

An Improved User Opinion Classification and Analysis On Social Media and Its Impact On National Unity in Nigeria (A Case Study of Facebook Users)

Alhaji Baba Umar^{1*}, Ibrahim Sani² and Musa Wakil Bara³

¹Department of General Studies

²Department of Statistics

³Department of Computer Science

Mai Idris Aloomo Polytechnic Geidam, Yobe State, Nigeria

*Corresponding author: babazobi@hotmail.com

Unity is the key for national security which is in turn the key for national development. Unity is the major single entity that Governmental is trying to avail since the independence of Nigeria in 1960. Even though there are many measures that the government introduced to promote unity among the diversified population of Nigeria, the lingering unity still faces challenges in many aspect ranging from religion to sectionalism. The recent emergence of social networking sites such as Facebook and Twitter has made it easier for the citizens to express their ideas and opinions on national matters. The user's opinions became a field of interest to researchers. As a result, researchers have classified users' opinions as either Positive, Negative or Neutral. Each Class of opinion has impact on the post where the opinion was expressed. The major challenge for analysing user opinions on social networking sites is the volume of the data. There are currently 30 million Facebook users in Nigeria with more than 30 million posts per day. The easiest way to collect these huge data is to automate the process. This research automatically captures and classifies the opinions of Facebook users in Nigeria on selected post that on national issues, and to study the impact of their opinions towards national unity. Data were automatically collected from Facebook using Facebook Graph API Application. Naïve Bayes classifier was used to classify the user opinions. The result showed that, social media is creating wide gap among the social media users in Nigeria, and if care is no taken, it will act as threat to national unity. When the result was analysed it was discovered that most of the negative comments are from same region which shows that, opinions of Nigerians are gradually being divided. Recommendations was given at the end of the analysis on how social media would be used to improve national unity.

Keywords: Social Media Sites; Naïve Bayes; Classifiers; Facebook.

I. Introduction

Normally, surveys and interviews are the major ways that governments use to understand public attitude; however, these methods are limited to small sample sizes, closed questions, and limited spatiotemporal granularity. In order to overcome these limitations, we argue that social media data can be used to obtain more, real-time insights into public sentiments and attitudes with considerable spatiotemporal granularity. Over half of the worldwide population, including approximately 70% the populations of Nigeria, are active social media users, and social media usage has significantly increased in recent times; for instance, Facebook usage increased by 37%. Since social media data are largely unstructured, they are amenable to the application of established artificial intelligence (AI) techniques such as machine learning, deep learning (DL) (Hussain et al., n.d.) , and natural language processing (NLP) (Adeel et al., 2020) to extract topics and sentiments from social media posts.

Opinion analysis is used to figure out the different views of people, it classifies the behavior of the given text. Therefore, it is useful in places where users' views are needed to examine what they are thinking about things e.g. review of the national constitution, national restructuring or election acts. User opinion mining is also being applied for educational data mining and product mining in educational institutions and financial markets and in social media to ensure whether people like their product or not.

It is important to note that, texts alone can be misleading in opinion/sentiment mining, because people use emoticons or smileys to express their opinion. Many researchers have already worked on opinion mining or sentiment analysis of social networking sites such as sentiment classification of Facebook status updates (Hussain et al., n.d.) sentiment analysis of Facebook comments (page comments) and

twitter data to analyse the users opinion on election result, party candidate and brands and sentiment analysis on Twitter messages (Tanna et al., 2020).

Our research focuses on the opinions of Facebook users on Facebook post tagged as #ENDSARS. We created our own Facebook Application called UserOipinionAnalyzer for automatic data collection, we applied data cleaning processes to clean our data for better analysis. We later used Naïve Bayes Classification Algorithm to classify the user opinions.

The rest of the paper was organized as follows: In section 2, we present the related scientific literature concerning the sentiment analysis in Social Networking Services. Research methodology including the description of our proposed system's architecture was presented in section 3. In section 4, we presented the sentiment analysis for opinion mining in Facebook. In section 5, we present our experimental results and a discussion about them. Finally, in section 6, we present the conclusions and our future work.

II. Review of Related Literature

This section describes the scientific literature related to user opinion/sentiments on social networking sites notably the Facebook. Social Networking Sites (SNSs) or Social media sites are Web sites that allow social interactions among people from across the world. Such media sites include but not limited to Facebook, Twitter, MySpace, WhatsApp as well as gaming sites and virtual worlds such as Club Penguin. The Social Media Sites allow people to construct a profile which can be public or semi-public within a limited system, articulate a list of other users with whom they share a connection, view and traverse their list of connections and those made by others within the system (Abdulahi, Jalil, & Lumpur, 2014).

An Improved User Opinion Classification and Analysis On Social Media and Its Impact On National Unity in Nigeria (A Case Study of Facebook Users)

Social media allow less privileged people the ability to post and express their opinion online at little (or no) cost. The Social media allow for the exchange of breaking news (such as Twitter), sharing ideas with liked minds (such as Facebook and LinkedIn) and private audio/video chats (such as WhatsApp and Telegram) and even playing online games such as Smeet App. The social media sites also make the provision for rich information of human interaction and collective behaviour, that attract much attention from various disciplines like psychology, politics, computer science, economics, and other cultural aspects of societies (Hu, X., & Liu, H., 2012). With the services given by social media to almost everybody, the numbers of users keep on increasing geometrically on daily basis.

The rapid growth of number of social media users and the type of information spread is alarming to most developing countries, either because it can be used for the organization of protests by the citizens (Seegerberg & Bennett, 2011), and street demonstrations by workers of public and private organizations and many social or religious movements as stated by Arriagada, (2012). Social media are also used to initiate and organize a broad spectrum of dissention activities, including consumer boycotts. The Egypt protest and demonstration has led to the overthrow of the Hosni Mubarak government (Wiest, 2011). The 2011 twitter revolution has resulted in many protests and demonstrations in Chile (Scherman, A., & Portales, U. D., 2014). The 2014 revolution in turkey was also not unconnected with the use of social media (Alikilic & Atabek, 2012). It is therefore worth to note that, the social media has impacts on national unity. If the aforementioned countries that virtually follow almost one religion and one tribe can fall into this chaos through the use of social media sites, what then would be said about African countries that have people with multiple religions, tribes and cultures like Nigeria?

Nigeria is a country that is blessed with multiple culture, religions and ethnics. Like any other country in the world, Nigeria has a lot of Facebook users, more than 16 million Nigerians use Facebook every month. There is need to study and analyse how Nigerians response to posts and updates on social media. Especially updates that have direct impact on national unity such as posts/updates that is political, tribal, religious and/or ethnical. For example the hash-tag #BringBackOurGirls has united Nigerians on common goal, while the hash-tag #ResumeOrResign and #ENDSARS were only accepted by some geographical zones in the South but was rejected in the North.

To study, classify and analyse the opinions of Nigerians on social media, we employ a machine learning based approach using an Opinion Mining Algorithms called Naïve Bayes classification algorithm, details on Machine Learning and Opinion Mining Algorithms are described in research methodology section.

There are a number of studies concerning opinion mining or sentiment analysis on social media users. For example, Jain, A. P., & Dandannavar, P. (2016) uses Twitter data for sentiment analysis. The use of Machine Learning Approach was used to determine negative and positive opinions of twitter users. However, using only twitter data for sentiment analysis has some disadvantages that include but not limited to: 1) twitter posts (called tweets) are restricted to only 140 characters; 2) there is less number of youths on twitter compare to Facebook; 3) twitter updates are mostly based on “what’s happening-right-now” and therefore cannot accurately identify opinions of many users who are not up-to-date on social media.

In Troussas et al., (2014), the researchers used Facebook Social networking site for sentiment analysis. The reason for using Facebook according to the researchers was that, up to 5000 characters are allowed for update.

An Improved User Opinion Classification and Analysis On Social Media and Its Impact On National Unity in Nigeria (A Case Study of Facebook Users)

However, Facebook also has its own downside; 5000 characters are too much especially if the targeted population is also large.

In this research, the users' opinion would be classified into three (3); Positive opinion, Negative opinion and Neutral.

III. Methodology

In this paper, we developed a framework that shows how data is extracted from the Facebook, then processed and classified into as positive, negative and neutral opinion. The Framework consists of eight different steps shown in fig. 1, which are:

- i. Facebook App
- ii. Facebook Comments Extraction
- iii. Data Collection
- iv. Language Detection
- v. Text Processing
- vi. Feature Extraction
- vii. Classification
- viii. Opinion Analysis

a. Facebook APP

First, a Facebook application called UserOpinion Analysis is made on the Facebook developer site. This application allows users to interact with Facebook or access Facebook data. In order to create our own application, three requirements must be met. In order to proceed, we need to first create an account on Facebook, then log in, and last register on the Facebook Developer website. Following that, we made our own Facebook application..

b. Facebook Comments Extraction

Following the creation of an application, we used the Facebook application to collect the data (comments) from the Facebook Academic group. The data is accessed by the application through the Facebook Graph API. Facebook

comments include more than just words; they can also include pictures, stickers, and emoticons like :-) :(.

c. Data Collection

Emojis and text are both included in the data that we have collected. Emoticons are smileys that users use to convey their emotions or facial expressions in comments, whereas text is the hashtags or tweets that group members enter. Through innovation, emotions have come to play a significant role in communication. The data description, which includes some of the tweets and emoticons utilised in this paper, is displayed in Table 1.

d. Language Detection

Language detection was employed to identify the language in the provided content. It is among the most crucial actions. There are numerous languages on Facebook, including Hausa, English, Igbo, and Yoruba. Users are not restricted to using just one language; they can leave comments in any of the aforementioned languages. For this reason, it is important to identify the language in which comments are left. In this work, a single language—English—has been analysed, and the language of the comments has been identified using the English character set, N-gram, and word list. Each language has a distinct collection of letters that aid in language recognition. N-Gram is a model for text analysis. Size of N-gram is 1, 2, 3 and so on (E.g. n=1, 2 or 3 etc.) called unigram, bigram and trigram respectively. The collected text is split out into words after that occurrence of 1, 2 and 3 gram were counted. And word list provides the list of other words in the family. For example word is achieve and their other related words are achievement, achievable, achieving, achieved etc.

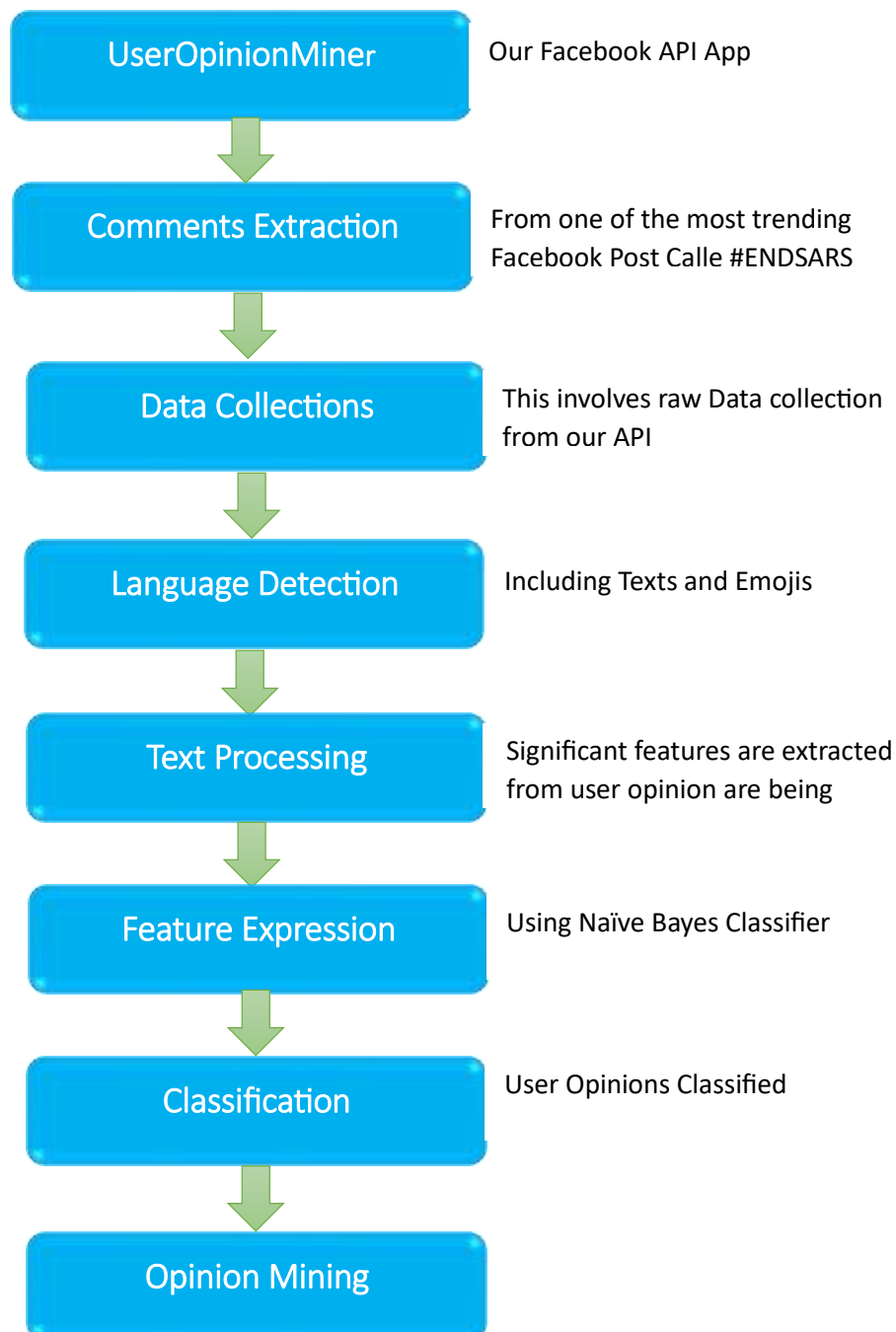


Figure 1 Data Collection Flow adopted from (Tanwani & Channa, 2017)

e. Text Processing

An Improved User Opinion Classification and Analysis On Social Media and Its Impact On National Unity in Nigeria (A Case Study of Facebook Users)

First, the material we took out of the Facebook Academic group was badly organised and included a lot of extraneous, useless language. It is in our best interests to process the data before using it in order to extract only the necessary information. Thus, the following procedures are involved in text processing: _ Removing URLs, extra spaces and special character present in the data.

_ Removing User name and user ID from the data.

_ Removing Hashtags and mention tags from the data.

_ Replacing all the upper case letters of collected data with the lower case letters.

f. Feature Extraction

It is a very important step in which, the data is sifted out. We can simply say that it is a reprocessing of the data, which gives the essential features from the gathered data. Following steps are used to filter out the data:

_ Remove all words that do not start with alphabetic letter.

_ Remove Stop words (pronoun, preposition articles, etc.) and punctuation present in the text. Stop words: is, am, are, the, at etc. Punctuations: Comas, single or double quotation mark, question mark etc.

_ Replacement of the all words into same meaning of full words which contains two or more repeated letters on them.

G. Classification

In this step, extracted features have classified using the Bayesian network to make decisions and to recognize the particular category of the tweet based on sentiment score. To give the input to the Bayesian network or to classify the

text firstly, each comment have tokenized in order to classify it in its suitable class by the given equation 1:

$$P(c|cm) = P(c)_p(w|c) \quad \text{----- (1)}$$

In this case, $P(c/cm)$ is the contingent score of class c in given comment cm , $P(c)$ is the earlier likelihood/score (prior probability) of class c , and $P(w/c)$ is the restrictive likelihood/score of token/word w in given class c . W is the words in the comment. The classifier will generate a class from the text with the highest likelihood value using equation 1. The outcome of each phrase's projected score for a specific class is reproduced by the (previous) score of that class, allowing us to choose which class another sentence should be assembled in. The classifier generates the class with the best back score after evaluating the aforementioned for each class. In the process of getting the classifier ready, those classes receive their previous class score..

h. Opinion Analysis

This stage involves analysing the specific category of a single comment, i.e., whether the comment is neutral, negative, or positive. Three groups were identified by the classifier from the data. Text containing good sentiments such as joy, love, or smile is classed as such; text containing negative sentiments such as hate, anger, or nasty tweets is classified as such; and text containing neutral sentiments such as cause, calm, or confusion is classified as such (citation).

IV. Results and Discussion

Here, firstly, we present the description of data in table 1 which contains some tweets and emoticons then show the extraction of the comments from one Facebook post tagged as #ENDSARS and perform opinion analysis on them.

Table 1: Data Description

Data Description	Positive	Negative	Neutral
Text Data	Great, Joy, Love, Smile, Success	Hell, enemy, Hate, Anger, Bad	Apology, Cause, Calm, Confusion, Sensitive
Emoticons	:) Or :-), ;), O:)	:(Or :-(, :(,	:p, :/, 8-), :D, :o

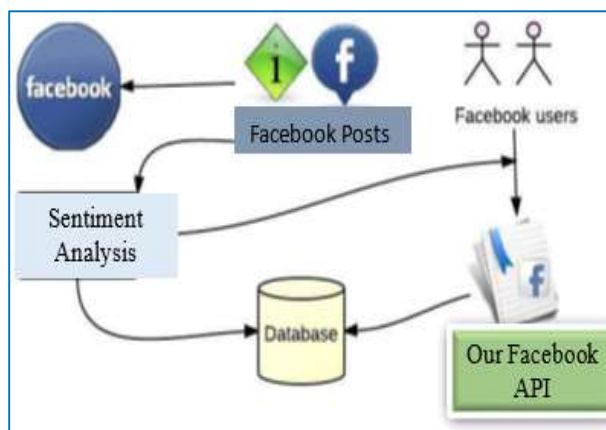
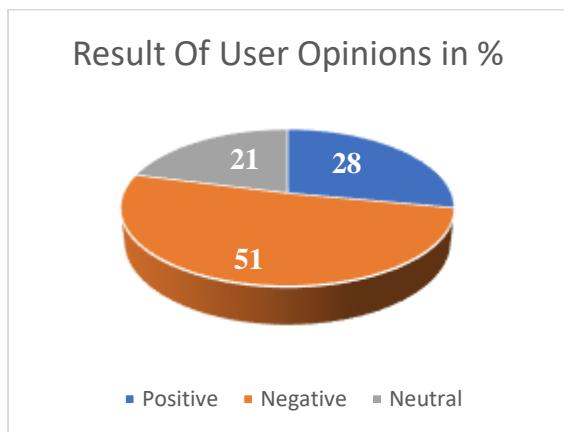


Figure 2 a & b Opinion of Facebook User on a post tagged as #ENDSARS

As shown in Fig. 2, the user opinion on the Facebook post under consideration shows that 50% of the commenters commented negatively, 28% commented positively while 21% are neutral about the post.

V. Conclusion

This research presented a novel opinion classifier called the Naïve Bayes Classification Algorithm. The algorithm was used to classify the opinions of Facebook user in Nigeria, one attracted topic tagged as #ENDSARS was selected and user options on the posts were automatically extracted, classified and analysed.

The result showed how opinions of Nigerians regarding posts are skewed based on regions. It is a known fact that Nigeria has 6 geographical regions- the Northeast, North central and Northwest which dominantly Hausa speakers, the Southwest consisting the Yorubas, Southeast occupied by the Igbos and South-South mostly the Ijaw speakers. Just like the

regions, user opinions are mostly proportional to the region pf the user.

We advise that, government can use the social media to create platforms that can unite the citizens rather than parting them apart.

VII. Acknowledgement

The researchers wish to acknowledge the Tertiary Education Trust Fund (TETFund) Nigeria for sponsoring this research work.



References

Abdulahi, A., Jalil, B., & Lumpur, K. (2014). A Study on the Negative Effects of Social Networking Sites Such as Facebook among Asia Pacific University Scholars in Malaysia, 5(10), 133–145.

Adeel, A., Gogate, M., & Hussain, A. (2020). Contextual deep learning-based audio-visual switching for speech enhancement in real-world environments. *Information Fusion*, 59, 163–170. <https://doi.org/10.1016/j.inffus.2019.08.0>

An Improved User Opinion Classification and Analysis On Social Media and Its Impact On National Unity in Nigeria (A Case Study of Facebook Users)

- Alikilic, O., & Atabek, U. (2012). Public Relations Review Social media adoption among Turkish public relations professionals : *A survey of practitioners* &. *Public Relations Review*, 38(1), 56–63. <http://doi.org/10.1016/j.pubrev.2011.11.002>
- Arriagada, A. (2012). The Social Media Basis of Youth Protest Behavior : *The Case of Chile*, 62, 299–314. <http://doi.org/10.1111/j.1460-2466.2012.01635.x>
- Hu, X., & Liu, H. (2012). Text analytics in social media. *Mining text data*, 385-414.
- Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., & Gogate, M. (n.d.). *Artificial Intelligence – Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States : Observational Study Corresponding Author* : 23(4), 1–10. <https://doi.org/10.2196/26627>
- Jain, A. P., & Dandannavar, P. (2016, July). Application of machine learning techniques to sentiment analysis. In *Applied and Theoretical Computing and Communication Technology (iCATecT), 2016 2nd International Conference on* (pp. 628-632). IEEE.
- Scherman, A., & Portales, U. D. (2014). Student and Environmental Protests in Chile : <http://doi.org/10.1111/1467-9256.12072>
- Segeberg, A., & Bennett, W. L. (2011). Social Media and the Organization of Collective Action : *Using Twitter to Explore the Ecologies of Two Climate Change Protests, 2011*(September), 197–215.
- Tanna, D., Dudhane, M., Sardar, A., Deshpande, K., & Deshmukh, N. (2020). Sentiment Analysis on Social Media for Emotion Classification. *Proceedings of the International Conference on Intelligent Computing and Control Systems, ICICCS 2020, Iciccs*, 911–915. <https://doi.org/10.1109/ICICCS48265.2020.9121057>
- Tanwani, N., & Channa, M. I. (2017). *Student Opinion Mining regarding Educational System using Facebook group*. 1–5.
- Troussas, C., Caro, J. D. L., Troussas, C., Virvou, M., Espinosa, K. J., Llaguno, K., & Caro, J. (2014). Sentiment analysis of Facebook statuses using Naive Bayes Classifier for language learning learning, (April). <http://doi.org/10.1109/IISA.2013.6623713>
- Wiest, J. B. (2011). Social Media in the Egyptian Revolution : *Reconsidering Resource Mobilization Theory*, 5, 1207–1224