**Chapter Two**

**Introduction to Literature Review**

Interest in mining sentiment and opinion in texts has risen steadily over the years (Greenwald, 2009), mainly because of the increased availability of information and personal opinion messages (Shin, 2017). In general, sentiment analysis is used to make predictions or measures in various fields such as the stock market, politics, and even social movements (Karol, 2007). Old studies of politics on social networks have ended up lacking or even lacking data because they can only take a small sample (Kiousis, 2000). For example (Cerezo, 2006), proved that sentiment analysis on political prediction can proved accurate on the 2009 German election, On the other hand (Han, 2012), failed to predict 2011 US Presidential election ranking.

**The Use of Social Media in Elections**

Twitter and other social media platforms play an important role in elections and other democratic conversations. Many studies examined the use of social media in elections. In an analysis of the 2012 US presidential candidates’ Facebook pages, (Bronstein, 2012) show that the medium was mainly used for mobilization of supporters. Candidates retain control of their message, posting information only on a small number of non-controversial topics. For the same election (Mascaro, 2012), studied the conversational features of the tweets and concluded that there was limited interaction between users. They describe the interaction as a one-sided conversation which elicits no response. The medium was mainly used for broadcasting of messages.

In a study of the 2013 Norwegian national election, (Kalsnes, 2016) examine the disparity between parties’ interaction strategy and online responsiveness. Findings from this research described the risk to negative online reputation and associated media attention as significant factors affecting parties’ interaction with voters online. However, (Graham, 2010) indicate evidence of interaction between candidates and voters in their study which compares how British and Dutch parliamentary candidates used Twitter during the 2010 general election. They show that Dutch candidates interact (@reply accounting for 47% of their tweets) with others more significantly than British candidates (@reply accounting for 32% of their tweets). For the 2014 Indian general election, (Jaidka, 2014) studied Twitter accounts of the top ten political parties. They used social media for self-promotion, mobilization of voters and posting real-time updates of their offline campaign activities. In a study of how social media was used during general elections in Nigeria and Liberia in 2011, (Smyth, 2013) find that social media was used to report problems at election units and provide updates about the election process, concluding that it helped to overcome scarcity of information.

The effect of social media use by candidates has also been considered. (Effing, 2011) studied the impact of social media on the Dutch election and found that although it did not influence voting behaviour in the 2010 local election, politicians with higher social media engagement received more votes within most political parties during the national election. Similarly, (Kruikemeier, 2014) investigated candidates’ online campaign styles during the 2010 Dutch national election and showed that candidates who used Twitter received more votes than those who did not.

In terms of using social media to predict election’s outcomes, (Tumasjan, 2010) use Linguistic Inquiry and Word Count (LIWC) text analysis software to analyse the content of messages mentioning political parties and candidates during the German federal election. Their result shows that online messages on Twitter closely mirror the offline political sentiments. In the same context. (DiGrazia, 2013) analyse tweets mentioning candidates for the U.S House of Representatives election and showed that there was a significant correlation between the candidates’ electoral performance and tweets mentioning the candidates. Looking at developing nations, Twitter users’ sentiment was used to forecast election outcomes in India, Pakistan, Indonesia and others (Kagan, 2015), (Ramadhan, 2014). Using Indonesia’s presidential election as a case study, (Prasetyo, 2015) claimed that Twitter forecast outperformed all the traditional polls at the national level.

In general, the majority of work in this area follows the same approach, combining Twitter sentiment analysis, the volume of tweets or mentions of a candidate and correlated these with electoral result (Prasetyo, 2015), (Ramadhan, 2014), (Tumasjan, 2010). In a closely related work, (Fink, 2011) studied a corpus of tweets during the 2011 Nigerian presidential election. They found that tweets’ sentiment was less accurate in predicting the support for the two major candidates but counts of tweets mentioning the candidates correlated with election results across the country’s six geopolitical zones. They mentioned the resolution of the users’ location as a challenging task and suggested analysis at the state level in the future.

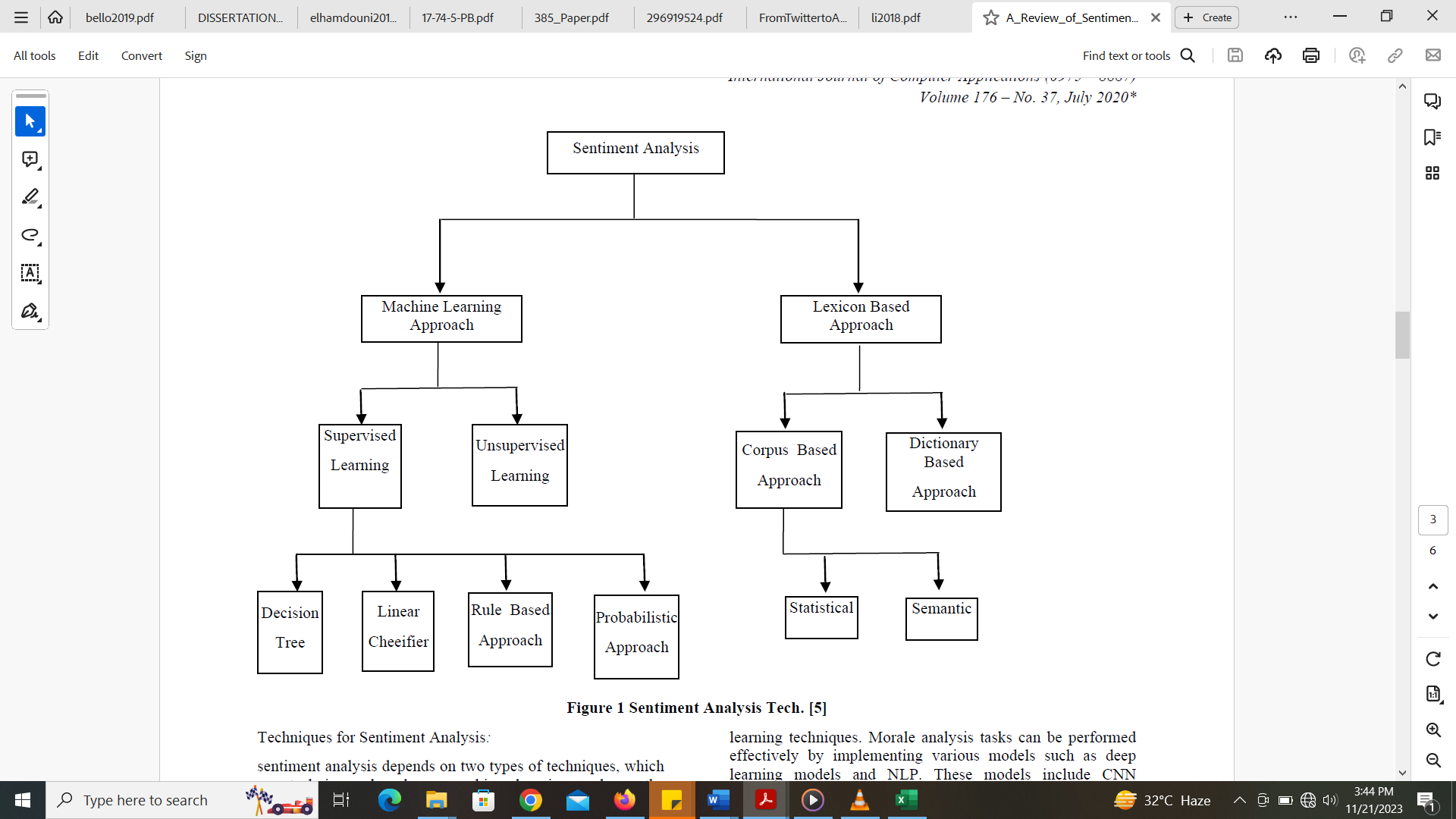
**Sentiment Analysis and Election Prediction**

Sentiment analysis has been used to predict the opinions of the citizens on US election using Twitter data (Wang, 2012). The authors used 17, 000 tweets to train their model (Naïve Bayes) and the model achieved a less than 60% prediction accuracy by classifying the tweets into positive, negative, neutral and not-sure and this assisted in analyzing real time tweets from the people which gave great insights about public opinions on each candidate. Another study (Neogi, 2021) used Twitter dataset to analyze tweets to get international opinions concerning a protest that happened in India conducted by the farmers and about 20,000 tweets was scraped to analyze and categorize tweets into positive, negative and neutral. Bag of words and TF-IDF were used to conduct the analysis and Bag of words outperformed TF-IDF. Other authors (Somula, 2016) and (Chandra, 2021) also established that Twitter can be used as an election indicator and has the ability to predict the favorite candidate of the people to emerge as the winner before the election is conducted. The people gave positive opinions about Donald Trump in almost all states in the United States prior to election. 1,000,000 tweets were collected from various users from different states and sentiment analysis was conducted. A study (Kausar, 2021) used Twitter data to check and conduct how different countries affected by the Corona virus coped with the situations. Tweets posted in English were analyzed to give the opinion and emotion of the people concerning the pandemic in their respective countries and 50,000 tweets were used in the study. Studies combined different features and used ensemble models to increase the detection of polarity of tweets for sentiment analysis and established that unsupervised and ensemble machine learning models outperformed other classical machine learning models in detecting opinions from text (Carvalho, 2021) and (Bibi, 2022). Authors have used Twitter data to examine political homophily in American Presidential Elections in 2016 where 4.6 million tweets were collected for analysis (Caetano, 2016). In some studies, authors used the ratio of positive message rate and negative message rate to predict the likely winner of a forthcoming election using twitter data and it showed that these opinions could be used to predict candidate that will emerge as the winner (Yavari, 2022). Twitter data has been used to establish that social media is not only used to express opinions, but it is also used to share ideas and opinions among other users. Authors used 100,000 tweets to predict German federal election in 2009 which could serve as a political indicator for the election (Tumasjan, 2010) and another study (Schmidt, 2021) also predicted the outcome of German presidential election of 2021 using 58,000 tweets which they established that traditional machine learning methods like Naive Bayes performed less than transformer-based models like the bidirectional encoder from transformers (BERT). Authors (Razzaq, 2014) performed social media sentiments about political parties to study and forecast Pakistan’s general election. They used supervised machine learning algorithms to classify tweets into positive, negative, and neutral. The findings of their experiment show that social media content can be a useful indication for identifying political behavior of various parties. In another study (Singh, 2017), 90154 tweets were analyzed and the results were compared with the actual election results, their model predicted the winning party accurately. Our overall goal in this study is to determine the sentiments in tweets with mentions of election-related words for the upcoming Nigeria 2023 elections and find if these tweets can provide meaningful insights regarding the election outcome. Some relative strengths have been identified from this study:

1. Size of dataset: It is noteworthy that bigger datasets enhance NLP models and improve their performance (Roberts, 2004). Having as large as two million tweets will enhance the statistical and predictive credibility of this study and also gives a more comprehensive view of the discourse.
2. Further NLU tasks such as topic modeling, tf-idf, context-based word-cloud, etc. help to provide a better applicability of the study vis-`a-vis the candidates and their impressions.
3. Social network analysis (SNA) (Al-Walid, 2019) is another strength of this study. This sees impression beyond the ones created by the candidates. It further analyses the network of verified friends around the candidates and their related activities towards their favorite candidates. Conducting SNA in this study for each respective candidate helps in determining homophily, assortativity, associativity, multiplexity and mutuality in the network of prominent friends of each candidate.
4. Further insights from the personal activities of the candidates were also used to measure their respective winning tendencies vis- `a-vis tweet pattern analysis, impression analysis, sentiment analysis of personal tweets etc. This is not prominently considered in the previous works as evident in the reviewed literature.
5. Timeliness is another strength of this study as this study pre-dates the actual election.

**Existing sentiment analysis techniques**

According to (Hamed, 2020), Sentiment Classification techniques are separated into two different techniques which is ML and Lexicon based Approaches.



Sentiment Analysis Techniques, source: (Hamed, 2020).

He went further to say that, sentiment analysis depends on two types of techniques, which are techniques based on machine learning and on the dictionary.

Techniques based on machine learning: In this type it is applied through extracting sentences and side levels. Features consist of Parts of Speech Marks (POS), n-gram, bi-gram, monogram and bag-of-words. Machine learning also has three aspects of sentence and side: Naive Bayes, Vector (SVM) and Maximum Entropy. and Lexicon-based technologies: In this type, they depend on trees through which decisions are made such as k-Nearest Neighborhood (k-NN), Conditional Random Field (CRF), Hidden Markov Model (HMM), and Class Dimensional Classification (SDC) and Chain Optimization (SMO) related to emotional classification methodologies.

The machine learning approach also has three categories:

a) supervised;

b) semi-supervised; and

c) non-supervised.

This approach is capable of automation and it can handle a large amount of data and therefore it is very suitable for the analysis of sentiment.

**Sentiment analysis tools and models.**

There are several existing sentiment analysis tools and models that have been developed and are widely used in various applications. These tools and models leverage natural language processing and machine learning techniques to analyze and classify the sentiment expressed in text data. Here are some notable ones:

1. VADER (Valence Aware Dictionary and sEntiment Reasoner): VADER is a lexicon and rule-based sentiment analysis tool designed for social media text. It is pre-trained on a wide range of data and can classify text as positive, negative, or neutral, as well as provide a sentiment intensity score.
2. TextBlob: TextBlob is a simple and easy-to-use Python library for processing textual data, including sentiment analysis. It provides a polarity score that indicates sentiment and a subjectivity score that measures how subjective or objective the text is.
3. NLTK (Natural Language Toolkit): NLTK is a popular Python library for natural language processing. It includes various sentiment analysis tools and models, such as the NaiveBayesClassifier and the SentimentIntensityAnalyzer.
4. IBM Watson Natural Language Understanding: This cloud-based service by IBM offers sentiment analysis among other NLP capabilities. It can analyze sentiment at the document, sentence, and entity level, and provides both positive and negative sentiment scores.
5. Amazon Comprehend: Amazon's Comprehend service offers sentiment analysis as part of its natural language processing capabilities. It can identify sentiment in a range of languages and provides a sentiment score.
6. Google Cloud Natural Language API: Google's NLP API includes sentiment analysis, which can determine the overall sentiment of a text as positive, negative, or neutral.
7. Stanford NLP Sentiment Analysis: The Stanford NLP toolkit offers a pre-trained deep learning model for sentiment analysis. It provides fine-grained sentiment labels, such as very positive, positive, neutral, negative, and very negative.
8. Hugging Face Transformers: Hugging Face offers a range of pre-trained transformer-based models for sentiment analysis, including BERT, GPT-2, and RoBERTa. These models can be fine-tuned for specific sentiment analysis tasks.
9. SentiWordNet: SentiWordNet is a lexical resource that assigns sentiment scores to words based on their meanings. It's used to calculate the sentiment of a piece of text by aggregating the sentiment scores of individual words.
10. Afinn: Afinn is a simple and lightweight sentiment analysis tool that uses a predefined list of words with associated sentiment scores to calculate sentiment in text.
11. Lexicon-based Models: Various sentiment lexicons, such as the Bing Liu's Opinion Lexicon and the NRC Emotion Lexicon, are used for sentiment analysis. These lexicons contain lists of words and their associated sentiment labels.
12. Deep Learning Models: There are deep learning models for sentiment analysis like Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) that can be trained on large datasets for more accurate sentiment classification.
13. OpenAI GPT-3: GPT-3, developed by OpenAI, is a large language model that can be fine-tuned for sentiment analysis tasks. It can generate human-like text and understand and generate sentiment.

**BERT, VADER VS RoBERTa MODEL**

BERT which stands for Bidirectional Encoder Representations from Transformers is a newer model based on the Transformers architecture. Compared to the transformer architecture, the BERT model only has encoders. The original publishers of BERT (Delvin, 2019) explain it as follows "BERT is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers". "As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific architecture modifications." (Delvin, 2019). As shown in figure 6, BERT is pre-trained on a large data set often publicly available text from the internet. The goal of pre-training is to make BERT learn "what is language" and "what is context". BERT learns the language by training on two unsupervised tasks simultaneously, these two are mass language model (MLM) and next sentence prediction (NSP). During the MLM phase, BERT takes in a sentence with random words filled with masks, the goal is to output the mask tokens that help BERT understand a bi-directional context within a sentence. In the case of the prediction of the upcoming sentence with NSP, BERT takes in two sentences and determines if the second sentence actually follows the first.

Bert can be trained on very specific NLP tasks, let’s take question-answering as an example. The things to be done is to replace the fully connected output layers of the network with a fresh set of output layers that can output the desired answer for a given question, Bert can then perform supervised training using a question-answering data set. The output parameters are the only part learned from scratch, the rest of the model parameters are slightly fine-tuned and as a result, the training time becomes faster (Delvin, 2019).

Sousa et al (Sousa, 2019) conducted a study on utilizing BERT for sentiment analysis in the stock market. The study investigated how sentiment analysis of texts, can be applied to financial news to improve the quality of decisions made by people in the financial industry. They note that BERT has achieved state-of-the-art results on various natural language processing tasks, including sentiment analysis.

The authors proposed evaluating BERT on financial news sentiment analysis problems to improve stock market prediction. The authors fine-tuned BERT with one additional output layer to create state-of-the-art models for sentiment analysis. They found that BERT was able to quickly analyze large volumes of financial news articles and provide valuable information for decision-making in the stock market (Sousa, 2019). The study is particularly relevant to this report as it demonstrates the potential of pre-trained models like BERT to improve the speed and accuracy of sentiment analysis in various industries.

**VADER**

Valence Aware Dictionary for sEntiment Reasoning (VADER) builds upon widely-known text analysis libraries known as polarity and intensity.

• Polarity: Looks at the polarity of different words and figures out if the word is positive or negative.

• Intensity: Looks at the intensity of different words and looks at how positive or how negative the word

is. It scales the various words from scale 1 to 9 with 1 being very negative, 5 as neutral and 9 being very positive.

The sentiment scores that Vader uses are a combination of both polarity and intensity and are specifically designed to analyze texts that are seen on social media such as Twitter. Texts from social media have features that we don’t see in other kinds of texts like acronyms, emoticons and emojis. The previous libraries didn’t have any kind of scores for these types of features and here is where Vader comes in. The first thing they had to do was to figure out what scores to assign to the unknown features. They solved this using various people to rate different types of acronyms and emojis and took the average of the results. Vader’s method for calculating compound scores involves evaluating each individual element in a sentence, adding up their scores, and then normalizing the sum to produce a final score on a scale of -1 to +1. However, Vader’s analysis extends beyond this basic approach, recognizing that other factors can influence the overall intensity of a text (Hutto, 2014). Two factors that Vader use is punctuation and capitalization. Adding things like exclamation points or using all caps to a sentence increases the intensity and Vader figured out a solution by how much a score should increase by including those factors. To develop a solution, they hired people to rate various sentences. Beginning with a baseline sentence lacking any of the target features, they gradually introduced each feature and had different individuals rate its impact. Through this process, Vader was able to precisely quantify the degree to which each factor should increase the

score.

Although VADER offers the advantage of being a quick and rule-based sentiment analysis tool, its limitations could potentially have a negative impact on the analysis of Samsung A53 reviews.

For instance, when dealing with complex or nuanced language such as reviews from Amazon, irony and sarcasm could be misinterpreted, misspelt or get grammatical errors in the text. With it being analyzed it can lead to important words or phrases being overlooked during sentiment analysis.

Potentially resulting in less accurate results (DeLancey, 2023). It could not be as accurate as other models that use machine learning techniques.

Anwar et al. from Lahore University of Management Sciences presented a study on the use of Twitter by QAnon supporters during the US Presidential Elections of 2020 (Anwar, 2021). The study collected over 12 million tweets for 46 consecutive days starting from August 1st to September 15th, 2020 containing the keywords "Trump", "Biden", or "Election2020". The tweets were analyzed using VADER to determine the overall sentiment of a text. The results showed that QAnon supporters were highly active on Twitter during the US Presidential Elections of 2020. They used Twitter as a platform to spread conspiracy theories and misinformation about the election process.

The sentiment analysis using VADER revealed that QAnon supporters had a highly negative sentiment towards Joe Biden and a highly positive sentiment towards Donald Trump (Anwar, 2021). The research, utilizing the VADER sentiment analysis tool, underscores the significance of social media platforms as influential spaces where political discourse, conspiracy theories, and misinformation can proliferate, ultimately shaping public perceptions and opinion

**RoBERTa**

RoBERTa builds on BERT’s language masking strategy, wherein the system learns to predict intentionally hidden sections of text within otherwise unannotated language examples. RoBERTa, which was implemented in PyTorch, modifies key hyperparameters in BERT, including removing BERT’s next-sentence pretraining objective and training with much larger mini-batches and learning rates. This allows RoBERTa to improve on the masked language modelling objective compared with BERT and leads to better downstream task performance (MetaAI, 2023). RoBERTa was built by the Ai department of Facebook (Meta Ai) in 2019.

The authors Liu et al (Liu, 2019) used a similar structure as the BERT model but have made some modifications compared to the BERT model. The data used to train the RoBERTa model consists of ≈ 160 GB, compared to the BERT model which consists of 16 GB. The datasets taken in the pre-training process were the following:

• BookCorpus (16 GB) (Zhu, 2015). BookCorpus, the same dataset trained on BERT, was one of the datasets to train ReBERTa. BookCorpus is a collection of over 11 thousand novel books with different genres.

• CC-News (76 GB) (Common, 2023) CC-news consists of 63 million English news articles from September 2016 to February 2019 (Liu, 2019). The total amount was estimated to be 76 GB of data.

• Openwebtext (38 GB) (Radford, 2019). Openwebtext is web content extracted from URLs shared on Reddit with at least three upvotes (Liu, 2019).

• Stories (31 GB) (Trinh, 2018). The Stories dataset contains a subset of CommonCrawl data filtered to match the story-like style of Winograd schemas (Liu, 2019). 

Figure 6: Comparison result between RoBERTa, BERT and XLNET (Liu, 2019).

The authors (Liu, 2019) have compared the development result for the RoBERTa model, the BERT model and the XLNET model. As shown in Figure 6, It has been noticed that the RoBERTa model got the best f1 score when training the model with 500k steps and with 8k in batch size.

The following sentiment analysis tool will therefore use the RoBERTa model as the Ai model when generating sentiment analysis.

Accuracy: RoBERTa, being a deep learning model, is expected to provide higher accuracy in sentiment analysis compared to VADER, which is based on a lexicon-based approach. You can evaluate the accuracy of both models on a labeled sentiment dataset to compare their performance.

Domain-Specificity: RoBERTa can be fine-tuned on domain-specific data, making it more adaptable to specific domains, while VADER has a more general-purpose lexicon.

Preprocessing: VADER requires minimal preprocessing, while RoBERTa may require careful tokenization and preprocessing.

Customizability: RoBERTa can be fine-tuned on custom sentiment analysis tasks, while VADER has fixed lexicon rules.

Speed: VADER is generally faster for sentiment analysis, whereas RoBERTa can be computationally intensive.

Robustness: RoBERTa can handle complex sentence structures and nuances better, while VADER might struggle with complex language.

Language Support: RoBERTa models can be used for sentiment analysis in multiple languages with appropriate pre-trained models, whereas VADER might be limited to English.

**Challenges of Sentiment Analysis in Nigerian Elections:**

Researchers have highlighted the challenges of sentiment analysis in the Nigerian context, including the use of multiple languages, dialects, and code-switching, which make sentiment analysis more complex.

The influence of local dialects and regional variations in sentiment expression adds a layer of nuance to the analysis.

**Impact of Social Media:**

Social media platforms, particularly Twitter and Facebook, have become significant sources of data for sentiment analysis during Nigerian elections.

Researchers have found that social media can be a powerful tool for political communication and sentiment expression, making it a rich source of data for analysis.

**Public Sentiment Trends:**

Researchers have identified trends in public sentiment during Nigerian elections, such as the expression of optimism, skepticism, or disillusionment with the electoral process.

Sentiment analysis has revealed shifts in sentiment dynamics during different phases of the election cycle, including the pre-election, election day, and post-election periods.

**Campaign Strategies:**

Sentiment analysis has been used to assess the effectiveness of campaign strategies and messaging. Researchers have investigated which campaign messages resonate with voters and how they influence sentiment.

**Predictive Power:**

Some studies have explored the predictive power of sentiment analysis in forecasting election outcomes. Researchers have investigated the correlation between sentiment and election results.

**Misinformation Detection:**

The detection and analysis of misinformation and disinformation have been integral to sentiment analysis in Nigerian elections. Researchers have sought to identify and combat false narratives and rumors.

**Regional Variation:**

Studies have examined regional variations in sentiment, considering the diverse cultural, linguistic, and political landscape of Nigeria. Sentiment analysis has revealed how issues are perceived differently in various regions.

**Demographic Differences:**

Researchers have investigated how sentiment varies across different demographic groups, including age, gender, and socioeconomic factors. This helps in understanding the diversity of voter opinions.

**Future Research Directions:**

Several studies have pointed to future research directions, such as the need for more robust sentiment analysis models, the exploration of real-time sentiment analysis, and the examination of the role of social media influencers.

**A review of sentiment analysis in political contexts.**

Sentiment analysis in political contexts has become an increasingly important area of research and application. It involves using natural language processing (NLP) and machine learning techniques to analyze and understand public sentiment and opinion regarding political issues, candidates, and elections. Here is a review of sentiment analysis in political contexts, highlighting its significance, challenges, and applications:

**Significance of Sentiment Analysis in Political Contexts:**

Real-Time Public Opinion Monitoring: Sentiment analysis enables the real-time monitoring of public sentiment and opinions on political matters. This can help political parties, candidates, and policymakers gauge public reactions and adjust their strategies accordingly.

Election Forecasting: Sentiment analysis can be used to predict election outcomes by analyzing public sentiment towards political candidates, parties, and issues. It provides insights into voter preferences.

Policy Analysis: Sentiment analysis can assist in understanding how the public perceives and reacts to policy changes and political decisions. This insight can inform policy-making and political discourse.

Crisis Management: During crises or emergencies, sentiment analysis can help government agencies and political leaders assess public sentiment and respond effectively to mitigate negative reactions.

Campaign Strategy: Political campaigns can benefit from sentiment analysis by identifying which issues resonate with voters and adapting their messaging and strategies accordingly.

**Challenges and Considerations:**

Language Variability: Political discussions often involve slang, regional dialects, and diverse language use. Sentiment analysis models must account for this variability.

Contextual Understanding: Sentiment analysis needs to consider the context of political discussions. A word that may be positive in one context can be negative in another (e.g., "revolution" can be positive or negative).

Bias and Polarization: Political discussions can be highly polarized, making it challenging to develop sentiment analysis models that avoid bias. Models may inadvertently reinforce existing biases.

Irony and Sarcasm: Political discourse frequently includes irony and sarcasm, which can be difficult for sentiment analysis models to detect and interpret accurately.

Data Collection: Gathering relevant data for sentiment analysis can be challenging. Social media platforms and online discussions can have a vast amount of data that needs to be collected and processed.

**Applications of Sentiment Analysis in Political Contexts:**

Election Polling: Sentiment analysis can be used to gauge public opinion on candidates, parties, and election issues. This information is valuable for election polling.

Issue Analysis: Analyzing sentiment can help identify public sentiment regarding specific political issues, enabling policymakers to address concerns.

Media Monitoring: Media organizations use sentiment analysis to understand how their audiences respond to political coverage, providing feedback for journalistic decision-making.

Political Advertising: Political campaigns use sentiment analysis to assess the effectiveness of advertising and messaging strategies.

Government and Policy: Government agencies use sentiment analysis to monitor public response to policies, crises, and government actions.

Public Diplomacy: Sentiment analysis can be used in the context of international relations to assess how foreign audiences perceive a country's political leaders and policies.

In conclusion, sentiment analysis in political contexts has grown in importance due to the prevalence of online discussions and social media platforms where political debates and opinions are expressed. Its significance lies in its potential to inform political decision-making, public engagement, and election strategies. However, it also faces challenges related to language variability, bias, and contextual understanding. When applied effectively, sentiment analysis can provide valuable insights into the ever-evolving world of politics.

**Previous studies on sentiment analysis in elections.**

Sentiment analysis in the context of elections has been the subject of various studies and research projects. Researchers have explored sentiment analysis to understand public sentiment, political discourse, and the impact of sentiment on electoral outcomes. Here are some notable previous studies and research papers related to sentiment analysis in elections:

Pak, A., & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. - This study examines sentiment analysis on Twitter during the 2010 French presidential elections, analyzing the sentiment expressed by Twitter users towards different candidates.

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment Analysis of Twitter Data. - The research focuses on sentiment analysis using Twitter data during the 2010 U.S. midterm elections to understand how Twitter users' sentiment correlated with election results.

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. - This study investigates the relationship between sentiment expressed on Twitter and public opinion polls during the 2010 U.S. midterm elections.

Barberá, P., & Rivero, G. (2014). Understanding the Political Representativeness of Twitter Users. - This research explores how sentiment and political behavior are linked on Twitter during the 2011 Spanish elections.

Gayo-Avello, D. (2012). No, You Cannot Predict Elections with Twitter. - This paper presents a critical perspective on the limitations of using Twitter data and sentiment analysis for predicting election outcomes.

Gruzd, A., & Roy, J. (2014). Investigating Political Polarization on Twitter: A Canadian Perspective. - This study examines sentiment and polarization in political discussions on Twitter during the Canadian federal elections.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. - This research investigates the predictive power of Twitter data and sentiment analysis for the 2009 German federal elections.

Joulin, A., Grave, E., Bojanowski, P., Mikolov, T., Bagheri, M., Lample, G., & Zemeckis, J. (2017). FastText.zip: Compressing text classification models. - While not specific to elections, FastText is an efficient tool for text classification and sentiment analysis, which can be adapted for election-related sentiment studies.

These studies represent a selection of research on sentiment analysis in elections, demonstrating the application of sentiment analysis in understanding public sentiment, predicting election outcomes, and studying political discourse on social media platforms. Researchers have used various methods, data sources, and techniques to conduct sentiment analysis in the context of elections, contributing to a growing body of knowledge on the subject.