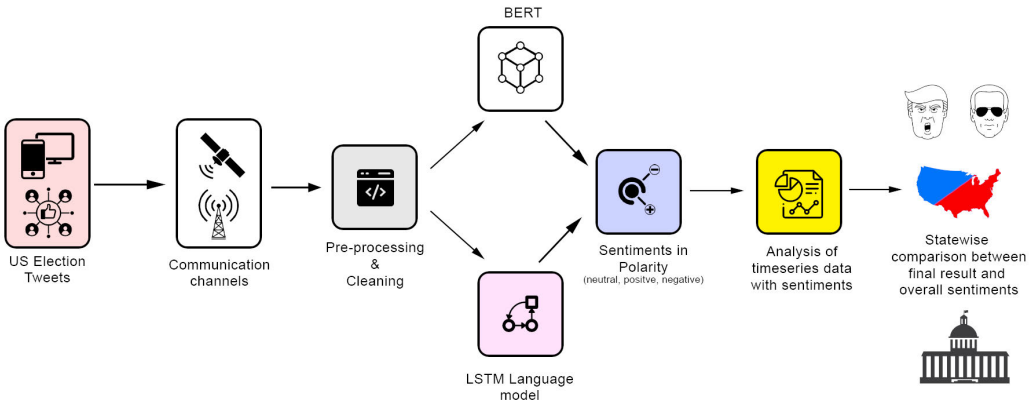



**FIGURE 1.** Framework for twitter-based sentiment analysis for US elections using LSTM and BERT language models.

different for BERT model since it can cater for more features when compared with LSTM. BERT provides attention to specific features in sentence due to being pre-trained from a large corpus and uses its own word embedding whereas LSTM uses word2vec.

In the case of the LSTM model, the overall approach is similar to the BERT model with minor changes in cleaning of dataset where we remove the stop-words, hashtags, uppercase letters and punctuation to extract better features. In model training for LSTM, the tweets are limited to 140 characters with embedding dimension to 32 using word2vec embedding, and the model is trained for 10 epochs. These hyper-parameter values have been determined from trial experiments. Note that the LSTM model is trained longer since BERT model is pre-trained and features knowledge from language corpus [55]. The model architecture showing number of trainable parameters of LSTM and BERT model is given in Table 2.

## III. RESULTS

### A. DATA ANALYSIS

We note that the Twitter dataset (1.17 million tweets) features tweets from 15th Oct 2020 - 11th November, 2020 which covers the first presidential debate to declaration of the final results [43]. Figure 2 presents a global visualisation by showing the locations of tweets. There is interest in the US elections from many different countries in the world with tweets from 40 different languages; however, a large proportion of the tweets are in English that originate from the US with 92,984 classified as English tweets using the Lang-detect library [57]. We note that only 47.25 percent of the data-set (544,885 out of 1,153,079) tweets has user location. We note that Twitter users can decide if they need to show their location. We find that majority of the tweets came from USA and Europe (Figures 2 and 3), followed by India which has a large population of growing internet and Twitter users.[4]

The information about the exact number of tweets for different location for Trump and Biden datasets is given in Figure 3, where we see that majority of the dataset is marked by location not available (NA). Figure 4 presents further details for the missing information (null values) in the number of tweets for Trump and Biden datasets showing missing information regarding user location and further details given by city, country, state and continent. Note that this information is shown by "Geo Data NA" in Figure 3.

Figure 5 shows the trend in the number of tweets per hour from 20th October to 20th November, 2020 which covers the span before the elections. We note that there is a major spike around 23rd of October (10-23) which is due to the election debate held at the Belmont University.[5] Similarly, from 1st November (11-01), the number of tweets gradually increases with major spike around 8th November which is due to Biden being declared "President Elect" by majority of the news organizations.[6]

Figure 6 presents the leading ten bi-grams and tri-grams from processed tweets from Trump and Biden datasets. We observe that the bi-grams are mostly descriptions of their respective roles and names. It is striking to see "joe - biden" as the second highest bi-gram in the Trump dataset along with "antitrump - please". The tri-grams on the other hand are more descriptive of support for Trump. In the case of Biden, we see "warning - awaits - u" and "video - warning - awaits" which seem to be either negative sentiments or sentiments showing concern.

### B. MODELLING AND PREDICTIONS

First, we provide model training prediction accuracy results that compare LSTM and BERT model using the IMDB dataset as shown in Table 3. Note that the training dataset is class balanced with 25,000 positive and 25,000 negative movie classified reviews [55]. BERT and LSTM models

[4]https://www.statista.com/statistics/255146/number-of-internet-users-in-india/

[5]https://www.theguardian.com/news/2020/oct/21/the-worlds-election-inside-the-23-october-guardian-weekly

[6]Joe Biden elected president: https://edition.cnn.com/politics/live-news/trump-biden-election-results-11-08-20/index.html

**TABLE 2.** LSTM and BERT model configuration.

| Model | Layer (type) | Output Shape | Param # |
|---|---|---|---|
| LSTM | Input Layer (Word Embedding) | (32) | 2948384 |
| | Hidden Layer 1 | (64) | 24832 |
| | Hidden Layer 2 | (64) | 4160 |
| | Output Layer | (1) | 65 |
| BERT [34] | | | 109482240 |

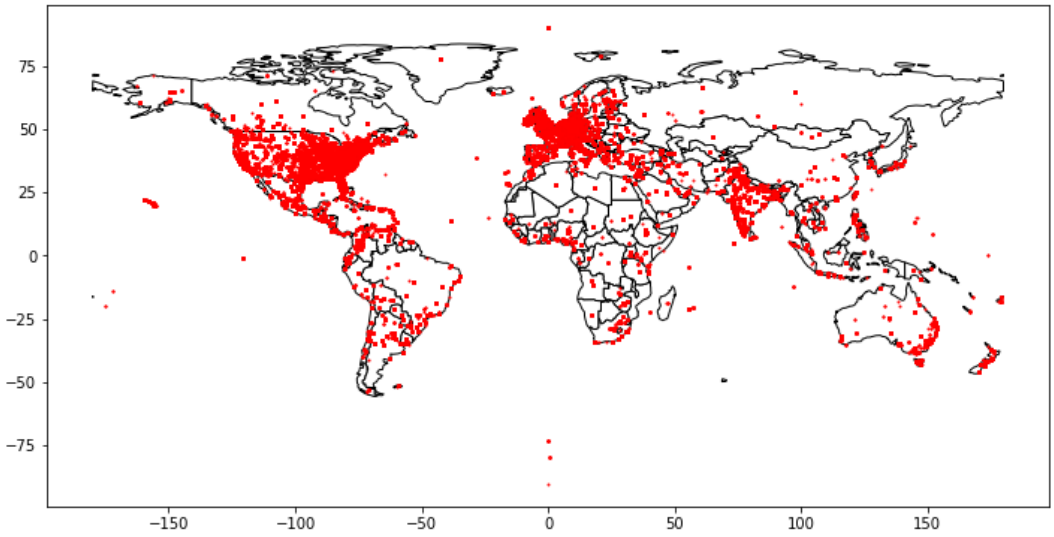



**FIGURE 2.** Tweet geo-location showing majority of tweets coming from USA, Europe and India.

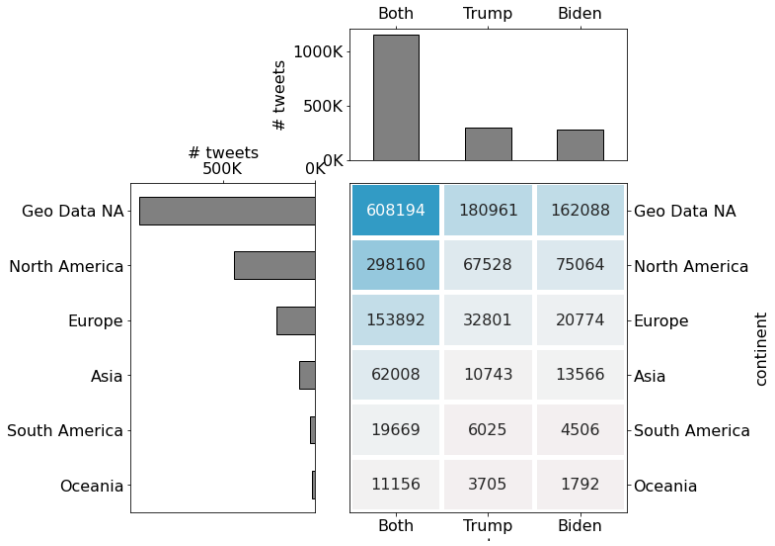



**FIGURE 3.** Geo-data for tweets for each candidate (Trump and Biden) across the major continents.

use learning rate of 1e-04 and 1e-05 for the Adam optimizer, respectively. We execute 30 experimental runs with different model parameter (weights and bias) initialization and consider different combinations of a batch size of the training dataset. We find that batch size of 64 provides better results for both models in terms of classification accuracy (mean and standard deviation) and the best F1-score documented in Table 3. We find that both BERT and LSTM provide similar training performance in terms of the F1 score.

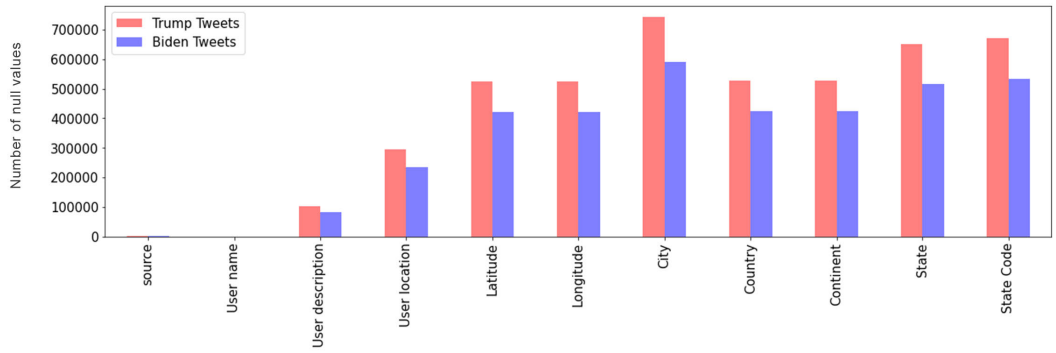



**FIGURE 4.** Number of missing (null) values present in tweets for the respective datasets (Trump and Biden) showing that there is mostly missing information regarding user location and further details given by city, country, state and continent. Note that this is giving further information about the "Geo Data not available (NA)" shown in Figure 3.

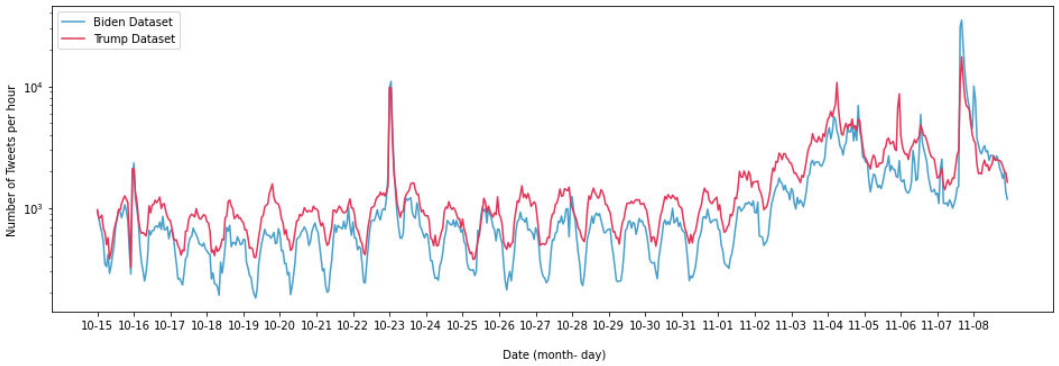



**FIGURE 5.** Tweet per hour from October 15th to 11th November, 2020.

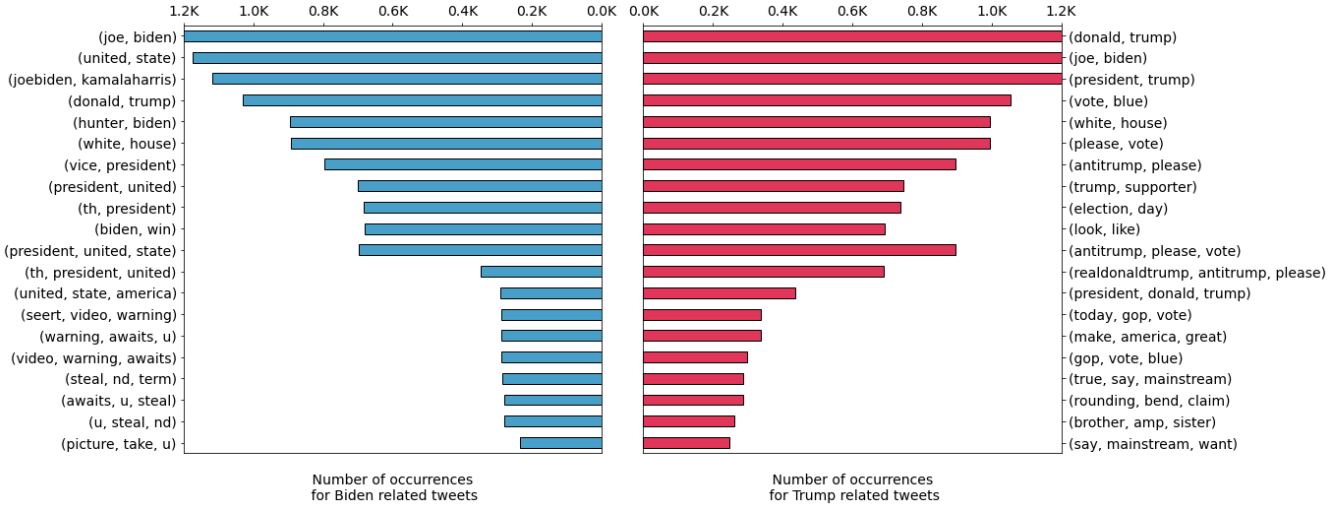
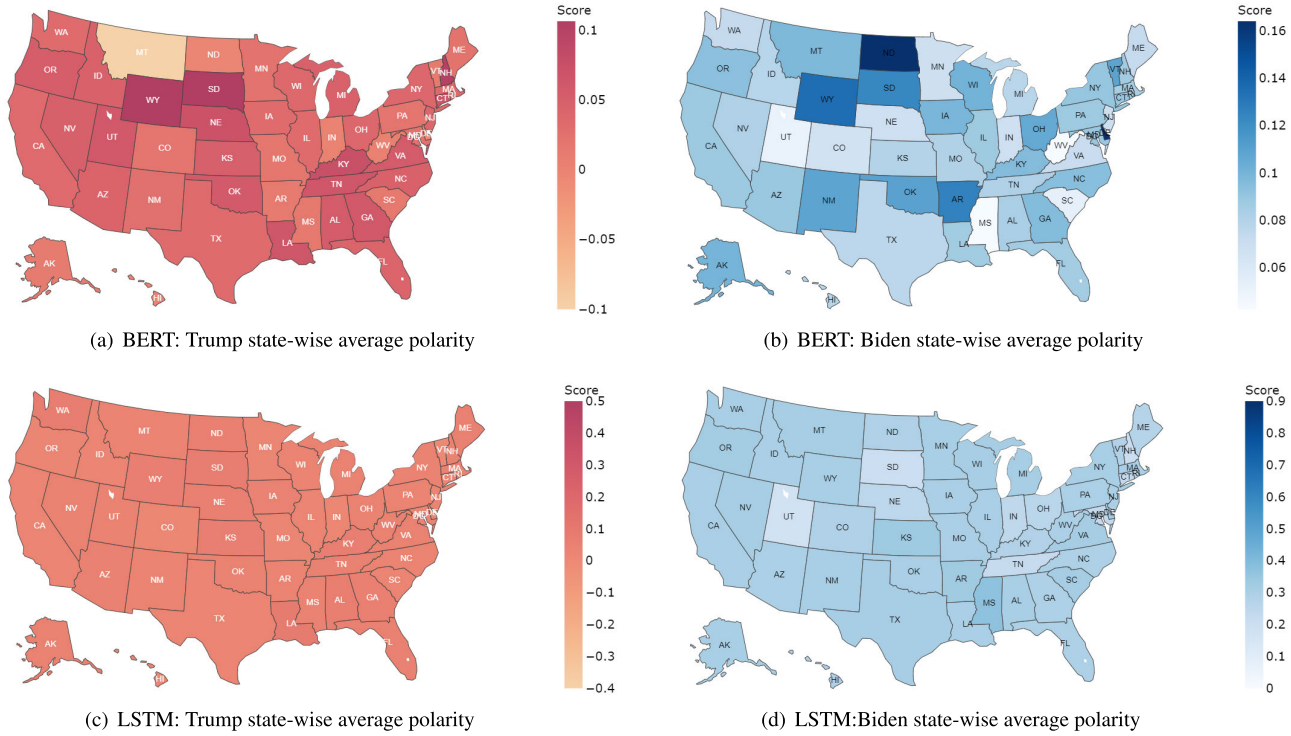



**FIGURE 6.** Top 10 Bi-grams and Tri-grams from processed tweets for Trump and Biden datasets.

Next, we present results with our trained models that features a binary classification dataset and sigmoid activation function in the output layer of the respective models. The predictions are transformed in the range $[-1,1]$ in the test phase to represent the sentiment polarity score. In this way, we develop a model for binary classification using training data which is then used for sentiment polarity prediction using the test data.

Figure 7 presents state-wise average polarity from predictions given by LSTM and BERT models. We calculate the state-wise average polarity by averaging the individual polarity score of the tweets for the respective states. Our




(a) BERT: Trump state-wise average polarity

(b) BERT: Biden state-wise average polarity

(c) LSTM: Trump state-wise average polarity

(d) LSTM:Biden state-wise average polarity

**FIGURE 7.** State-wise sentiment polarity predictions by LSTM and BERT models for Trump and Biden.

**TABLE 3.** Result after 30 experiments using IMDB dataset given different initial learning rate (LR) and batch size of the training dataset.

| Model | Batch Size | Mean (Std) Accuracy(%) | F1 Score |
|---|---|---|---|
| BERT | 64 | 87.45 (1.74) | 75.7 |
| | 128 | 86.60 ( 1.40 ) | 70.2 |
| LSTM | 64 | 88.10 (2.11) | 76.5 |
| | 128 | 85.62 (1.23) | 68.6 |

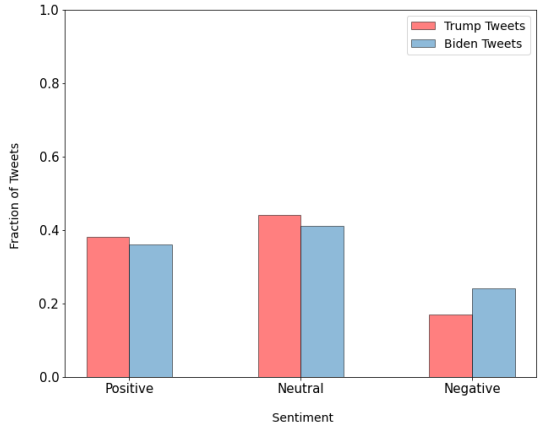
definition of a contentious state is when the absolute ratio between either Trump or Biden average state-wise polarity score of the electoral state is less than $\alpha$. In order to determine the best value of $\alpha$, we ran trial experiments from which we selected $\alpha = 1.5$, after comparing with past election results (2016 US presidential election) in order to capture the state's electoral history. The goal is to capture prior voting information by states to determine the winning candidate.

In Figure 7 (Panel a) BERT for Trump dataset, we notice that in the state of Montana (MA), the average polarity is much lower (below $-0.1$) when compared to the rest of the states. In the Biden dataset (Panel b), we find that there is no state with negative average polarity. In LSTM prediction for the Trump dataset, the total sentiment polarity for most of the states is positive (Panel c). In (Panel d), the case of the Biden dataset shows no negative polarity in predictions. So far, we can assert that both LSTM and BERT models have predicted positive average polarity for the Biden dataset, while the Trump dataset have some states with negative polarity.

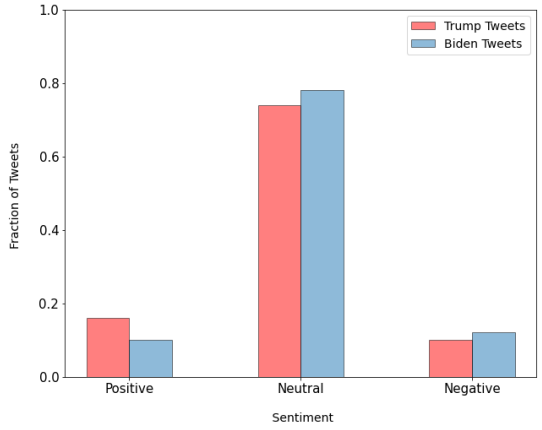
Finally, we present results that show the predictions based on BERT and LSTM language models. The polarity score is based on the sentiments (negative/positive) and normalized final score for individual electoral states. The normalisation range is between $[-1, 1]$ and defined by $p = \frac{x}{\sqrt{x^2+\alpha}}$; where, $\alpha$ is user-defined constant, $x$ is the sentiment score, and $p$ is the polarity. Based previous research [58], we use $\alpha = 15$ which approximates the maximum expected value of $x$.

Figure 8 shows the prediction by giving the percentage of tweets by positive, neutral, and negative sentiments for respective candidates using BERT and LSTM models. In the case of BERT, we observe that both candidates have similar level of neutral tweets and positive tweets, with a lower number of negative tweets where Biden has more negative tweets than Trump. In the LSTM model, we find that the number of negative and positive tweets is similar, but there is a large influx of neutral tweets, which is almost double when compared with the BERT model. These predictions show model bias which can be due to the model architecture and also due to the information that was already present in the pre-trained BERT model. We can quantify these predictions only by comparing the actual election results, which will be shown in the later section of this paper.

Table 4 shows the prominent electoral state's average sentiment and their actual result comparison based on the BERT model (Figure 7, Panel a and b) for Trump and Biden state-based polarity score. We find that in the top 3 states given by positive polarity score, the chances of winning for Trump are




(a) BERT




(b) LSTM

**FIGURE 8. Fraction of tweets by positive, neutral, and negative sentiments for respective candidates using LSTM and BERT.**

higher in Wyoming, New Hampshire, and Kentucky. We note that Trump lost the state of New Hampshire as given by the actual result since Biden's score is close, which is also the case of Kentucky. In the actual result, we observe that Biden won the state of New Hampshire while Trump won in Wyoming and Kentucky. Moreover, New Hampshire has been one of the swing states that implies that either the Democratic or Republican presidential candidate can win [59].

Similarly, Wyoming gained its position in the Top 3 positive states for both Trump and Biden. Also, Delaware is a top positive state for Biden and one of the most hostile states for Trump, and thus it is not surprising that Biden won with a good margin in actual results. In the case of North Dakota, the previous result could not be reinstated because it is one of the most favorable states for Biden, and has been a hostile state for Trump. However, Trump won probably due to his dominance, and we need further analysis to explain why he won by an excellent margin.[7] The negative sentiment score could be due to the aggressiveness of Trump supporters against the opposition candidate and supporters, while Biden focused on liberal political views that has been more inclusive to minority groups and promoted climate change policies. Nevertheless, Biden's campaign failed due to the far-right Trump supporters in that electoral state (North Dakota).

In Table 6, we observe that words such as 'pron', 'begging', 'sick', and 'china', led to a negative sentiment score which indicates that the user tried to be either aggressive or are using whataboutery to defend their candidate. Hence, this shows that Twitter has been used as a medium to impose political opinion rather than discussing a viewpoint.

In Table 3, we find that the LSTM model provides a good competition with BERT in terms of the performance accuracy on the IMDB training dataset; however, in the test case (Figure 7, 8, and 9), they perform differently. The difference in performance may be due to BERT's complex architecture and pre-existing knowledge gained by training from more than 300 million parameters from a large corpus; hence, having better semantic information relating to the context of the tweets for capturing the sentiment polarity.

The results from the BERT model in Figure 9 show that sentiment analysis via Twitter can provide a good framework for modeling election results. If we compare the BERT model results (Panel a) with actual results (Panel d), we find that the BERT model has been successful in distinguishing the Trump and Biden states and some contentious states. We note that the LSTM model could not fully capture the situation due to the large number of neutral sentiments (Figure 7 - Panel b), and hence it has performed poorly in Figure 9 (Panel b) when compared with actual results (Panel d). Several factors such as net presidential approval, GDP growth in the second quarter of the election year, and a "term" penalty for the incumbent party can help in improving the prediction. While social media such as Twitter can give insights into how people vote, it must be noted that a large percentage of voters do not express themselves in social media. The factors such as distribution of tweets in terms of count, language, location play a vital role which is evident from our results. We note that some of the previous models indicated that it would be tough for Biden to win the elections [62]. Moreover, the multi-factor Twitter analysis predicted Republican's (Trump) winning the elections [41]; however, our BERT model indicates that Biden had more chances of winning (Figure 9, Panel a).

Table 5 provides average sentiment and their actual result based on the BERT model, where either the polarity ratio [Biden/Trump, Trump/Biden] of less than 1.5 determines if the state will be contentious. A contentious state can indicate if the state will be a swing state. Looking at 11th November media reports about potential swing states (Figure 9, Panel c), we find that the BERT model (Figure 9, Panel a) provides accurate results in highlighting five out of the eight states as contentious (Arizona, Michigan, Wisconsin, Minnesota, Pennsylvania, North Carolina, Florida, and Georgia). The LSTM model (Figure 9, Panel b) gives good information about swing states, but we need to ignore the model as it is

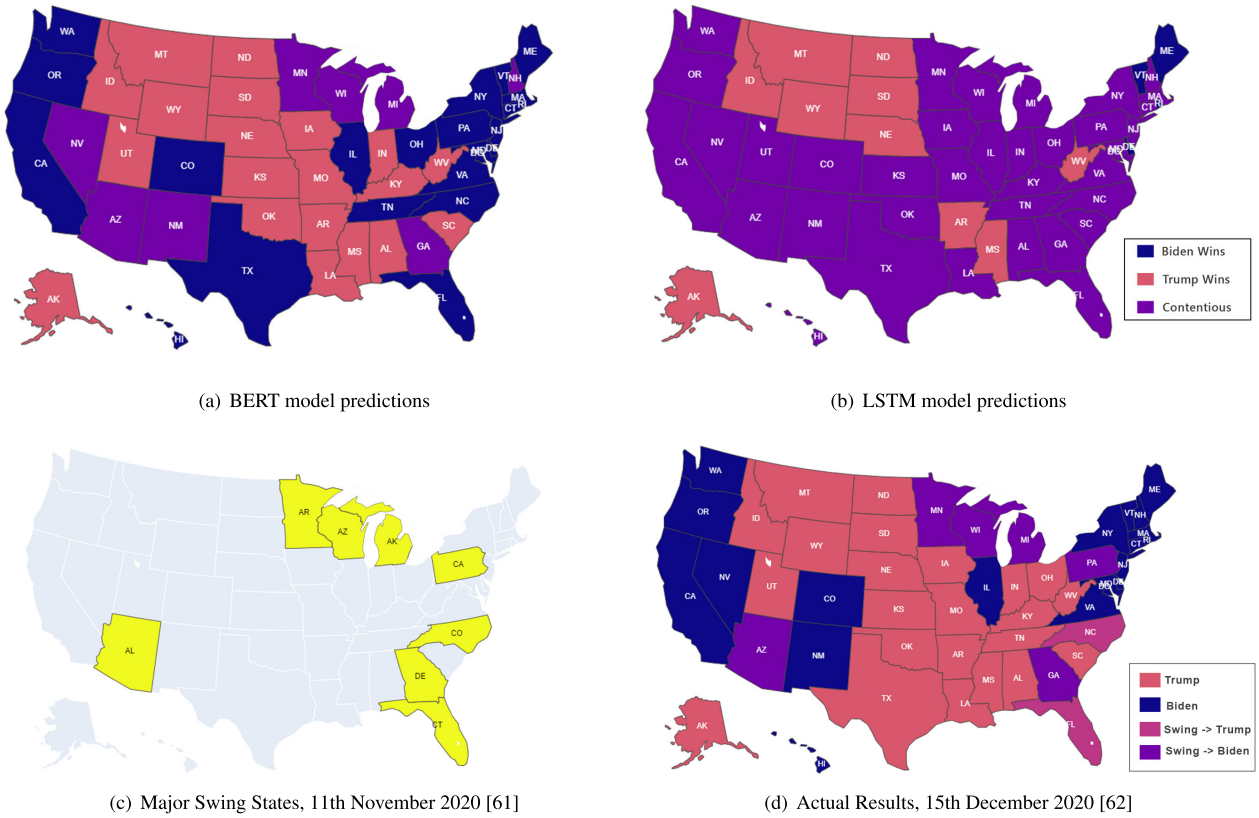
[7]Why Trump's 2020 dominance in North Dakota signals long road for state Democrats: https://www.thedickinsonpress.com/news/government-and-politics/6748856-Why-Trumps-2020-dominance-in-North-Dakota-signals-long-road-for-state-Democrats

**TABLE 4.** Prominent state's average sentiment and their actual result comparison based on BERT model Figure 7 (Panel a and b).

| Top 3 Positive State | | | | | | | |
|---|---|---|---|---|---|---|---|
| Donald Trump (Figure 7, Panel a) | | | | Joe Biden (Figure 7, Panel b) | | | |
| State | Trump Score | Biden Score | Actual Result | State | Trump Score | Biden Score | Actual Result |
| Wyoming | 0.106353 | 0.134895 | Trump 70.4% | Delaware | -0.01333 | 0.164065 | Biden 58.8% |
| New Hampshire | 0.100506 | 0.086244 | Biden 52.9% | North Dakota | -0.01333 | 0.163476 | Trump 65.5% |
| Kentucky | 0.074034 | 0.095935 | Trump 62.1% | Wyoming | 0.106353 | 0.134895 | Trump 70.4% |
| Top 3 Least Positive State | | | | | | | |
| Donald Trump (Figure 7, Panel a) | | | | Joe Biden (Figure 7, Panel b) | | | |
| State | Trump Score | Biden Score | Actual Result | State | Trump Score | Biden Score | Actual Result |
| Montana | -0.10017 | 0.097477 | Trump 56.9% | West Virginia | 0.003656 | 0.042460 | Trump 68.6% |
| Delaware | -0.01333 | 0.164065 | Biden 58.8% | Mississippi | 0.014113 | 0.04309 | Trump 57.5% |
| North Dakota | -0.00297 | 0.163476 | Trump 65.5% | Utah | 0.061679 | 0.050787 | Trump 58.1% |

**TABLE 5.** Swing state's average sentiment and their actual result comparison based on the BERT model. Note that either the polarity ratio [Biden/Trump, Trump/Biden] of less than 1.5 determines if the state will be contentious, which can give insights if the state will be a swing state.

| States | Trump Score | Biden Score | $B_{score}/T_{score}$ | $T_{score}/B_{score}$ | Our Result | Actual Result |
|---|---|---|---|---|---|---|
| Arizona | 0.053 | 0.069 | 1.307 | 0.765 | Contentious | Trump 49.1% |
| Michigan | 0.0573 | 0.077 | 1.355 | 0.738 | Contentious | Biden 50.6% |
| Wisconsin | 0.059 | 0.089 | 1.499 | 0.667 | Contentious | Biden 49.4% |
| Minnesota | 0.0428 | 0.063 | 1.483 | 0.674 | Contentious | Biden 52.4% |
| Pennsylvania | 0.029 | 0.088 | 2.973 | 0.336 | Biden | Biden 50.0% |
| North Carolina | 0.047 | 0.0942 | 2.018 | 0.495 | Biden | Trump 49.9% |
| Florida | 0.042 | 0.085 | 2.059 | 0.485 | Biden | Trump 51.2% |
| Georgia | 0.061 | 0.091 | 1.489 | 0.671 | Contentious | Biden 49.5% |




(a) BERT model predictions

(b) LSTM model predictions

(c) Major Swing States, 11th November 2020 [61]

(d) Actual Results, 15th December 2020 [62]

**FIGURE 9.** State-wise results for all electoral states.

deplorable when it comes to Biden state wins when compared to actual results (Figure 9, Panel d). Table 5 shows that in most of the cases, the BERT model correctly predicts the swing states as contentious. In some cases, the prediction turns out to be incorrect, such as Florida and North Carolina, where Trump won by a mere 1-2% margin. It is also evident

**TABLE 6.** Example of tweet scores in favour or against Trump and Biden.

| Biden Supporters | |
|---|---|
| Tokenized Tweet | Sentiment score |
| ['nypost', 'censorship', 'censored', 'twitter', 'manipulate', 'us', 'election', 'favor', 'joebiden', 'trump', 'ccp', 'china', 'porn', 'twitter', 'always', 'fine', 'jack', 'vijaya', 'dickc', 'katies', 'marciadorsey', 'jack', 'sick'] | -0.1488 |
| ['wrong', 'cory', 'booker', 'brilliant', 'final', 'questioning', 'trump', 'nominee', 'amy', 'coney', 'barrett', 'amyconeybarrett', 'corybooker', 'barrett', 'booker', 'trump', 'kamalaharris', 'joebiden', 'scotus', 'supremecourtconfirmation'] | 0.1333 |
| ['icecube', 'sellout', 'long', 'black', 'people', 'going', 'vote', 'democrats', 'nothing', 'us', 'look', 'democrats', 'really', 'mean', 'means', 'satan', 'joebiden', 'admitted', 'pedophilia', 'trump', 'tosaveamerica'] | -0.1763 |
| ['twitter', 'since', 'censoring', 'free', 'speech', 'regards', 'joebiden', 'email', 'scandal', 'opened', 'parler', 'app', 'account', 'done', 'without', 'motivation', 'joebiden', 'parler', 'censorship', 'twitter', 'conservative'] | 0.4001 |
| **Trump Supporters** | |
| Tokenized Tweet | Sentiment score |
| ['trump', 'student', 'used', 'hear', 'years', 'ten', 'years', 'heard', 'china', 'know', 'many', 'asked', 'many', 'said', 'sir', 'know', 'millions', 'like', 'million', 'um'] | 0.3333 |
| ['clady', 'minutes', 'long', 'time', 'ago', 'omarosa', 'never', 'represented', 'black', 'community', 'thereidout', 'cried', 'trump', 'begging', 'job'] | -0.1083 |
| ['meiselasb', 'wonder', 'drugs', 'trump', 'takes', 'masses', 'burgers', 'maybe', 'still', 'dexamethasone', 'high', 'currently', 'high', 'november', 'latest', 'low', 'trumpcovid', 'trumpisalaughingstock', 'trumpisaloser', 'votehimout', 'votebluetoendthisnightmare'] | 0.2050 |
| ['jimjordan', 'devinnunes', 'mattgaetz', 'johncornyn', 'billbarr', 'donaldtrump', 'amp', 'rest', 'railed', 'indicted', 'amp', 'go', 'jail', 'perpetuating', 'fraud', 'american', 'people', 'perfectly', 'legal', 'unmasking', 'general', 'flynn', 'make', 'sick', 'lockthemup'] | -0.1128 |

that this might be due to them being swing states from the 2016 elections. In Florida and North Carolina, the normalized final score ($p$) is above 2 in Table 5, which is close to the borderline vote count in actual results. Generally, it would be challenging for such models to predict the outcomes of contentious and swing states due to limited data. Only a tiny fraction of tweets have state information. We also need to note that not all voters would express themselves on Twitter.

## IV. DISCUSSION

The primary purpose of our study was to understand the nature of the political discourse that took place on Twitter during the elections, such as sentiments expressed, frequently mentioned terms, and popular tweets/retweets. We utilized user attributes such as tweet ID, retweet count, date of joining Twitter, user followers count and observe that the tweets' overall sentiment has been positive especially for Biden. In the case of Trump (Figure 7, Panel a), the sentiment polarity score has been negative in some states, indicating the nature of his campaigns that targeted global issues that promoted abuse.

According to our exploratory data analysis, we find that although Twitter is a popular tool for political discussions and debates, a minimal number of users dominate this platform. Figure 3 gives the comparison between the number of tweets with various geo-locations, where only 26.12% are within the US out of 1.17 million tweets. Hence, most users (who shared their geo-location) are simply following trends and discussions through tweets. It seems that most users who shared their geo-location from US origin have been passive and did not actively participate in conversations during the peak of the US election campaigns.

Modeling and forecasting electoral results only with tweets is a very challenging task. The US 2020 election was held during the COVID-19 pandemic with significant travel restrictions, and uncertainty in vaccination and economic activity [38]. There has been a significant rise in unemployment and geopolitical tensions, especially with China's trade apart from restricted migrations and the development of a border with and Mexico, given Trump's policies. These led to the polarising viewpoint in social media not just from US users, but from all the users worldwide, which has been covered by Google leading news coverage (top stories) [63]. The coverage of the US 2020 campaign and elections was dominant in international news, and hence there were massive tweets regarding the elections worldwide.

Our model has a major limitation where it only provides a prediction based only on a small section of the society that expresses themselves on social media about their political views. Our aim was to predict and provide a general viewpoint of the society based on sub-sampling from a population using novel language models powered by deep learning.

The BERT and LSTM models have been trained on the same dataset (IMDB dataset), and hence it is fair to compare their training performance. Although the framework can incorporate other models, in order to maintain a fair comparison we need to ensure that the other models use the same dataset and similar word embedding. We note that we use the basic BERT model known as BERT-base; however, in principle the framework can incorporate larger models such as BERT-large with fine tuning [64].

Furthermore, large pre-trained models would be more suitable when more data is available from the elections. Our framework uses BERT-base since it is publicly available and our dataset is not so large that it requires a larger model. Other models can be used to enhance the framework further. These include 1.) pre-trained models such as *embeddings from language models* (ELMO) [65] that use complex characteristics such as syntax and semantics in word embedding and 2.) word embeddings such as contextualized word vectors (CoVe) [66].

## V. CONCLUSION AND FUTURE WORK

Our study highlighted that discussion on the social media platform can be helpful in understanding crowd behavior and viewpoint during elections. We analyzed approximately 1.2 million tweets associated with the US 2020 presidential elections. After modeling and analyses, we found that sentiment analysis can form a general basis for modeling election outcomes. The BERT model indicated that Biden had a better chance of winning based on the tweets during the electoral campaigns. We find that the BERT model has been accurate in determining Trump, Biden, and the contentious states. Hence, given more data and geographical information, sentiment analysis could be helpful in predicting election results.

In future work, we can expand this study by detailed geolocation analyses, which can significantly increase the number of tweets for the given states. Furthermore, by dividing US states into rural and urban areas, we can further refine our location sentiment analyses as rural; urban divide plays a crucial role in elections. The framework can also be extended to other areas other than general elections, including smaller-scale elections involving cities and states. The framework can also be used to understand public viewpoints regarding emerging political issues such as COVID-19 travel restrictions, lock-downs, and vaccination strategies.

## DATA AND CODE

We provide Python-based open source code and data for further research.[8]

## AUTHOR CONTRIBUTIONS STATEMENT

Rohitash Chandra devised the project with the main conceptual ideas and experiments and contributed to overall writing, literature review and discussion of results. Ritij Saini provided implementation and experimentation and further contributed to results visualization and analysis. Rohitash Chandra and Ritij Saini contributed equally to this work.

[8]https://github.com/sydney-machine-learning/sentimentanalysis-USelections

## REFERENCES

[1] M. S. Lewis-Beck, "Election forecasting: Principles and practice," *Brit. J. Politics Int. Relations*, vol. 7, no. 2, pp. 145–164, May 2005.
[2] W. Ascher, "Political forecasting: The missing link," *J. Forecasting*, vol. 1, no. 3, pp. 227–239, Jul. 1982.
[3] K. L. Remmer, "The political economy of elections in Latin America, 1980–1991," *Amer. Political Sci. Rev.*, vol. 87, no. 2, pp. 393–407, 1993.
[4] S. Ansolabehere and J. M. Snyder, Jr., "The incumbency advantage in US elections: An analysis of state and federal offices, 1942–2000," *Election Law J.*, vol. 1, no. 3, pp. 315–338, 2002.
[5] G. Terhanian, J. Bremer, R. Smith, and R. Thomas, "Correcting data from online surveys for the effects of nonrandom selection and nonrandom assignment," Harris Interactive, Chicago, IL, USA, White Paper, 2000, pp. 1–13. [Online]. Available: https://innovativepublichealth.org/wp-content/uploads/Correcting_Data_Nonrandom.pdf
[6] B. K. Kaye and T. J. Johnson, "Research methodology: Taming the cyber frontier: Techniques for improving online surveys," *Social Sci. Comput. Rev.*, vol. 17, no. 3, pp. 323–337, 1999.
[7] V. Vehovar and K. L. Manfreda, "Overview: Online surveys," in *The SAGE Handbook of Online Research Methods*. Surrey, U.K., 2008, pp. 143–160. [Online]. Available: http://mr.crossref.org/iPage?doi=10.4135%2F9780857020055.n10
[8] J. Blasius and M. Brandt, "Representativeness in online surveys through stratified samples," *Bull. Sociol. Methodol./Bull. de Méthodologie Sociologique*, vol. 107, no. 1, pp. 5–21, Jul. 2010.
[9] G. Enli, "Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and clinton in the 2016 US presidential election," *Eur. J. Commun.*, vol. 32, no. 1, pp. 50–61, Feb. 2017.
[10] D. Lilleker, K. Koc-Michalska, and N. Jackson, "Social media in the UK election campaigns 2008–2014: Experimentation, innovation, and convergence," in *The Routledge Companion to Social Media and Politics*. New York, NY, USA: Routledge, 2015, pp. 325–337.
[11] E. Skogerbø and A. H. Krumsvik, "Newspapers, Facebook and Twitter: Intermedial agenda setting in local election campaigns," *Journalism Pract.*, vol. 9, no. 3, pp. 350–366, 2015.
[12] K. S. Jones, "Natural language processing: A historical review," in *Current Issues in Computational Linguistics: In Honour of Don Walker*. Cambridge, U.K., 1994, pp. 3–16. [Online]. Available: https://link.springer.com/chapter/10.1007/978-0-585-35958-8_1
[13] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," *IEEE Comput. Intell. Mag.*, vol. 9, no. 2, pp. 48–57, May 2014.
[14] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1532–1543. [Online]. Available: http://www.aclweb.org/anthology/D14-1162
[15] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan, "A system for real-time Twitter sentiment analysis of 2012 us presidential election cycle," in *Proc. ACL Syst. Demonstrations*, 2012, pp. 115–120.
[16] A. Severyn and A. Moschitti, "Twitter sentiment analysis with deep convolutional neural networks," in *Proc. 38th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2015, pp. 959–962.
[17] E. Kouloumpis, T. Wilson, and J. Moore, "Twitter sentiment analysis: The good the bad and the omg!," in *Proc. 5th Int. AAAI Conf. Weblogs Social Media*, vol. 5, no. 1, 2011, pp. 1–3. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/14185/14034
[18] A. Giachanou and F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," *ACM Comput. Surv.*, vol. 49, no. 2, pp. 1–41, Jun. 2016.
[19] A. Agarwal, D. Toshniwal, and J. Bedi, "Can Twitter help to predict outcome of 2019 Indian general election: A deep learning based study," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. India: Springer, 2019, pp. 38–53. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-43887-6_4
[20] A. Suciati, A. Wibisono, and P. Mursanto, "Twitter buzzer detection for Indonesian presidential election," in *Proc. 3rd Int. Conf. Informat. Comput. Sci. (ICICoS)*, Oct. 2019, pp. 1–5.
[21] K. K. Mohbey, "Multi-class approach for user behavior prediction using deep learning framework on Twitter election dataset," *J. Data, Inf. Manage.*, vol. 2, no. 1, pp. 1–14, Mar. 2020.

[22] P. Vijayaraghavan, S. Vosoughi, and D. Roy, "Automatic detection and categorization of election-related tweets," in *Proc. 10th Int. AAAI Conf. Web Social Media*, 2016, pp. 703–706.
[23] M. Li, E. Perrier, and C. Xu, "Deep hierarchical graph convolution for election prediction from geospatial census data," in *Proc. AAAI Conf. Artif. Intell.*, 2019, vol. 33, no. 1, pp. 647–654.
[24] C. Manning and H. Schutze, *Foundations of Statistical Natural Language Processing*. Cambridge, MA, USA: MIT Press, 1999.
[25] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in *Mining Text Data*. Chicago, IL, USA: Springer, 2012, pp. 415–463. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4614-3223-4_13
[26] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, 2014.
[27] D. M. El-Din Mohamed Hussein, "A survey on sentiment analysis challenges," *J. King Saud Univ.-Eng. Sci.*, vol. 30, no. 4, pp. 330–338, 2018.
[28] F. V. Ordenes, B. Theodoulidis, J. Burton, T. Gruber, and M. Zaki, "Analyzing customer experience feedback using text mining: A linguistics-based approach," *J. Service Res.*, vol. 17, no. 3, pp. 278–295, 2014.
[29] F. Greaves, D. Ramirez-Cano, C. Millett, A. Darzi, and L. Donaldson, "Use of sentiment analysis for capturing patient experience from free-text comments posted online," *J. Med. Internet Res.*, vol. 15, no. 11, p. e239, Nov. 2013.
[30] A. Mittal and A. Goel, "Stock prediction using Twitter sentiment analysis," Stanford Univ., Stanford, CA, USA, Tech. Rep. CS229 (2011), 2012, pp. 1–5.
[31] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2016, pp. 606–615.
[32] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2017, *arXiv:1706.03762*. [Online]. Available: http://arxiv.org/abs/1706.03762
[33] T. Wolf *et al.*, "Transformers: State-of-the-art natural language processing," in *Proc. Conf. Empirical Methods Natural Lang. Process., Syst. Demonstrations*, 2020, pp. 38–45.
[34] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*. [Online]. Available: http://arxiv.org/abs/1810.04805
[35] T. Wang, K. Lu, K. P. Chow, and Q. Zhu, "COVID-19 sensing: Negative sentiment analysis on social media in China via BERT model," *IEEE Access*, vol. 8, pp. 138162–138169, 2020.
[36] Z. Gao, A. Feng, X. Song, and X. Wu, "Target-dependent sentiment classification with BERT," *IEEE Access*, vol. 7, pp. 154290–154299, 2019.
[37] W. Yang, H. Zhang, and J. Lin, "Simple applications of BERT for ad hoc document retrieval," 2019, *arXiv:1903.10972*. [Online]. Available: http://arxiv.org/abs/1903.10972
[38] L. Baccini, A. Brodeur, and S. Weymouth, "The COVID-19 pandemic and the 2020 US presidential election," *J. Population Econ.*, vol. 34, no. 2, pp. 739–767, Apr. 2021.
[39] Hitkul, A. Prabhu, D. Guhathakurta, J. jain, M. Subramanian, M. Reddy, S. Sehgal, T. Karandikar, A. Gulati, U. Arora, R. Ratn Shah, and P. Kumaraguru, "Capitol (Pat)riots: A comparative study of Twitter and parler," 2021, *arXiv:2101.06914*. [Online]. Available: http://arxiv.org/abs/2101.06914
[40] L. Xiguang and W. Jing, "Web-based public diplomacy: The role of social media in the Iranian and Xinjiang riots," *J. Int. Commun.*, vol. 16, no. 1, pp. 7–22, 2010.
[41] I. Sabuncu, M. Ali Balci, and O. Akguller, "Prediction of USA November 2020 election results using multifactor Twitter data analysis method," 2020, *arXiv:2010.15938*. [Online]. Available: http://arxiv.org/abs/2010.15938
[42] I. Sabunchu, *USA Nov. 2020 Election 20 Mil. Tweets (With Sentiment and Party Name Labels) Dataset, IEEE Dataport*, 2020. [Online]. Available: https://dx.doi.org/10.21227/25te-j338
[43] M. Hui, *US Election 2020 Tweets, IEEE Dataport*, 2020. [Online]. Available: https://www.kaggle.com/manchunhui/us-election-2020-tweets
[44] L. Cram, C. Llewellyn, R. Hill, and W. Magdy, "UK general election 2017: A Twitter analysis," 2017, *arXiv:1706.02271*. [Online]. Available: http://arxiv.org/abs/1706.02271
[45] R. Chandra and A. Krishna, "COVID-19 sentiment analysis via deep learning during the rise of novel cases," *PLoS ONE*, vol. 16, no. 8, Aug. 2021, Art. no. e0255615, doi: 10.1371/journal.pone.0255615.
[46] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proc. Conf. Human Lang. Technol. Empirical Methods Natural Lang. Process. (HLT)*, 2005, pp. 347–354.
[47] A. Devitt and K. Ahmad, "Sentiment polarity identification in financial news: A cohesion-based approach," in *Proc. 45th Annu. Meeting Assoc. Comput. Linguistics*, 2007, pp. 984–991.
[48] A. Montejo-Ráez, E. Martínez-Cámara, M. T. Martín-Valdivia, and L. A. Ureña-López, "Ranked WordNet graph for sentiment polarity classification in Twitter," *Comput. Speech Lang.*, vol. 28, no. 1, pp. 93–107, Jan. 2014.
[49] Z. Waseem and D. Hovy, "Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter," in *Proc. NAACL Student Res. Workshop*, 2016, pp. 88–93.
[50] H. Watanabe, M. Bouazizi, and T. Ohtsuki, "Hate speech on Twitter: A pragmatic approach to collect hateful and offensive expressions and perform hate speech detection," *IEEE Access*, vol. 6, pp. 13825–13835, 2018.
[51] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," 2013, *arXiv:1310.4546*. [Online]. Available: http://arxiv.org/abs/1310.4546
[52] L. De Vine, G. Zuccon, B. Koopman, L. Sitbon, and P. Bruza, "Medical semantic similarity with a neural language model," in *Proc. 23rd ACM Int. Conf. Conf. Inf. Knowl. Manage.*, Nov. 2014, pp. 1819–1822.
[53] V. Vargas-Calderón and J. E. Camargo, "Characterization of citizens using Word2vec and latent topic analysis in a large set of tweets," *Cities*, vol. 92, pp. 187–196, Sep. 2019.
[54] S. Yilmaz and S. Toklu, "A deep learning analysis on question classification task using Word2vec representations," *Neural Comput. Appl.*, vol. 32, pp. 2909–2928, Jan. 2020.
[55] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning word vectors for sentiment analysis," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics: Hum. Lang. Technol.* Portland, OR, USA: Association for Computational Linguistics, Jun. 2011, pp. 142–150. [Online]. Available: http://www.aclweb.org/anthology/P11-1015
[56] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[57] N. Shuyo. (2010). *Language Detection Library for Java*. [Online]. Available: https://code.google.com/archive/p/language-detection/
[58] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. 8th Int. AAAI Conf. Web Social Media*, E. Adar, P. Resnick, M. D. Choudhury, B. Hogan, and A. H. Oh, Eds. Atlanta, GA, USA: AAAI Press, 2014, pp. 216–225. [Online]. Available: http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf
[59] D. W. Beachler *et al.*, *Presidential Swing States: Why Only Ten Matter*. Lexington Books, 2015.
[60] T. Seymat. (2020). *US Election: These Eight Battleground States Will Decide Whether Trump or Biden Wins*. Euronews. https://www.euronews.com/2020/11/03/us-election-these-eight-battlegroun d-states-will-decide-whether-trump-or-biden-wins
[61] C. Canipe, A. J. Levine, and S. Hart. (2020). *U.S. Election Results, Reuters*. [Online]. Available: https://graphics.reuters.com/USA-ELECTION/RESULTS-LIVE-US/jbyprxelqpe/15/12/2020
[62] B. Jérôme, V. Jérôme, P. Mongrain, and R. Nadeau, "State-level forecasts for the 2020 us presidential election: Tough victory ahead for biden," *PS: Political Sci. Politics*, vol. 51, no. 1, pp. 1–4, 2020.
[63] A. Kawakami, K. Umarova, and E. Mustafaraj, "The media coverage of the 2020 US presidential election candidates through the lens of Google's top stories," in *Proc. Int. AAAI Conf. Web Social Media*, vol. 14, Wellesley, MA, USA, 2020, pp. 868–877. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/7352
[64] C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to fine-tune BERT for text classification?" in *Proc. China Nat. Conf. Chin. Comput. Linguistics*. Springer, 2019, pp. 194–206.
[65] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," 2018, *arXiv:1802.05365*. [Online]. Available: http://arxiv.org/abs/1802.05365
[66] B. McCann, J. Bradbury, C. Xiong, and R. Socher, "Learned in translation: Contextualized word vectors," 2017, *arXiv:1708.00107*. [Online]. Available: http://arxiv.org/abs/1708.00107