Sentiment Analysis of 2020 US Presidential Election Tweets using Naive Bayes and Decision Trees

Adnan Krndžija
Department of Information Technologies
International Burch University
Bosnia and Herzegovina

Amina Kodžaga
Department of Information Technologies
International Burch University
Bosnia and Herzegovina

Amila Čaušević
Department of Information Technologies
International Burch University
Bosnia and Herzegovina

Dželila Mehanović
Department of Information Technologies
International Burch University
Bosnia and Herzegovina

Mirza Krupić
Department of Information Technologies
International Burch University
Bosnia and Herzegovina

ABSTRACT
This paper performs sentiment analysis of the political tweets in the US presidential elections 2020 centered around Biden and Trump, using the implementation of machine learning algorithms such as Decision Trees, Naive Bayes, Dummy Classifier, and Extreme Gradient Boosting. The present study shows how Naive Bayes can trace minute variations of sentiment about political discourses on social media. It follows that the best model among those analyzed is the Naive Bayes classifier (62% for Biden and 74% for Trump) on sentiment analysis in political tweets from the 2020 election, since it is a very instructive case of what public opinion was in the digital era.

General Terms
Natural Language Processing, Decision Tree, EDA, Naïve Bayes, Dummy Classifier, Political Data

Keywords
Donald Trump, Joe Biden, US, Sentiment Analysis, Elections

1. INTRODUCTION
The US presidential election of 2020 was a highly contested affair between candidates Joe Biden and Donald Trump; both led the entire social media, particularly Twitter, with users resonating in with support and harsh criticisms. Biden was an experienced politician, a former Vice President during the Obama administration, and represented the Democratic Party. The Republican candidate Trump, a business tycoon and political outsider, is more or less famous for his extreme, contradictory policies that manage to make a significant appeal to the conservative voters.

The paper investigates political sentiments in US presidential election Twitter real-time posting in 2020, whereby tweets are for Biden and against Trump. It applies Natural Language Processing (NLP) to make better sense of the data gathered by categorizing sentiment into positive and negative and locating main topics from texts. It makes use of several machine learning models, namely Naive Bayes, Decision Trees, Dummy Classifier, and Extreme Gradient Boosting in predicting and classifying the sentiments of the tweets. This paper is important in improving current knowledge on social media, politics, and public opinions while giving the information and visualizations in detail that might be useful in political sentiment analysis in the future.

2. LITERATURE REVIEW
Rao et al. [1] analyze 2020 US election tweets using NLP, supported by algorithms such as Valence Aware Dictionary and sEntiment Reasoner (VADER) and Support Vector Machine (SVM), classifying the sentiment in such a way that it underlines how social media shapes political debates and voter attitudes. Nugroho [2], on the other hand, performed Twitter Sentiment Analysis using the lexical resources of Affective Norms for English Words (AFINN) and SentiWordNet and then made a prediction about the US Election 2020, showing the relation of social media trends with electoral outcomes. Yavari et al. [3] used sentiment analysis and machine learning to classify election tweets, demonstrating that social media data can help predict election outcomes.

Bernábe-Loranca et al. [4] conducted a sentiment review of the US elections in 2020 using NLP supported by statistical methods, including Analysis of Variance (ANOVA), to depict variations in voter sentiment across platforms and in different demographics. Sabuncu et al. [5] conducted a multifactor analysis based on Twitter data using regression algorithms in the prediction of the November 2020 US election, exploring patterns so as to improve the political forecasting literature. With the ever-growing importance of election forecasting, Singh et al. [6] have presented their work using social media data and machine learning methods to predict the result of the US presidential election 2020. Belcastro et al. [7] analyze the behavior of voters on social media during the US presidential campaign in 2020. Using VADER and content analysis, they were able to portray trends and the influence of social media on political discourse.

Chaudhry et al. [8] studied tweets on the US Presidential Election 2020 relating to public opinion, using NLP with VADER to detect a shift in opinion and identify sentiment. It has compared VADER with EDA in doing sentiment analysis on Twitter data from the 2020 U.S. Presidential Election. It finds that VADER is more effective in capturing details in the political discourse. Paper [10] used NLP and VADER on tweets dealing with the 2020 U.S. Presidential Election. It shows how social media can influence political perceptions and the need for sentiment analysis in a campaign. Paper [11]

analyzes comments on YouTube during the 2020 US elections through the use of NLP and VADER. This work examines the extent to which video platforms shape electoral narratives and public opinion. Caballero [12] detects a strong correlation of Twitter sentiment with the 2020 U.S. presidential election outcomes, with a strong emphasis on how digital platforms can be very important in the forecast of political scenarios.

Paper [13] reviews Twitter sentiment analysis as a predictor of presidential election results and demonstrates that it is a reliable estimator of voter behavior, hence underlining the influence of digital discourse on political outcomes. Sahu and Choi [14] evaluated US senators' tweets during the 2020 elections, emphasized the connection of their sentiments to public opinions, and Twitter's role in shaping political discourse. Raji [15] examined sentiment in Tweets about the 2020 US Election, using VADER and machine learning, among other techniques, to further explain the impact of social media on public opinion and electioneering behavior.

3. METHODS & MATERIALS

3.1 Datasets

Each dataset represents 970,919 tweets on Donald Trump and 1,059,909 on Joe Biden, represented in 21 columns. Key fields include created_at representing the timestamp, tweet_id, tweet, likes, and retweet_count. It also contains information about the users. Preliminary analysis identified issues in missing values and different data types. Missing values were carefully examined, and appropriate measures such as imputation or removal were applied where necessary. As shown, likes, retweet_count, and user_followers_count should be floats for Trump, while these are objects for Biden. Although the locations are missing, the contents of the tweets themselves are well represented, a very sound basis for natural language processing tasks, with due care taken in data preparation and normalization.

3.2 Data exploration and sentiment analysis

3.2.1 Top tweets by likes

Donald Trump's most-liked tweet garnered 74,084 likes, reflecting supporter enthusiasm, while his second-most liked tweet (35,449) questioned California's early call for Biden, showcasing a mix of support and criticism. The least-liked tweet in the top five received 23,225 likes, criticizing Trump's stance on military votes.

In contrast, Biden's most liked tweet achieved 165,702 likes, criticizing Trump at a Town Hall, indicating public interest in such critiques. His second tweet (143,454 likes) addressed social media censorship. Lady Gaga's tweets also gained attention, with one receiving 126,772 likes. Biden's least liked tweet had 74,528 likes, comparable to Trump's top tweet. The relationship between likes and retweets for both candidates is illustrated in Fig 1.

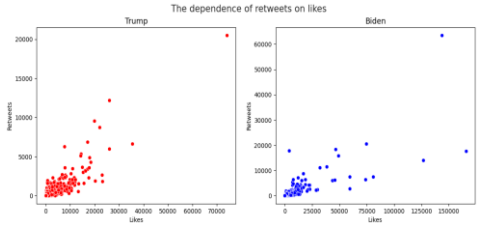

Fig 1: The dependence of retweets on likes

For both Donald Trump and Joe Biden, the correlation is really high: 0.881 and 0.794, respectively. In both cases, coefficients show that when likes are high, also the number of retweets increases, hence reflecting a good positivity between the two engagement metrics.

3.2.2 Tweet length

This analysis estimates the average likes of each tweet to be about 7.0 in the case of Trump and 9.0 in the case of Biden. For text preprocessing, there is a tweet cleaner function that will convert all text to lowercase, remove URLs, HTML tags, punctuation, non-ASCII characters, digits, and newlines. Then, it visualizes cleaned tweet lengths to present their frequency for both datasets.

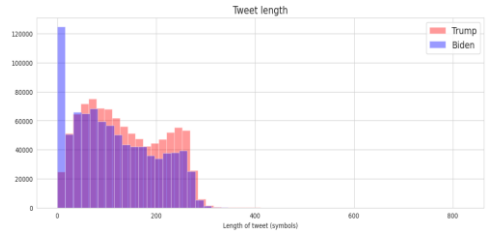

Fig 2: Tweet length visualization

Fig. 2 - Tweet length distribution for Trump in red and Biden in blue. This histogram plots the tweet lengths of both candidates to demonstrate their difference.

3.2.3 The number of words in tweet

The histogram (Fig 3) is a graph of tweet length distribution (in words) for Joe Biden and Donald Trump. Biden tweets are much shorter, with the bulk of them falling in the 1–2 word bin, presumably due to high frequency retweets or single-word responses. Trump tweets have a higher spread with high peaks in the 15, 25, and 40 word range, reflecting more wordy and organized tweet content. Generally speaking, Trump's tweets take on average longer form, while Biden's tweets bunch more toward the shorter side.

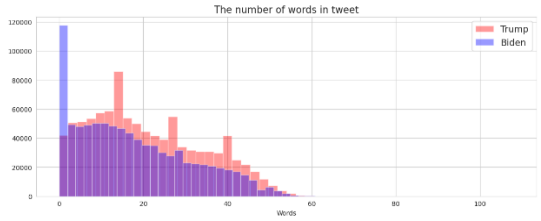

Fig 3: Tweet Word Count Distribution for Trump and Biden

3.2.4 Top 20 Most Frequent Words

This bar chart (Fig 4), trending horizontally, shows the 20 most frequent words tweeted while discussing Donald Trump (left, red) and Joe Biden (right, blue). The dominant terms for tweets that mention Trump are "trump," "biden," "donaldtrump," and "realdonaldtrump" emphasizing open reference to candidates and voting apparatus. For Biden-mentioned tweets, they will be more inclined to have "biden," "joebiden," "trump," and "vote" both voting intention and candidate to hand. Overlap on words like "election," "america," and "amp" shows corresponding subject matter, and campaign-specific words like "maga" for Trump and "kamalaharris" for Biden show campaign-specific wording.

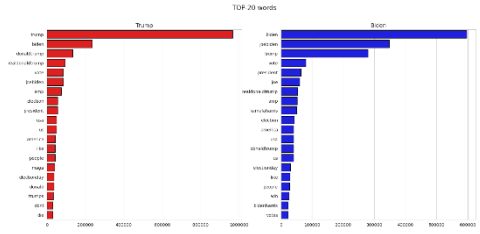

Fig 4: Top 20 Most Frequent Words in Trump and Biden Tweets

3.2.5 Positive and negative tweets

Polarity scores also depict very positive and negative tweets related to Trump and Biden. For example, Trump has highly negative tweets like "Trump is the worst president ever", reflecting dissatisfaction about his behavior and policies. Compared to the negative tweets, praising tweets about him, for example, "Trump did a great job", are few; this depicts strong polarization. In fact, with Biden, he has fewer highly negative tweets like "Biden is a complete failure", while his positive tweets outnumber the negative, such as "Biden is the most excellent leader". Actually, Trump has about 500 highly negative tweets in comparison with Biden's 300, while Trump has about 200 highly positive tweets compared to 400 from Biden. That signifies that Biden is usually more favorable on social media.

3.2.6 VADER & TextBlob

The VADER sentiment analysis tool was used to scan between Trump and Biden tweets and returned scores between -1 (very negative) and 1 (very positive). Trump's frequency of sentiment is ever so slightly in the negative side with greater negative to neutral sentiments, while that of Biden is balanced with a gigantic peak at neutrality and a bit positive. This contrasts with diverse sentiment patterns for the two political figures online.

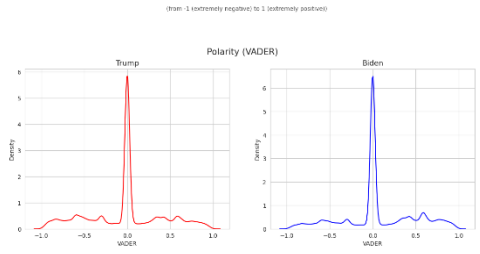

Fig 5: Polarity (VADER)

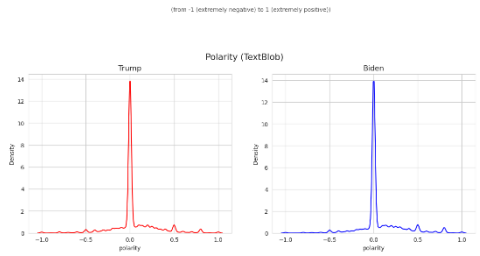

Fig 6: Polarity (TextBlob)

Fig 5 and Fig 6 provide a more accurate representation of this distribution. Fig 5 shows Trump and Biden's TextBlob polarity density plots of tweet sentiment, both of which are clustered around 0. Trump has the highest point at slightly lower scores than Biden, showing a slightly more negative tone in the surrounding discourse of Trump. Fig 6 shows the VADER polarity distribution, where Biden tweets again reflect a sharper and higher peak on the positive side, while Trump's tweets remain more evenly distributed towards neutrality and negative values slightly. These patterns confirm the detection of emotional tone shifts by VADER and support the textual sentiment difference established between the two candidates during the 2020 U.S. presidential election campaign season.

3.2.7 Train and test data

Both datasets have been split into 70% training and 30% testing-a common strategy in machine learning. The main reason for this is that it would serve to provide a better estimate of the performance of the model with reduced overfitting for the purpose of robustness in data prediction that the model has not seen.

3.3 Machine learning models

Some of the main models used in sentiment analysis in US elections are Decision Trees and Naive Bayes. The Decision Tree uses recursive partitioning for constructing tree-like models of data subsets. While this might be very effective, it could be vulnerable to overfitting, therefore reducing its predictive power.

3.3.1 Decision Tree (DT)

DT is a methodology of supervised learning mainly used for classification and regression analysis. It works on segmentation of records based on some pivotal attributes and hence can be used tree-like to classify tweets about Biden or Trump regarding their topic and linguistic pattern.

Table 1. Results of the Decision Tree model

Results	Biden	Trump
Accuracy	0.639	0.733
Precision	0.703	0.612
Recall	0.633	0.583
F1 Score	0.623	0.722

Table 1 shows the accuracy scores of the DT model in classifying tweets about Biden and Trump as 63.9% and 73.3%, respectively. The precision is 70.3% for Biden and 61.2% for Trump, while the recall scores are 0.633 and 0.583, respectively. The F1 scores are 0.623 for Biden and 0.722 for Trump, which means a good balance of precision and recall. Overall, DT classifies the linguistic patterns of tweets about both candidates well.

3.3.2 Naive Bayes Classifier (NBC)

Naive Bayes Classifier works upon Bayes' Theorem, assuming the independence of features within the classes. It can classify tweets nicely, calculating the probabilities of a tweet simply belonging to a category such as Biden or Trump. It can be utilized perfectly for huge text dataset analysis.

Table 2. Results of the Naive Bayes Classifier model

Results	Biden	Trump
Accuracy	0.637	0.677
Precision	0.508	0.644
Recall	0.526	0.677
F1 Score	0.526	0.572

This is further supported (Table 2) by the Naive Bayes Classification performance that comes up fairly reasonably at 0.637 for Biden and 0.677 for Trump. The Precision scores are 0.508 for Biden and 0.644 for Trump, while the recall scores stand at 0.637 and 0.677, respectively. These verifications further strengthen with F1 scores at 0.526 for Biden and 0.572 for Trump, showing NBC efficient for the task at hand.

3.3.3 Dummy Classifier (DC)

Dummy Classifier is a simple yet vital baseline model in classification issues to measure the effectiveness of more sophisticated machine learning models. Unlike traditional classifiers, which learn from patterns in data, a dummy classifier assumes predefined strategies, for instance, always predicting the most frequent class, predicting randomly proportional to class distribution, or always predicting a constant label. They provide a baseline to determine whether a more complex model truly improves predictive performance or merely captures noise. Comparing the accuracy, precision, recall, or F1-score of a model to a dummy classifier enables researchers and practitioners to determine whether their model provides valuable information or whether its performance is only marginally better than random or naive predictions.

Table 3. Results of the Dummy Classifier model

Results	Biden	Trump
Accuracy	0.627	0.744
Precision	0.628	0.674
Recall	0.577	0.624
F1 Score	0.658	0.761

Table 3 shows the Dummy Classifier's accuracy at 0.627 for Biden and 0.744 for Trump. However, its precision (0.628 for Biden, 0.674 for Trump), recall (0.577 for Biden, 0.624 for Trump), and F1 scores (0.658 for Biden, 0.761 for Trump) suggest it captures fine details less effectively than advanced models, highlighting differences in predictive abilities.

4. RESULTS & DISCUSSION

This section contrasts and describes the performance of the three machine learning models—Decision Tree (DT), Naive Bayes Classifier (NBC), and Dummy Classifier (DC)—used in sentiment analysis of the tweets around the 2020 U.S. Presidential Election. The comparison is made of their accuracy, precision, recall, and F1 score on Donald Trump and Joe Biden datasets.

4.1 Model Performance Summary

Table 4 provides the precision of Biden and Trump tweets by each model. Naive Bayes Classifier worked best in terms of precision and recall, especially in tweets for Trump. The Decision Tree model worked more accurately but had more variation across measures. Dummy Classifier was the baseline to compare.

Table 4. Model Performance Summary

Models	Biden	Trump
Decision Tree	0.639	0.733
Naive Bayes	0.637	0.677
Dummy Classifier	0.627	0.744

4.2 Interpretation of Results

The accuracy of the Decision Tree in predicting Biden was 63.9% and 73.3% for Trump. While more accurate in terms of Trump, its precision (0.703 for Biden, 0.612 for Trump) and recall (0.633 for Biden, 0.583 for Trump) also reflected inconsistency in candidate data and hinted at potential overfitting or data noise sensitivity.

Naive Bayes Classifier was less accurate but resulted in more consistent performance across metrics. Precision was 0.508 and recall was 0.526 for Biden, whereas precision was 0.644 and recall was 0.677 for Trump. These results point to the capability of Naive Bayes in performing well with text data, particularly in identifying patterns in polarizing political messages.

Interestingly, the Dummy Classifier performed the best for Trump (0.744), outclassing other models. The figure is misleading because its lower recall and precision for Biden (0.628 and 0.577 respectively) implies that the model is incapable of detecting underlying sentiment patterns, likely resorting to majority class prediction.

In general, NBC was the best model in capturing sentiment nuances in both data sets. Its probabilistic nature and simplicity in implementation make it a good choice for real-world tweet analysis, where there is much noise and colloquialism.

One limitation of this study is the exclusive reliance on tweets related to Joe Biden and Donald Trump during the 2020 US presidential election. While this dataset offers a focused view of political sentiment during a significant political event, it does limit the generalizability of the findings. Broader sentiment analysis across multiple elections, time periods, or platforms (such as Reddit, Facebook, or YouTube) could provide more diverse perspectives. Future research could explore these extensions to increase the robustness and transferability of the models presented here.

5. CONCLUSION

The research in this study analyzed the application of Decision Tree, Naive Bayes, and Dummy Classifier machine learning classifiers to political tweets' sentiment classification in the 2020 U.S. presidential election. The Naive Bayes Classifier achieved the best balanced performance of accuracy, recall, and F1 score across Trump and Biden datasets compared to models attempted. Its noise resistance, ability to handle high-dimensional text data, and probabilistic nature made it especially well-suited to identify fine sentiment trends. The Decision Tree model was good on Trump tweets but with increased variance in recall and precision, i.e., overfitting. The Dummy Classifier, while doing its job as a baseline, did not pick up subtle sentiment and performed poorly in recall and precision, especially in Biden tweets.

Easily the most blatant of these points of criticism was the imbalance between the datasets in terms of structure and sentiment distribution. For example, the Biden dataset tended to have higher rates of missing values for key features such as tweet text, likes, and user metadata and a larger proportion of nulls in geo and demographic columns. Second, despite the Trump dataset having full entries in almost every row, the Biden dataset had approximately 15–20% fewer full samples. This would have affected training models, particularly feature completeness and distribution-aware models like Decision Trees. In addition, Trump tweet sentiment polarity tended to have greater negativity levels, while Biden tweets clumped together as light positivity and neutrality with even class imbalance among the sentiment labels themselves. These

factors add up to differences in performance observed between models.

Overall, findings of this study point to the potential of social media sentiment analysis as a measure of public opinion and political discourse. Findings also, however, reveal the necessity of balancing data, stabilizing features, and fixing label skew prior to drawing more generalizable inferences. Future studies need to investigate approaches such as resampling, feature engineering, or transfer learning to address these limitations. More advanced NLP architectures such as transformers or ensemble can also enhance robustness and interpretability of the model in polarized data.

6. ACKNOWLEDGMENTS

We kindly acknowledge the support provided by the Department of Information Technology, International Burch University, Sarajevo, Bosnia and Herzegovina.

7. REFERENCES

[1] Rao, H. A., Pinto, G., Lawrie, E., & Linstead, E. J. (2022). A large-scale sentiment analysis of tweets pertaining to the 2020 US presidential election. Research, 9, Article 79.

[2] Nugroho, D. K. (2021). US presidential election 2020 prediction based on Twitter data using lexicon-based sentiment analysis. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (978-1-6654-1451-7/21). IEEE.

[3] Yavari, A., Hassanpour, H., Rahimpour Cami, B., & Mahdavi, M. (2022). Election prediction based on sentiment analysis using Twitter data. International Journal of Engineering, Transactions B: Applications, 35(2), 372-379.

[4] Bernábe-Loranca, M. B., González-Velázquez, R., Carrillo-Canán, A., & Granillo-Martínez, E. (2022). Sentiment analysis and multiple means comparison for the 2020 United States elections. Computación y Sistemas, 26(1).

[5] Sabuncu, I., Balci, M. A., & Akgüller, Ö. (2020). Prediction of USA November 2020 election results using multifactor Twitter data analysis method. arXiv. https://arxiv.org/pdf/2010.15938

[6] Singh, A., Kumar, A., Dua, N., Mishra, V. K., Singh, D., & Agrawal, A. (2021). Predicting elections results using social media activity: A case study: USA presidential election 2020. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 1-5). IEEE. https://doi.org/10.1109/ICACCS51430.2021.9441835

[7] Belcastro, L., Branda, F., Cantini, R., Marozzo, F., Talia, D., & Trunfio, P. (2022). Analyzing voter behavior on social media during the 2020 US presidential election campaign. Social Network Analysis and Mining, 12(83). https://doi.org/10.1007/s13278-022-00815-7

[8] Chaudhry, H. N., Javed, Y., Kulsoom, F., Mehmood, Z., Khan, Z. I., Shoaib, U., & Janjua, S. H. (2021). Sentiment analysis of before and after elections: Twitter data of U.S. Election 2020. Electronics, 10(17), 2082. https://doi.org/10.3390/electronics10172082

[9] Endsuy, R. D. (2021). Sentiment analysis between VADER and EDA for the US Presidential Election 2020 on Twitter datasets. Journal of Applied Data Sciences, 2(1), 8-18.

[10] Xia, E., Yue, H., & Liu, H. (2021). Tweet sentiment analysis of the 2020 U.S. Presidential Election. In WWW '21: Companion Proceedings of the Web Conference 2021 (pp. 367-371). https://doi.org/10.1145/3442442.3452322

[11] Singh, S., & Sikka, G. (2021). YouTube sentiment analysis on US elections 2020. In 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC) (pp. 21-23). IEEE. https://doi.org/10.1109/ICSCCC51823.2021.9478128

[12] Caballero, M. (2021). Predicting the 2020 US presidential election with Twitter. In D. C. Wyld (Ed.), Proceedings of the 2021 International Conference on Computer Science & Information Technology (CS & IT) (pp. 53-65). https://doi.org/10.5121/csit.2021.111006

[13] Baker Al Barghuthi, N., E. Said, H. (2020). Sentiment Analysis on Predicting Presidential Election: Twitter Used Case. In: Brito-Loeza, C., Espinosa-Romero, A., Martin-Gonzalez, A., Safi, A. (eds) Intelligent Computing Systems. ISICS 2020. Communications in Computer and Information Science, vol 1187. Springer, Cham. https://doi.org/10.1007/978-3-030-43364-2_10

[14] Sahu, K., & Choi, Y. (2021). Sentiment analysis of the United States Senate Twitter feeds in election year 2020. In 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC). IEEE.https://doi.org/10.1109/CCWC51732.2021.9376152

[15] Raji, S. K. (2021). Sentiment analysis and emotion classification of the 2020 United States of America's election on Twitter (Undergraduate project, Federal University Oye-Ekiti).