PERCEPTION MINING AND SENTIMENT ANALYSIS OF POLITICAL SOCIALIZATION AMONG TWITTER USERS IN THE 2023 NIGERIA GENERAL ELECTION

Nicholas Ude Eze
Department of Computer and Robotics Education, University of Nigeria, Nsukka,
Nsukka, Enugu, 234, Nigeria
nicholas.eze@unn.edu.ng

Ifeoma Joanes Onodugo
Department of Management, University of Nigeria, Enugu Campus,
Nsukka, Enugu, 234, Nigeria
ifeoma.onodugo@unn.edu.ng

Stella Ifeoma Osondu
Department of Science Education, Michael Okpara University of Agriculture, Umudike,
Abia State, Nigeria, 234, Nigeria
ifeoma.osondu@mouau.edu.ng

Ekwutosi Ozioma Chukwu
Department of Vocational and Entrepreneurship Education, University of Nigeria, Nsukka,
Nsukka, Enugu, 234, Nigeria
ekwutosi.chukwu.pg92474@unn.edu.ng

Ndukanma Ngozika Agape
Department of Computer and Robotics Education, University of Nigeria, Nsukka,
Nsukka, Enugu, 234, Nigeria
ngozika.ndukanma@unn.edu.ng

Kelechi Patricia Onyishi
Department of Computer and Robotics Education, University of Nigeria, Nsukka,
Nsukka, Enugu, 234, Nigeria
kelechi.onyishi.pg94189@unn.edu.ng

Lovina Ifeoma Eze
Department of Computer and Robotics Education, University of Nigeria, Nsukka,
Nsukka, Enugu, 234, Nigeria
ifeomaezelov@gmail.com

Abstract

The study analyzes the applicability and political use of Twitter using sentiments and content (textual) analysis with the purpose of examining the pattern of online communications among Nigerian voters during the run up to the 2023 Nigerian General Elections (NGE23) to make prediction for winners. Naive Bayes, Support Vector Machine, and Random Forest were utilized to determine sentiment analysis for English tweets, while ICT specialists were employed to determine content analysis for the three key Nigerian languages – Igbo, Hausa, and Yoruba tweets. The results of two vectorization algorithms were compared to those of English tweets and our finding reveals that overall sentiments were positive throughout the study period, though negative sentiments appeared to be slightly higher among the English language speakers than among the Nigerian Language speakers.

***Keywords*: Content Analysis, Political Socialization, Perception Mining, Sentiment Analysis, Social Media, Tribal Sentiments, Twitter**

1. Introduction

The phenomenon of political socialization in Nigeria has a long history and the preparation of individuals for their roles in the political world is also as old as politics itself. That is to say, both political authorities and ordinary citizens of Nigeria have been subject to the practices and outcomes of political socialization regardless of political regime type. Given an understanding of political socialization as an induction into the political culture of a regime that comprises 1) cognitive, 2) normative and 3) affective components Almond, (2000), Gidengil, Wass & Valaste (2016), there are different and alternative explanations with regards to how and when the people form their political attitudes and ideas about government toward change and whether these attitudes changes throughout their live time. One of the important explanation concerns the understanding of the fact that since Nigeria gained independence from the British government in 1960, politicians from the north, such as those who speak Hausa or Fulani, have had more success governing Nigeria than their southern counterparts, such as those who speak Yoruba or Igbo.

This is due in part because of the North and South's diverse experiences with imperial rule which has led to marginalization cry by certain part of the country (Lazarus, Button, 2022). Also politicians of all stripes in Nigeria have been linked to corruption and nepotism which have an impact on how political power brokers in Nigeria distribute the "national cake" (national riches) both legitimately and illegally, as well as how elections are managed and won with total disregard to the youth (Nwabueze, 2020). These has aggravated the Nigerian youth who have remained unemployed for many years after graduation (Olubusoye, Salisu & Olofin, 2022). Aside youth unemployment and labour market flexibility, the effects of the macroeconomic uncertainties in Nigeria and the Central Bank of Nigeria (CBN) monetary policy which led to naira scarcity, fuel scarcity etc CBN, (2015), Ososuakpor, (2021) that threw Nigerians into sufferings; the lingered dispute between the federal government and the University based unions (ASUU, NASU, SSANU and NAAT) Monogbe, Monogbe, (2019), that led to students staying out of school throughout the year 2022 all frustrated the students (youths) and sent them on a mission to get even.

However, the youths' experiences have remained their best teacher; and desperation was never their best option. With the advent of information communication technology (ICT), Nigerian youths both those staying in the rural and urban areas have stepped aside from the track to teach politicians what the political race to the 2023 election victory looks like using the innovation. What Nigerians expected was just to have peaceful coexistence among the various political parties – the Labor Party (LP), Peoples Democratic Party (PDP), All Progressive Congress (APC), All Progressive Grand Alliance (APGA), Action Alliance (AA) and others that are going to participate in the 2023 general election (INEC, 2022). Also what Nigerians expected was that the 2023 general election is going to wear a different look in Nigeria because the electorates, especially the youths have resolved to cast their votes for tested and trusted individuals who will not joke with the confidence reposed on them, as against the erstwhile practice of 'vote your caste' which is mainly based on religion and region which has often led to supporting parties with candidates whose lifestyles are questionable (Karleen, 2020).

However, as the 2023 general election draws nearer, the presidential candidates of the various political parties too were not resting on their oars. All the political parties in Nigeria were making frantic efforts to make a breakthrough in the election. The supporters of various presidential candidates made the political scene very popular too. For example, the youth organized numerous solidarity marches in support of their candidates which according to Miriam, (2019) is a mark of contentious moves and a strong degree of politicization. They had not only done that, but also their approaches cum methodologies had endeared their candidates to people of all classes. One of their approaches cum methodologies being the use of social media.

Social media become a more important source of political knowledge for Nigerian youths at the moment. Numerous studies are cautiously optimistic about how social media affect young people's civic and political engagement (Delli, 2000, Lupia, & Philpot, 2005). According to experts, the interaction, use, and accessibility of social media, as well as web innovations like blogs, YouTube, and social networking sites, all have the potential to revive young people's dwindling rates of political engagement. Nigerian youths used the internet as a tool for political action through peer mobilization during the online campaigns against the bad government because of its participatory, interactive as well as cost effective nature (Acemoglu & Robinson, 2013). Social media thus was described as democratic because of the unfettered access of its users to send and receive political campaign information (Alegu, Maku, Adelaja, & Rasheed, 2019). However, such political campaigns information at the time was characterized with more or less, an avenue for peddling deliberate incitements, intolerance, parody and falsehood which should not have a place in a decent political campaign against the perceived opponents all in an attempt for political parties to markets their choice candidates. Politicians embarked on battle of wits as they struggled for the minds of the voters on whom their victory depended on to elevate what has been a 'beer parlour' gossip and social media lies being propagated by supporters using the instrumentality of tribal sentiment.

Tribal sentiment according to Mohan, Mihir, (2022) refers to a feeling-driven attitude, thinking, or judgment (positive or negative) of voters or electorates against specific tribes or entity. However, because these feelings-driven attitudes, thinking, or judgements are mostly expressed in text especially using Twitter, it is always difficult to define the orientation of the expressed sentiment in other to determine if the text expresses the positive, negative or neutral sentiment of the user about the entity in consideration (Fang, Zhan, 2015). This is however, one of the fundamental problems of sentiment analysis.

Sentiment analysis refers to a branch of study that studies people's attitudes toward specific things (Wankhade, Rao & Kulkarni, 2022; Yue, Chen, Li, Zuo, Yin, 2019; Yousif, Niu, Tarus, Ahmed, 2019). It is a machine learning technique in social media and digital marketing that detects polarity. Polarity on the other hand refers to the strength of a perception *w*hich could be positive or negative (Yadav, Vishwakarma, 2020). Polarity categorization during sentiment analysis, according to Fang, Zhan, (2015), Liu, (2010) and Pak, Paroubek, (2010), is one of the most difficult aspects of Natural Language Processing (NLP). The dearth of work in this area is probably due in part to a lack of computational power, and also due in large part to the lack of large publicly available online dataset with sufficient size and diversity of specific perceptions, to indicate whether such perception is positive, negative, or neutral, to perform an in-depth analysis. The only exception till date is the Stanford Sentiment 140 Tweet Corpus (Stanford, 2014).

However, few authors have attacked the problem of sentiment detection and analysis primarily to determine winners in elections Kreiss, (2016); Pramuk, (2016); Burnap, Gibson, Sloan, Southern & Williams, (2016); Yaqub, Chun, Atluri, & Vaidya, (2017b); Enli, (2017); Ahmed, Jaidka & Cho, (2018); Xenos, Scheufele, Brossard, Choi, Cacciatore, Yeo, et al. (2018); Lee & Xenos, (2019) and to gauge public perceptions of politicians Mohammed, Zhu, Kiritchenko, & Martin, (2015); Yaqub, Chun, Atluri & Vaidya, (2017a); but not within the Nigerian political context. In our paper, we focus on the applicability and political use of Twitter with the purpose of examining the pattern of online communications among Nigerian electorates during the run up to the 2023 Nigerian General Elections (NGE23) for the task of sentiment analysis. This is because Twitter was used to enhance the election through voter education and dissemination of information on political party's policy preferences, views of politicians and electorates (Yaqub, Chun, Atluri & Vaidya, 2017a). This means that, aside from Twitter being more communicative than prior web versions, the Twitter app is unlike any other social networking site in terms of functionality and performance (Enli, 2017).

The notion of Twitter emerged in the early 1970s, when authors such as Alvin Toffler predicted that web 1.0 media will shift from professional content producers to passive content consumers, and that more interactive content producers and consumers (prosumers) would be desirable (Toffler, 1970). Tim Berners Lee, the creator of the World Wide Web, later echoed these sentiments, saying, "I have always envisaged the information universe as something to which everyone has instant and intuitive access, and not just to explore, but to contribute" (Berners-Lee, 2000). This term contradicts the assumption that content is created by one group and consumed by another. Instead, content is created and consumed by the same group (Toffler, 1970). What this also mean is that users of the Twitter app are participating in the co-creation of material and can be considered co-creators, allowing them to play a number of roles such as content consumer, commenter, creator, and collector (Malik, Heyman-Schrum, & Johri, 2019). However, users as content consumers occupy a passive role in which they read texts, look at pictures, listen to music, or watch videos online. On the contrary, users as content creators play an active role, contributing a variety of items that enable opinion sharing utilizing various forms such as words, audios, photos, videos, links, and so on (Kamal, Sajid & Iftikha, 2015). This has been done in a unidirectional and asymmetrical manner, allowing people to connect without formal notification. The provision of enormous internet data via Application Programming Interfaces (APIs) has always helped in collecting and evaluating such contents created (Valenzuela, Correa & Gil de Zúniga, 2017).

Twitter provides three Application Programming Interfaces (APIs) that allow for the collection and evaluation of the tribal sentiments as well as Tweets on political socializations created. The REST API, for example, allows for the collection of user status data and information; the Search API allows users to query specific Twitter material; and the Streaming API allows users to collect Twitter content in real-time (Twitter, 2014). However, there are a number of difficulties with these internet data that could cause the sentiment analysis process to be disrupted (Fang, Zhan, 2015). The first concern is that because users (voters) can freely express their thoughts, the quality of those perceptions cannot be assured because most voters will likely transmit useless, irrelevant, and fraudulent perceptions about a candidate rather than sharing topic-related perceptions or making issue-based arguments (Liu. 2014; Jindal, Liu, 2008; Mukherjee, Liu, Glance, 2012). For example, in the run-up to the 2023 general election in Nigeria, over 400 Twitter identities were identified daily churning out facts and fake news to criticize the opposition party, with over three thousand Tweets (Lazarus & Button, 2022). All of these activities contribute to Twitter's global user base, which topped 500 million users with over 328 million active users (Statista. 2018). The second concern is that people are gullible, and they are often very keen in expressing their thoughts, emotions, or judgments, as well as spreading fake news about certain entities or services, on Twitter. As a result, their reviews on Twitter platforms can have a significant impact on voters making informed and optimal decisions when voting for a candidate in an election. The third concern here is the fear that such evaluations and fake news could be utilized by persons who want to promote behaviors that are harmful to a country's corporate existence (Schone, Parkinson & Goldenberg, 2021).

However, it is too soon to say if this was the beginning of a permanent political revolution as the run up to the 2023 Nigerian General Election marked the beginning of Twitter's tremendous effect in political campaigns in the country's political history, with overwhelming supporters of the various political party's Twitter account being the popular at the time. In fact, the election campaign is dubbed the 'Twitter election' because the Twitter platform serves as a major platform for politicians especially from the Labour Party (LP) to mobilize Nigerians, resulting in a record turnout of about 90% youth voters. Many voters see this presidential candidate's emergence as engendering a mass movement that will have a similar if not stronger impact like the 2020 #ENDSARS protest in Nigeria, while others see him as one that will evolve a new Nigeria, with the capacity to change the way things are been done by upholding a progressive ideology that is not based on tribe, religion or political leaning.

However, both the ruling coalition and the opposition used Twitter for campaigning and citizen outreach, and were primarily characterized by all forms of issue-based and negative clues against each other in an attempt to ousting the long-serving government of the All Progressive Congress (APC) and vice versa. Following that, individual politicians followed suit, and as at the time of conducting this research, practically all of the main political parties' politicians, both old and young, are on Twitter, with the most well-known among them being Femi Fani Kayode (FFK), Mr. Sowore, Senator Shehu Sani, Senator Dino Meleye, Reno Omokri among others.

1.1. *Purpose of the Study*

The main goal of this paper was to determine the role twitter played for the winner to emerge in the run-up to the 2023 Nigeria General Election. Specifically, the study will (1) describe the largest publicly available dataset for performing in-depth predictive analysis using Twitter analytics; (2) report baseline performance of the Tweets of the four major languages spoken in Nigeria, namely English, Igbo, Hausa, and Yoruba; and (3) investigate the performance of several machine learning algorithms (Nave Bayes, SVM, and Random Forest) with two word embedding approaches (Word2Vec and String2Vec) for the English Tweets; (4) Determine the sentiments of Igbo, Hausa, and Yoruba Tweets using content analysis; and (5) identify the top trending keywords for all languages, both in terms of total analysis and sentiments.

1.2. *Contributions*

The following are the contributions of our study to the body of literature:

1. In our work, we propose a method to automatically collect a corpus of texts containing positive and negative attitudes, and to classify the texts without the need for human interaction.
2. We analyze the corpus of data collected statistically.
3. We create a sentiment categorization algorithm for all Tweets using the gathered corpora.
4. To demonstrate the effectiveness of our suggested strategy and show that it outperforms earlier proposed methods, we carry out experimental assessments on a set of genuine Twitter posts.

2. Relevant work

Opinion mining and sentiment analysis have gained popularity as a result of the growth of blogs and social networks. The existing body of research was summarized in great detail in (Pang & Lee, 2008). Current methodologies and procedures for opinion-oriented information retrieval are described in the authors' survey. There hasn't been many research on opinion mining, however, that have considered blogs or even microblogging. Web-blogs are used by the authors of Yang, Hsin-Yih Lin & Hsin-His, (2007) to generate a corpus for sentiment analysis, and they use the emotion icons that are assigned to blog entries as indicators of readers' moods. The authors looked at a variety of approaches to determine the overall sentiment of the content after using SVM and CRF learners to identify emotions at the sentence level. As a result, the winning strategy is chosen by setting its baseline at the feeling conveyed in the paper's final sentence. J. Read created a training set for the sentiment categorization in using emoticons like ":-)" and ":- (" (Read, 2005). For this project, the author collected emoticon-heavy sentences from Usenet newsgroups. Happy emoticons were classified as "positive" emoticons, whereas sad or angry emoticons were classified as "negative" emoticons. On the test set, two emoji-strained classifiers, SVM and Naive Bayes, were able to get up to 70% accuracy. Before doing a sentiment search, the authors of Go, Huang & Bhayani, (2009) used Twitter to acquire training data. The strategy is comparable to that of (Read, 2005). The authors create corpora by obtaining "positive" and "negative" samples using emoticons, and then they use a variety of classifiers. The Nave Bayes classifier using a mutual information measure for feature selection produced the best results. On their test set, the authors were able to get up to 81% accuracy. With three classes ("negative," "positive," and "neutral"), the technique performed poorly.

2.1. *Political Socialization and Social Media Penetration in Nigeria*

Nigeria is a multi-ethnic country with ethnic groupings such as Igbos, Hausas, Yorubas, and others. English is the official language of communication in Nigeria, despite the fact that there is no dominant national language. The vast majority of English speakers are bilingual and of a certain ethnic background. It is often thought that English speakers have a higher socio-economic position (i.e., better education, standard of living, income level, etc.) than the remainder of the population who speak local Nigerian languages. In truth, English-speaking Nigerians are generally viewed as arrogant, boastful, highly westernized, and outsiders probably because of their high education (Lee, Wong & Azizah, 2010). According to reports, Nigerians' use of social media is influenced by their education (Al-Ansari, 2006). People with higher education spend more time on the Internet than those with lower education and they account for over 61% of Nigerians who utilize social media to express their views especially on political issues and electioneering campaigns (Hartley, 2007).

According to surveys, the Internet and Nigerians' higher socioeconomic position are changing the demography of politically active citizens (Lazarus, Button, 2022). To give one example, among Nigerians, those with a high socioeconomic position are most likely to utilize the Internet for political goals (Jennings& Zeitner, 2003; Krueger, 2006). However, research indicates that young people who become interested in politics through an internet campaign are not more likely to remain so (Krueger, 2002). Furthermore, despite the fact that some politicians are using the Internet in their election campaigns, the majority still disseminates information in the same ways they have in the past rather than switching to interactive formats that are more effective in swaying young people (Iyengar, & Jackman, 2004; Xenos, 2008). The issue does not get much better when one considers alternative news sources. Younger Nigerian voters are growing more and more responsive to controversial shows like The O'Reilly Factor as well as entertainment "news" shows like The Daily Show. Audience engagement has also been associated to more pessimistic perceptions regarding politicians, elections, and the media, even though there is some evidence that these alternative news sources increase viewers' political literacy (Baumgartner, & Morris, 2006; Baek, & Wojcieszak, 2009).

However, the Trump presidential campaigns demonstrated promise that political campaigns could use new media to efficiently organize adolescent engagement using email, texting, online social networking, and online finance (Enli, 2017). This is reflected in the increased usage of the Internet for political socialization and participation in Nigeria during the recent decade. For example, Internet penetration among Nigerians within the past four years was estimated to be 46.59% in 2019, 49.14% in 2020, 51.44% in 2021 and 53.51% in 2022 of the country's population, according to the most recent Internet users' statistics for Nigeria, compared to 28.9% of all internet users in Africa (Joseph, 2022). Younger Nigerians between the ages of 20 and 34 were the highest internet users (53.6%), followed by those between the ages of 35 and 49 (24.7%) Statista, (2018), suggesting that younger generations are inversely related to internet access and use among users (Kaur & Balakrishnan, 2018). Corroborating, Eze, Obichukwu & Nnadi, (2018) infers that these young users are digital natives who grew up on the Internet and are at ease utilizing it to suit their information demands. Thus Twitter, among other internet resources, could meet their political campaign needs, allowing them to create, collaborate, contribute, connect, share, and participate in certain entities (Eze, Olive, Ada, 2017). However, despite the fact that Instagram and Facebook is Nigeria's most popular social media network (77.9% and 71.2%) respectively, with telegram users accounting for 50.3%, a vast majority of younger Nigerian voters (71.8%) were discovered to actively exchange political information and campaigns online, primarily via Twitter (57.4%) during the run up to the 2023 general election in Nigeria (Doris, 2021).

2.2. *Theoretical Framework*

In this study, Everett Roger's Diffusion of Innovation (DOI) theory was used to examine how members of a social group accept and make decisions about new innovative ideas (Rogers, 1962). The spreading process involves both social media and interpersonal communication channels. According to this theory, Twitter usage was widely implemented in real-world circumstances, with the adaptation of the usage culture playing a key role wherever the theory is applied. This paradigm, as first provided and later adapted in IT adoption studies Giachanou, Crestani, (2016); Nguyen, Nguyen, (2020); Huq, Ali & Rahman, (2017) provides a valuable analytical framework that may be applied to investigate individual Twitter adoption and assimilation rates. The framework proposed the following four elements:

- Innovations – an individual's perception of a new idea, practice, or object that can also be an inspiration to try something new or bring about social change,
- Communication Channel – a channel by which messages are passed from one person to another. Innovations propagate across the population through the medium of communication. It can take any shape, such as word of mouth, SMS, or any literary form,

- Time – refers to the amount of time it takes for individuals to adjust to new technologies in a culture.
- Social System – Interconnected network groups get together to address challenges for a common aim in a social system. The term "social system" refers to all of the elements that go into making up a society, such as religion, institutions, and social groups etc.

The DOI theory has been tested, and the four parts have been proven to describe how people recognize a need, seek for, and accept new technology in each of these frameworks, confirming Roger's assertion that DOI settings influence adoption and usage of innovation.

3. Methodology

In order to satisfy this study's long and short term objectives, we employed data collected between March 2022 and February 2023. Before being used for analysis, each product review was evaluated and scored on a positive, negative, or neutral scale. Following that, the reviews were assessed on a five-point scale ranging from 5 to 1. Finally, we proposed a mathematical model for calculating sentiment scores. The effectiveness of well-known machine learning approaches was evaluated and compared. Then there were comparison studies on the effectiveness of various strategies based on online user reviews from Nigerian voters. The approaches utilized in this investigation consists of seven steps, starting with data gathering and ending with model evaluation. The next parts go over each of the seven steps as discussed in the subsequent units.

3.1. *Step 1 – Data Collection*

With the help of the Twitter API, we gathered a corpus of text messages and created a dataset with three classes: objective texts, negative feelings, and positive sentiments (neutral sentiments). We used the same process as in to gather both favorable and negative sentiments (Read, 2005; Go, Huang & Bhayani, 2009). Two categories of emoticons were found on Twitter:

- Happy emoticons: ":-)", ":)", "=)", ":D" etc.
- Sad emoticons: ":-(", ":(", "=(", ";(" etc.

The two different types of corpora that were gathered were utilized to train a classifier to distinguish between positive and negative emotions. In order to collect a corpus of objective posts, we retrieved all text messages from Twitter accounts between March 2022 and September 2022, using top trending political hashtags that were monitored throughout the observed period (i.e. #TakeBackOur9ja, #OBIdient, #OurMumuDonDo, #docs4peterobi, #PleaseVoteChange, #PowerMustShift etc). Approximately 311,417 (i.e. 65 percent English versus 7 percent Yoruba, 18 percent Igbo, and 10 percent Hausa) Tweets were gathered as shown in table 1. To gather a training set of objective texts, we queried accounts for all 44 hashtags. However, Tweets, Retweets (a type of endorsement by sharing the Tweet with others), mentions (mentioning someone explicitly, e.g. @NickEze), responses to Tweets, user locations, number of followers and following, language, date, and time were all fetched using a Python script.

3.2. *Step 2 - Pre-Processing and Tweet Statistics*

Each message is limited to 140 characters by the terms of the Twitter site, therefore it is typically just one phrase long. As a result, we presume that an emoticon in a message symbolizes an emotion for the entire message and that this emotion is tied to all of the words in the message. We only utilize English in our study. However, as Twitter API allows users to define the language of the retrieved posts, our solution is simply adaptable to other languages. To improve the accuracy of the analysis, irrelevant details (i.e. noise) were deleted from the dataset. This is because the overall sentiment may not always be conveyed through the repetitive usage of keywords. So for general information retrieval purposes, the frequency of a keyword's occurrence is a more suitable attribute. For instance, the dataset was filtered and cleaned using the binary feature of an n-gram. Bigrams, trigrams, and unigrams have also been employed, though. Contrarily, Dave et al. found that bigrams and trigrams were more effective at categorizing the polarity of product reviews (Dave, Steve & Pennock, 2003). However, when assessing the emotion of movie reviews, Pang et al. found that unigrams outperformed bigrams (Pang, Lee & Vaithyanathan, (2002). So, we sought for the appropriate Twitter data parameters. On the one hand, trigrams and other high-order n-grams should be better at capturing the patterns of emotional expression. On the other hand, unigrams should provide a good representation of the data. Therefore, in order to extract n-grams from the Twitter post, we used the following criteria:

- Only English, Igbo, Hausa, and Yoruba Tweets (Tweets containing fewer than three Nigerian words were considered as English, and vice-versa). Other languages' Tweets were discarded.
- Emojis, emoticons, special characters (@,!, &) and urls were eliminated because the current study concentrated on textual analysis.
- Short Tweets were removed (i.e. fewer than three words).
- The removal of Tweets and Retweets from official news media (e.g., Aso Rock, the Presidency, etc.) because the majority of these Tweets contained only official announcements.
- Duplicate entries were removed.

Thereafter, a total of 272,225 Tweets were generated as a result of the data pre-processing (i.e. 176,950 English, 27,223 Hausa, 19,051 Yoruba, 49,001 Igbo). The basic breakdown of the Tweets gathered after processing is shown in Table 1. Retweets exceeded original Tweets in most cases; nonetheless, this is a regular occurrence on Twitter, where the majority of users tend to repost messages that they find significant or relevant (Ahmed, Jaidka, & Cho, 2018). A sample of about 40% of Tweets from each of the languages was thereafter chosen for further analysis, yielding a total of 108,889 (i.e. 70,780 English Tweets against 50,109 Tweets from other Nigerian languages) as shown in table 1. This is in line as prescribed by (Nworgu, 2016), who stated that the sample size of 40% from a population is enough to ensure representation of the population that ranges from 2000 (Two thousands) and above. It is therefore on this sample Tweets that our report was based on.

	Before Processing				After Processing				Final Sample Tweets			
Language	English	Igbo	Hausa	Yoruba	English	Igbo	Hausa	Yoruba	English	Igbo	Hausa	Yoruba
Tweets	135,652	22,367	14,391	10,785	108,765	36,985	19,590	12,030	43,506	14,794	7,836	4,812
Re-Tweets	66,769	33,688	16,751	11,014	68,185	12,016	7,633	7,021	27,274	4,806	3,053	2,808
Total	**202,421**	**56,055**	**31,142**	**21,799**	**176,950**	**49,001**	**27,223**	**19,051**	**70,780**	**19,600**	**10,889**	**7,620**
Grand Total	**311,417**				**272,225**				**108,889**			

Table 1: Dataset for English and Nigerian Language Tweets (N= 108,889): Source: Result of Data Processing

3.3. *Step 3 - Data Annotation*

The 38,109 Tweets in the three Nigerian languages (i.e. 19,600 Igbo, 10,889 Hausa and 7,620 Yoruba) were then forwarded to human experts for annotation and further pre-processing. The Tweets were manually labeled as positive, neutral, or negative by three linguistic specialists. They were also provided a sample annotation to use as a guide. The linguistics specialists explored all the Tweets (i.e., Igbo, Hausa, Yoruba) and reported 2 different interpretations towards a similar value of intercoder reliability. This possibly is because the percentage of data used in the intercoder reliability test and the identity of intercoder varies. The rationale for the intercoder reliability otherwise known as intercoder agreement is simply to determine the extent to which independent coders evaluate a characteristic of a message or tweet and reach the same conclusion (Tinsley & Weiss, 2000). However, this paper reported a replicability of coding only in conflicting circumstances (i.e. Expert 1 = positive; Expert 2 = negative) was the third expert's remark used. There were no occurrences with three labels (Expert 1 = positive, Expert 2 = negative, and Expert 3 = neutral).

3.4. *Steps 4 and 5 - Additional pre-processing and vectorization*

Because machine learning techniques require further preparation of textual input, steps 3 and 4 on the graphical methodology in figure 1 were applied to the English Tweets. Stop words (e.g., and, is, this, etc.) were removed, followed by tokenization, which divides a given text into smaller segments or tokens where strings like "OBIdient" for example gets translated into "OBI" and "dient,". Taking a stem (i.e. removing suffixes or prefixes: e.g. electing - elect). Lemmatization was also done, followed by POS tagging, which assigns each word to a grammatical category (e.g. noun, verb, adjectives, etc.) (Sun, Luo & Chen, 2017).

In calculating the probability that POS-tags will appear in diverse sets of texts, the classifier based on the POS distribution derives posterior probability. Although POS depends on n-grams, we assume conditional independence between n-gram properties and POS information in order to simplify computations as thus:

$$P(s|M) \sim P(G|s) \cdot P(T\,|S)$$

where G is a group of n-grams used to represent the message and T is a group of POS tags used in the message. The conditional independence of the n-grams is taken for granted thus:

$$\mathrm{P}(G|s) = \prod_{g \in G} \mathrm{P}(g|s)$$

Assuming POS-tags are similarly conditionally independent, we have:

$$\mathrm{P}(T|s) = \prod_{t \in G} \mathrm{P}(t|s)$$

$$\mathrm{P}(s|M) \sim \prod_{t \in G} [\mathrm{P}(t|s). \prod_{g \in G} \mathrm{P}(g|s)]$$

We then determine each sentiment's log-likelihood as thus:

$$L\left(\frac{S}{M}\right) = \sum_{g \in G}^{\infty} log\left(P\left(\frac{g}{s}\right)\right) + \sum_{t \in G}^{\infty} log\left(P\left(\frac{t}{s}\right)\right)\Big)$$

Finally, two-word embedding algorithms, Word2Vec and String2Vec, were investigated in order to convert the text into numerical representations. Word embedding employs vectors to represent words, accounting for semantic relationships and ensuring that a word that appears more frequently in the vector representation is closer.

3.5. *Step 6 - Sentiment Analysis*

The sentiment of the English Tweets was analyzed using several supervised machine learning approaches (i.e., employing labelled data), including Naive Bayes, Support Vector Machine (SVM), and Random Forest. Researchers have used these strategies to tackle a variety of categorization problems (Kaur & Balakrishnan, 2018; Yordanova & Kabakchieva, 2017). SVM, for example, is one of the most popular supervised machine learning algorithms in sentiment analysis because of its high accuracy and ability to handle large datasets (Huq, Ali, & Rahman, 2017). Word2Vec and String2Vec, two-word embedding techniques, were used to train and test all of the algorithms. The data was divided into 90-10 groups (i.e. 90% - training versus 10% - testing). Python was used to carry out all of the tasks.

3.6. *Step 7 - Evaluation*

Finally, the accuracy (number of instances correctly predicted) and F-score (harmonic mean of precision and recall), which are represented by Eqs. (1) and (2) below, were used to evaluate the performance of all three algorithms.

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

$$\textbf{F – Score} = \frac{2TP}{2TP + FP + FN} \quad 2$$

False negatives, false positives, true positives, and true negatives are represented by FN, FP, TP, and TN, respectively.

Because F-score considers both recall and accuracy, these metrics were not employed independently to evaluate the algorithms (Shaojun, Jin, Ruixu & Guijun, 2012). The F-score is a better way to assess categorization ability since it shows the weighted average of precision and recall (Giachanou, Crestani, 2016; Nguyen, Nguyen, 2020). However, better sentiment classifications are indicated by higher accuracies and F-scores as depicted in table 2.

For the Period March 2022 – June 2022			
Models	**Metrics**	**Word2vec**	**String2vec**
Naïve Bayes	Accuracy	62.75	56.15
	F-Score	60.91	54.18
		Word2vec	String2vec
Support Vector Machine	Accuracy	58.57	52.34
	F-Score	56.34	50.08
		Word2vec	String2vec
Random Forest	Accuracy	57.53	54.26
	F-Score	56.81	51.53

Table 2: Model Evaluation of Sentiment Analysis for English Language Speakers:

The sentiment analysis results for English Tweets as shown in table 2 revealed that Naive Bayes employing Word2Vec outperformed the other classification models, with an average accuracy of 62.75%. In fact, regardless of the timescales, the introduction of Word2Vec resulted in greater accuracies and F-scores for all of the models. Support Vector Machine and Random Forest, on the other hand, performed comparably well.

3.5 Content Analysis

Because there is no sentiment analysis tool for Nigerian languages, the sentiment of Nigerian speakers was assessed using a content analysis approach. Two graduate assistants conversant in the three Nigerian languages were handed the 38,109 Tweets encompassing the three timelines as part of this strategy. A sample annotation, similar to the English annotation was provided. The assistants were also asked to search the tweets for emerging themes like information, incitement, intolerance, or parody. The measure of agreement was calculated using Krippendorff's alpha which yielded an inter coder reliability of 0.83. However, because no machine learning techniques were used, there are no accuracy measurements for the Nigerian Languages Tweets.

4. Results and Discussion

This section presents the findings of the sentiment and content analyses, starting with the top themes identified without taking emotions into account, then Nigerians' perceptions based on their sentiments during this pre-election 2023 period. Finally, the most important issues for each of these feelings are discussed. It is important to highlight that the sentiment analysis results are all based on the best classification model, which is Nave Bayes with Word2Vec.

4.1. Top Trending Keywords for English and Nigerian languages (Igbo, Hausa, Yoruba) Speakers

Tables 3 and 4 indicate the top keywords for English and Nigerian language speakers, respectively, without taking into account their sentiments.

Keyword	Nos	Percentage (%)
OBIdient	14,754	20.84
PVC	7,172	10.13
Voter Registration	1,251	1.77
Corruption	11,007	15.54
APC	13,532	19.17
INEC	2,142	3.02
Election	3,261	4.61
PDP	9,064	12.8
Vote	6,628	9.35
Bandit	1,969	2.77

Table 3: The top 10 keywords for the English speakers. (N=70,780): Source: Result of Data Processing of English Language keywords

In table 3, OBIdient, APC, corruption, were among the most popular topics within the period of conducting this study. Looking closer, the keyword "OBIdient" emerged mostly, and additional examination of the sample Tweets revealed a pattern among Nigerians making jokes about the ill health of contesting politicians and the fear that votes may not count. This could be seen on the low percentage rate of vote and election. However, there were also clear trends for specific terms like INEC (Independence National Electoral Commission), PVC and Voter Registration as well as bandit which was airing throughout the country at the moment.

Keyword	Nos	Percentage (%)
OBIdient	10,119	26.55
Change	3,102	8.14
PDP	1,366	3.58
Consumption	2,397	6.28
Production	8,982	23.57
Muslim-Muslim ticket	5,168	13.56
Vote	2,209	5.8
PVC	1,105	2.9
INEC	304	0.81
Christian-Muslim ticket	3,357	8.81

Table 4: The top 10 keywords for the Nigerian Language speakers. (N=38,109): *Source: Result of Data Processing of Nigerian Language keywords*

In the case of Nigerian language speakers (Table 4), the majority of Twitter users also appeared to be predominantly supporters of Mr. Peter Obi, as seen by the high occurrences of 'OBIdient' – 10,119, (i.e., let's support Peter Obi) during the period. The majority of communications within this period concentrated on Mr. Peter Obi's word '*consumption to production'* mantra signifying a break in the long wait Nigerians faced for this change to happen. Indeed, many Twitter users made light of the extended wait for this change by comparing it to their own experiences waiting for white collar jobs that are not available.

However, most Nigerian language speakers' comments were loaded with jokes and sarcasm based on their previous experiences also in respect to the electoral process, the failure of smart card reader during accreditation in the 2015 general election, and the wait in result releases (frequency is unavailable because experts were not asked to identify Tweets based on sarcasm/irony). Furthermore, regardless of the timelines, all the groups' Tweets contained overtly religious and ethnicity sentiments as depicted in the Muslim-Muslim ticket and Christian – Muslim tickets, showing that the majority of Nigerian language Twitter users were racist and religious in their outlook and attitude. But, if the country undergoes more economic, religious and political turmoil, this may not be sustainable in the long run.

Although most of the keywords tended to be similar among English speakers, with the most appearing to be supporters of the ruling party, as seen by terms like 'change' mantra of the APC which refers to the fact that Nigeria entered a new era in its history in 2015 and will continue, with the term denoting a new sense of hope and change for Nigerians. However, the Independent National Electoral Commission (INEC) and PVC was a hot topic for all the parties though a quick glance at the sample Tweets reveals Nigerians' dissatisfaction with INEC's handling and the slow pace of the PVC registration.

4.2 Public Opinion Among Speakers of English and Nigerian Languages

The conclusions of the analysis undertaken to examine if public opinion changes between English and Nigerian language speakers, as well as to discover the most popular keywords based on their attitudes, are presented in this part.

Positive	%	Neutral	%	Negative	%
OBIdient	17	Nigeria	19	Help	15
PVC	15	Vote	14	People	15
Voters Register	12	PVC	13	Time	14
Votes	10	Election	9	Nigeria	11
Elections	10	Candidate	9	Vote	10
INEC	9	INEC	8	Election	9
PDP	9	Time	7	Party	8
APC	8	New Nigeria	7	Country	6
Corruption	7	Party	7	Ticket	6
Bandit	6	Ticket	7	Twitter	6

Positive	%	Neutral	%	Negative	%
OBIdient	17	Nigeria	23	Help	14
Change	16	Vote	16	People	12
PDP	15	PVC	11	Time	12
Consumption	11	Election	11	Nigeria	10
Production	11	Candidate	10	Vote	10
New Nigeria	8	INEC	6	Election	9
Hope	6	Time	6	Party	9
Hope	6	New Nigeria	6	Country	8
INEC	5	Party	5	Ticket	8
PVC	5	Ticket	5	Twitter	7

Table 5: Sentiment analysis and top 10 keywords for the English Language and Nigerian Language speakers

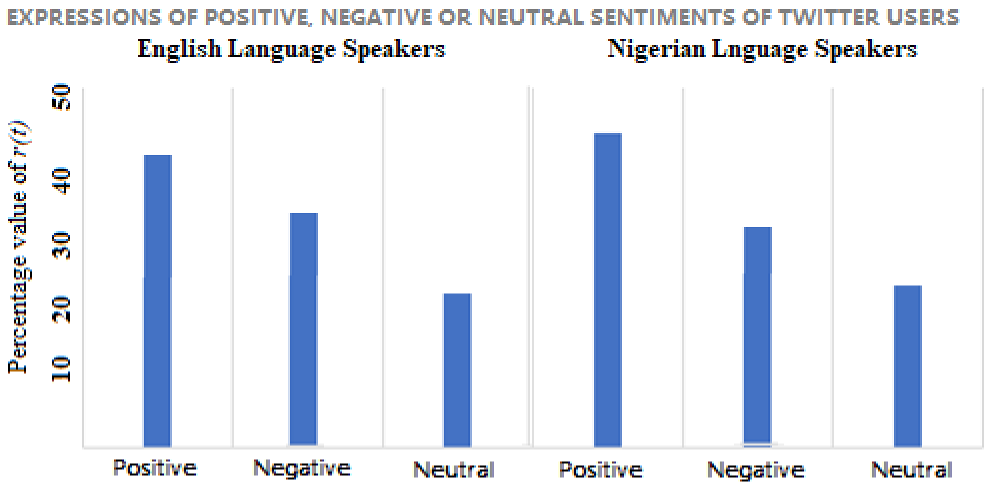


Figure 1: Cumulative graph in figure 1 showing the percentage value of *r(t)*

Table 5 shows the sentiment analysis of the top 10 keywords for the English Language and Nigerian Language speakers followed by the cumulative graph in figure 1 showing the percentage value of *r(t)*. From the graph, despite the fact that public opinion appears to be overwhelmingly supportive for both languages, the majority of Nigerian language speakers appear to be opposed to the ruling party (i.e. 44% and 45% for the English language and Nigerian language tweets respectively). The majority of the Tweets were also official in nature, with the least amount of neutral emotion observed (e.g. candidates). Election, INEC, PVC and vote were frequently used in both factions' messaging, as was to be expected given the nature of the event. According to keywords like vote, there appears to be a strong sense of nationalism among English speakers over the fate of their country (a possible huge reminder for Nigerians to vote wisely for the sake of the country).

When examining the unfavorable feelings for all the groups, it can be seen that INEC was cited frequently, which corresponds to the data reported in Table 5. This is likely due to a number of reasons, including INEC's actions in making it harder for PVC registration within the period as a result of lack of widespread of ad hoc staff as well as network failures at different registration points, and their insistence on completing the registration within one month instead of extending the exercise for at least three months' as pleaded by the House of Representatives, so that many people can register. However, if INEC insists, this could result in a very slim chance of disenfranchisement of eligible voters. These many saw as unfair. Simply put, INEC was one of the most despised institutions at the moment, as evidenced by the following sample tweets:

Speakers who speaks English (negative sentiment) - *I believe we can all agree that INEC is a filthy disgrace to our country and rights as Nigerians, regardless of our political views.*
Nigerian language speakers (bad emotion) - *Please, INEC, be democratic and transparent; what good is it if all you do is pointless?*

In contrast, many Nigerians have however made jokes about this network failures, indicating that they have not forgotten about the tragedy. "*Guys, prepare ready with your hotspots, it might suddenly blackout*," as an example of a translated tweet.

We further examined the sample Tweets based on the top keywords identified, and it can be concluded that, despite minor differences in sentiments between English and the other Nigerian language speakers, both showed patriotism toward the country by primarily communicating about providing basic amenities like free transportation, security as well as financial support to those who need to travel to their home states to vote and arranging carpool sessions prior to the election, thus making mockery about the electoral process. All the political parties have strong support base, all working hard to win. However, our effort here is to present a model of this kind that will help in determining the pattern of voters with the aim of forecasting winners in elections especially the incoming 2023 general election in Nigeria.

5.0. Conclusion and Future Direction

Using sentiment and content analysis, the study looked at Twitter communication during pre - Nigeria's 2023 general election. With accuracy of 62.75 %, Naive Bayes with Word2Vec outperformed the other classification models. Overall, findings show that overall sentiments were positive throughout the study period in favour of the Labour Party which led to the candidate of the Labour Party scoring the highest votes in the election. However, though negative sentiments appeared to be slightly higher among the Nigerian language speakers than among English language speakers against the ruling APC, the majority of the top terms were connected to voting (i.e. more of an encouragement), carpooling, and congratulatory sentiments to the Labour Party (LP) candidates. Also when it came to negative sentiments, the focus was primarily on the Independent National Electoral Commission (INEC) because of their previous conducts of elections as well as their inept management of the PVC registration process. Finally, our findings shed light on a unique component of the election, providing valuable insights and directions for future research on political usage of social media, particularly in developing countries like Nigeria in forecasting election winners. To the best of our knowledge, this is the first study in the literature to look into the predictive effects of Twitter in election victory in Nigeria's political history. This is therefore how our study contributes to the existing knowledge.

5.1. *Limitations*

The research is not without flaws. The Tweets were first obtained using popular hashtags. Though this removed prejudices as compared to utilizing single hashtags as in Ahmed, Jaidka, & Cho, (2018), it may also be beneficial to investigate the themes and discourses around individual hashtags. A hashtag linked to a specific politician, such as Ahmed Bola Tinubu, for example, or a political party, such as Labour Party, could reveal useful information about Nigerian Twitter users' perceptions and sentiments about that entity. Second, past study has demonstrated that Twitter variables such as the number of followers (and thus popularity), the number of retweets, and user engagement are all linked to good moods (Chu, Majmundar, Allem, Soto, Cruz, & Unger, 2019). Future research could look at these linkages because they were not examined in the current study.

Third, because social media data contains a lot of noise, such as spam and other bot activity, pre-processing it before any classifications or regressions is essential. Despite the fact that typical precautions were taken in this study, no specific responsibilities were assigned, such as checking for spam activity. This is something that could be looked upon in the future. It is also worth mentioning that the purpose of the current study was to evaluate Nigerians' moods between the period March 2022 to February 2023; (i.e. during the run-up to the 2023 general election), therefore the models were not tested for future mood forecasts in Nigeria yet. The findings of this study will therefore serve as a forecast model for the future elections in Nigeria. However, in the future, this is an exciting research topic to examine. Finally, because the study focused mostly on electorates that used twitter, the findings should not be extended to all Nigerians.

6. *Funding*

This study was not supported by any funding authority

6.2. *Conflict of Interest*

The authors have no conflicts of interest to declare.

References

[1] Acemoglu, D & Robinson, J.A. (2013). Why Nations fail: The origins of power, prosperity and poverty. London: Profile Book.
[2] Ahmed, S., Jaidka, K. & Cho, J. (2018) "Do birds of different feather flock together? Analyzing the political use of social media through a language-based approach in a multilingual context", *Computers in Human Behavior* 86 (2018) 299-310.
[3] Al-Ansari, H. (2006). Internet Use by the Faculty Members of Kuwait University. The Electronic Library 24 (6): 791-803.
[4] Alegu, J.C., Maku, B.S., Adelaja, K.I., & Rasheed, A.A. (2019). Social media, propaganda and electioneering campaigns: The Nigerian experience (1999 – 2019). *International Journal of Sociology and Political Science,* 2,1. Pp 4-13
[5] Almond, G.A, (2000). The Study of Political Culture. In: Crothers, L., Lockhart, C., (eds) Culture and Politics. New York: Palgrave Macmillan.
[6] Baek, Y.M and Wojcieszak, M.E (2009). Don't Expect Too Much! Learning from Late Night Comedy and Knowledge Item Difficulty. *Communication Research.* 36,6. https://doi.org/10.1177/0093650209346805
[7] Baumgartner, J.C and Morris, J.S (2006). The Daily Show Effects: Candidates Evaluations, Efficacy, and American Youth. *American Politics Research.* 34,3. 341-367. Doi:10.1177/1532673x05280074
[8] Berners-Lee, T. (2000). Weaving the Web: The Original design and ultimate destiny of the World Wide Web by its inventor: Harper San Francisco.
[9] Burnap, P., Gibson, R., Sloan, L., Southern, R. & Williams, M. (2016) "140 characters to victory? Using Twitter to predict the UK 2015 General Election Electoral Studies" 41, 230-233.
[10] CBN, (2015). Modeling The Impact of Macroeconomic Uncertainty on the Conduct of Monetary Policy. https://www.cbn.gov.ng/out/2017/rsd/modeling%20the%20impact%20of.pdf

[11] Chu, K.H., Majmundar, A., Allem, J.P., Soto, D.W., Cruz, T.B., & Unger, J.B. (2019). Tobacco use behaviors, attitudes, and demographic characteristics of tobacco opinion leaders and their followers: Twitter analysis. Journal of medical Internet research, 21(6), e12676.
[12] Dave, K., Steve, L., and Pennock, D.M. (2003). Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In WWW '03: Proceedings of the 12th international conference on World Wide Web, pages 519–528, New York, NY, USA. ACM.
[13] Delli, C (2000). Gen Com: Youth, Civic Engagement, and the new information Environment. *Political Communication*, 17,4. Pp. 341-349. Doi:10.1088/10584600050178942
[14] Doris,D.S. (2021). Leading Social Media Platforms in Nigeria 2021. *Statista*.
[15] Enli, G. (2017). "Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US Presidential election". *European Journal of Communication*, 32(1), 50-61.
[16] Eze, N., Olive, N., Ada, O. (2017). Impact of Digital-Divide on Access and Use of Information Communication Technology (ICT) Resources for Research Productivity among Students of Universities in South East Zone of Nigeria. *Computer Engineering and Intelligent Systems,* 8(6).
[17] Eze, N.U., Obichukwu, U.P., and Nnadi, H.S., (2018). Competency-based learning (CBL) and animation-based learning: a comparative study on students' academic achievement, interest and retention in auto mechanics. *International Journal of Engineering &Technology*, 7 (4), pp 6558-6564. doi: 10.14419/ijetv7i4.28755
[18] Fang, X., Zhan, J., (2015). Sentiment Analysis Using Product Review Data. *Journal of Big Data, 2(1)*. Doi:10.1186/s40537-015-0015-2
[19] Giachanou, A. Crestani, F. (2016) "Like it or not: A survey of twitter sentiment analysis methods", *ACM Computing Surveys (CSUR)*, 49, 28
[20] Gidengil, E., Wass, H., & Valaste, M. (2016). Political socialization and voting the parent–child link in turn out. *Political Research Quarterly*,69(2), 373–383
[21] Go, A., Huang, L and Bhayani, R (2009). Twitter sentiment analysis. Final Projects from CS224N for Spring 2008/2009 at The Stanford Natural Language Processing Group
[22] Hartley, D. (2007). Use and Users of Digital Resources. EDUCAUSE Quarterly 30 (4): 12-20. Available at http://www.educause.edu/ir/library/pdf//EQM0742.pdf. (Accessed 20 June 2011).
[23] Huq, M. R., Ali, A., & Rahman, A. (2017). "Sentiment analysis on Twitter data using KNN and SVM". *International Journal of Advanced Computing Science and Application*, 8(6), 19-25.
[24] INEC, (2022). List of Candidates for 2023 General Election. https://inecnigeria.org/list-of-candidates-for-2023-elections/
[25] Iyengar, S and Jackman, S.D (2004). Technology and Politics: Incentives for youth Participation. *Political Participation,* 24
[26] Jennings, M.K and Zeitner, V. (2003). Internet Use and Civic Engagement: A longitudinal Analysis. *Public Opinion Quarterly,* 67,3. 311-334.
[27] Jindal N, Liu B (2008) Opinion spam and analysis. In: Proceedings of the 2008 International Conference on, Web Search and Data Mining, WSDM '08. ACM, New York, NY, USA. pp 219–230
[28] Joseph, J., (2022). Nigeria: Online usage penetration 2016-2026. Retrieved from https://*www.statista.com/statistics/484918/internet-user-reach-nigeria/*.
[29] Kamal, A.S., Sajid, Y.Z., & Iftikha, H.J., (2015) Competence and usage of Web 2.0 technologies by higher education faculty of Education.
[30] Karleen, J.W. (2020). The Importance of Candidate and District Characteristics: A Study of Pachakutik Campaigns in Ecuador. *Oxford Academic*, 96-146. https://doi.org/10.1093/oso/9780190068844.003.0004
[31] Kaur, W. & Balakrishnan, V. (2018) "Improving sentiment scoring mechanism: a case study on airline services", *Industrial Management & Data Systems*, 118, 1578-1596.
[32] Kreiss, D. (2016). "Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle". *New Media & Society*, 18(8), 1473-1490.
[33] Krueger, B.S, (2002). Assessing the Potential of Internet Political Participation in the United States: A Resource Approach. *American Politics Research.*30, 476-498.
[34] Krueger, B.S, (2006). A Comparison of Conventional and Internet Political Mobilization. *American Politics Research,* 34,6. https://doi.org/10.1177/1532673x06290911
[35] Lazarus, S., Button, M. (2022). Tweets and Reactions: Revealing the Geographies of Cybercrime Perpetrators and the North-South Divide. *Cyberpsychology, Behavior, and Social Networking*, 25,8, p.504-511.
[36] Lee, K. S., Wong, F. F., & Azizah, Y. A. (2010). "The English language and its impact on identities of multilingual Nigerian undergraduates". *GEMA Online Journal of Language Studies*, 10(1), 87-101.
[37] Lee. S., & Xenos, M. (2019) "Social distraction? Social media use and political knowledge in two U.S. Presidential elections", *Computers in Human Behavior*, 90 (2019) 18–25.
[38] Liu B (2010) Sentiment analysis and subjectivity. In: Handbook of Natural Language Processing, Second Edition. Taylor and Francis Group, Boca
[39] Liu B (2014) The science of detecting fake reviews. http://content26.com/blog/bing-liu-the-science-of-detectingfake-reviews/
[40] Lupia, A and Philpot, T.S, (2005). Views from inside the net: How websites Affects Young Adults' Political Interest. *The Journal of Politics,* 67,4.: 1122-1142. Doi:10.1111/i.1468-2508.2005.00353.x
[41] Malik, A Heyman-Schrum, C., & Johri, A. (2019). Use of Twitter across Educational settings: A review of the literature. International Journal of Educational Technology in Higher Education, 16(36). https://doi.org/10.1186/s41239-019-0166-x
[42] Miriam, H. (2019). Solidarity as a Field of Political Contention: Insights from Local Reception Realities. *Social Inclusion,* 7(2): 74. DOI:10.17645/si.v7i2.1975
[43] Mohammed, S. M., Zhu, X., Kiritchenko, S., & Martin, J. (2015). "Sentiment, emotion, purpose, and style in electoral tweets". *Information Processing & Management*, 51(4), 480–499.
[44] Mohan, D.M., Mihir, N.M. (2022). Verbal Sentiment Analysis and Detection using recurrent neural network: Hybrid Computational Intelligence for Pattern Analysis. *Advanced Data Mining Tools and Methods for Social Computing, 2022,* pp 85-106.
[45] Monogbe, B. O., Monogbe, T.G. (2019). ASUU Strike and Nigerian Educational System: An Empirical Investigation of the Nigerian Tertiary Institution. *American Journal of Social Sciences and Humanities*, 4(1): 56-67.
[46] Mukherjee A, Liu B, Glance N (2012) Spotting fake reviewer groups in consumer reviews. In: Proceedings of the 21st, International Conference on World Wide Web, WWW '12. ACM, New York, NY, USA. pp 191–200.
[47] Nguyen, H.T. Nguyen, L.M. (2020) "ILWAANet: An Interactive Lexicon-Aware Word-Aspect Attention Network for aspect-level sentiment classification on social networking", *Expert Syst. Appl*., 146, 113065.
[48] Nwabueze, P.O (2020). Impact of Corruption On Nigeria's Economic Development. *International Journal of Engineering Applied*

Sciences and Technology, 5, 3. Pp.39-44

[49] Nworgu, B.G (2016). *Educational research: Basic issues and methodology (Revised Edition). Nsukka: University Trust publishers.*

[50] Olubusoye, O.E., Salisu, A.A. & Olofin, S.O. (2022). Youth Unemployment in Nigeria: Nature, causes and solutions. *Quality and Quantity*. https://doi.org/10.1007/s11135-022-01388-8

[51] Ososuakpor, J. O (2021). The Effects of Market and Macroeconomic Uncertainties on Corporate Investment Decisions in Nigeria. *International Journal of Management and Sustainability*, *10*(4), 104–113. https://doi.org/10.18488/journal.11.2021.104.104.113

[52] Pak, A., Paroubek, P. (2010) Twitter as a corpus for sentiment analysis and opinion mining. In: Proceedings of the Seventh conference on International Language Resources and Evaluation. European Languages Resources Association, Valletta, Malta

[53] Pang, B.O., Lee, L and Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 79–8

[54] Pang, B.O and Lee, L. (2008). Opinion mining and sentiment analysis. Found. Trends Inf. Retr., 2(1-2):1–13

[55] Pramuk, J. (2016). "Trump spent about half of what Clinton did on his way to the presidency". CNBC. Retrieved from *http://www.cnbc.com/2016/11/09/trump-spent-about-half-of-what-clinton-did-on-his-way-to- the-presidency.html*

[56] Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In ACL. The Association for Computer Linguistics.

[57] Rogers, E.M. (1962). Diffusion of Innovations. New York: Free Press of Glencoe

[58] Schone, J.P., Parkinson, B. & Goldenberg, A. (2021). Negativity Spreads More than Positivity on Twitter After Both Positive and Negative Political Solutions. Affective Science, 2. Pp 379-390. https://doi.org/10.1007/s42761-021-00057-7

[59] Shaojun, H., Jin, C., Ruixu, G., & Guijun, W. (2012, December). The Capability Analysis on the Characteristic Selection Algorithm of Text Categorization Based on F1 Measure Value. In *2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control* (pp. 742-746). IEEE.

[60] Stanford (2014) Sentiment 140. http://www.sentiment140.com/

[61] Statista. (2018). "Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2018 (in millions)". Retrieved from https://*www.statista.com/statistics/282087/number-of-monthly-active-twitter-* users/.

[62] Sun, S., Luo, C., & Chen, J. (2017). "A review of natural language processing techniques for opinion mining systems". *Information Fusion*, 36, 10-25.

[63] Tinsley, H. E. A. & Weiss, D. J. (2000). Interrater reliability and agreement. In H. E. A. Tinsley & S. D. Brown, Eds., *Handbook of Applied Multivariate Statistics and Mathematical Modeling,* pp. 95-124. San Diego, CA: Academic Press.

[64] Toffler, A., (1970). Future shock. Random House, New York

[65] Twitter (2014) Twitter apis. https://dev.twitter.com/start

[66] Valenzuela, S., Correa, T., & Gil de Zúniga, H. (2017). "Ties, likes, and Tweets: Using strong and weak ties to Explain differences in protest participation across Facebook and twitter use". *Political Communication*, 35(1), 117-134.

[67] Wankhade, M., Rao, A.C.S, & Kulkarni, C.A. (2022). A Survey on Sentiment Analysis Methods, Applications, and Challenges. *Artificial Intelligent Review*. https://doi.org/10.1007/s10462-022-10144-1

[68] Xenos, M (2008). New Mediated Deliberations: Blog and Press Coverage of the Alito Nominations. *Journal of Computer-Mediated Communication,* 13,2. 485-503. https://doi.org/10.1111/i.1083-6101.2008.00406x

[69] Xenos, M. A., Scheufele, D. A., Brossard, D., Choi, D. H., Cacciatore, M., Yeo, S., et al. (2018). "News media use and the informed public in the digital age." *International Journal of Communication*, 12, 706–724.

[70] Yadav, A., Vishwakarma, D.K. (2020). Sentiment Analysis using Deep Learning Architecture: a review. *Artificial Intelligence Review*, 53(6):4335-4385

[71] Yang, C., Hsin-Yih Lin, K and Hsin-Hsi C (2007). Emotion classification using web blog corpora. In WI '07: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, pages 275–278, Washington, DC, USA. IEEE Computer Society.

[72] Yaqub, U., Chun, S.A., Atluri, V. & Vaidya, J. (2017a) "Analysis of political discourse on twitter in the context of the 2016 US presidential elections", *Government Information Quarterly 34 (2017)* 613–626.

[73] Yaqub, U., Chun, S., Atluri, V., & Vaidya, J. (2017b). "Sentiment based analysis of tweets during the US presidential elections". *Proceedings of the 18th annual international conference on digital government research* (pp. 1–10). ACM.

[74] Yordanova, S., & Kabakchieva, D. (2017). "Sentiment Classification of Hotel Reviews in Social Media with Decision Tree Learning". *International Journal of Computer Applications*, 158, 1-7.

[75] Yousif, A., Niu, Z., Tarus, J.K., Ahmed, A. (2019). A Survey on Sentiment Analysis of Scientific Citations. *Artificial Intelligence Review* 52(3):1805-1838

[76] Yue, L., Chen, W., Li, X., Zuo, W., Yin, M., (2019). A Survey of Sentiment Analysis in Social Media. *Knowledge Information System*, 60(2): 617-663

Authors Profile

Dr. Nicholas Ude Eze received his B.Sc degree in Computer Education from University of Nigeria, Nsukka in 2004 and MSc degree in Computer Science from MS University, India in 2012. He also holds a PhD in Computer Science from MS University India and currently a Senior Lecturer in the Department of Computer and Robotics Education, University of Nigeria, Nsukka. His areas of interest include Empirical and Evidence-based Software Engineering, Software Process Improvement, Organizational Learning, Robotics and Artificial Intelligence, Machine Learning and Human Computer Interaction.

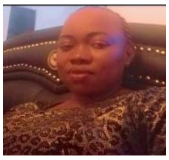
Mrs. Ifeoma Joanes Onodugo is a chattered Human Resources Practitioner with several years of experience in the banking industry. She has led various teams in providing Human Resources Services to high end clients in the area of training and development, recruitment of top management personnel, and preparing strategic HR documents. She is currently a Lecturer in the Department of Management, University of Nigeria, Enugu campus (UNEC). She is finishing up her PhD program in Management from the same University. She holds an MSc degree in Human Resources Management from Sheffield Hallam University UK in 2010, a PGD in Business from Nottingham Trent University, UK in 2009, and a HND in Management from Lagos State Polytechnic, Lagos in 2005. She is a member of Chattered Institute of Personnel Development (CIPD), UK in 2010. Her research interest is on human resource management and management sciences.

Mrs Stella Ifeoma Osondu obtained B.Sc and M.Ed degrees in Computer Education from University of Nigeria, Nsukka in 2005 and 2013 respectively. She is currently concluding her Doctorate Degree in Computer Education in Ebonyi State University, Abakaliki and currently a lecturer in Science Education Department, Michael Okpara University of Agriculture, Umudike, Abia State. Her research interests are on Systems Analysis, Data Analysis and information Technology.

Miss. Ekwutosi Ozioma Chukwu received her B.Sc degree in Business Education from University of Nigeria, Nsukka in 2019 and currently undertaking her MSc degree in Vocational and Entrepreneurship Education, University of Nigeria Nsukka. Her area of interest is Business Management, Management of Technology and Innovation, Entrepreneurship Development.

Mrs. Ndukanma Ngozika Agape received her B.Sc degree in Computer Education from University of Nigeria, Nsukka in 2008 and currently undertaking her MTech degree in Computer and Robotics Education, University of Nigeria Nsukka. Her area of interest is Hardware Maintenance.

Mrs. Kelechi Patricia Onyishi received her B.Sc degree in Computer Science from University of Abuja Nigeria, in 2011. She also received Post Graduate Diploma in Computer and Robotics Education from University of Nigeria, Nsukka in 2019 and currently undertaking her MTech degree in Computer and Robotics Education, University of Nigeria Nsukka. Her area of interest is software development.



Mrs. Lovina Ifeoma Eze received her B.Sc degree in Computer Science from Kano State University of Science and Technology, Nigeria, in 2020. and is currently undertaking her MTech degree in Computer and Robotics Education, University of Nigeria Nsukka. Her area of interest is Networking and Communication.