



A COMPREHENSIVE DATA ANALYSIS ON FUDMA ASUU WHATSAPP GROUP CHAT

*¹Abubakar Ahmad, ¹Nuhu Abdulhafiz A. and ²Abubakar Abdulkadir

¹Department of Computer Science and Information Technology, Federal University Dutsin-Ma, Katsina State.

²Department of Computer Science Isah Kaita College of Education Dutsin-Ma, Katsina State

*Corresponding Author's: abubakarahmad82@gmail.com; aahmad1@fudutsinma.edu.ng

ABSTRACT

WhatsApp is an instant messaging application for information exchange in real time. It is a medium for communication and interaction among individuals, groups, institutions and business partners. Enormous amount of information is generated by WhatsApp in velocity, volume and variety which can serve as a source for various analyses, prediction and for other purposes. In this paper, dataset was collected from WhatsApp Group Chat, FUDMA ASUU MATTERS (FAM), a chat group of lecturers from Academic Staff Union of University (ASUU), Federal University Dutsin-Ma, Katsina state Nigeria. The primary goal is to present detailed analysis of the WhatsApp group chat to ascertain the level of involvement and participation by members in the group. Detailed analysis of fact such as the number of messages sent in different format, the most active date and time as well as the most active user(s) is to be investigated. Text classification method with Python and Jupyter notebook was used. The Python libraries applied include, Numpy, Pandas, Matplotlib and Seaborn. The result has shown that the level of participation of members compared to top ten members is by far uneven as only the top ten members accounted for more than half of the cumulative messages sent over a period of fourteen months. The research encourages members to be actively involved instead of allowing few members to dominate the platform. It is better to be an active contributor rather than remaining as a passive onlooker

Keywords: Data Mining, Sentiment Analysis, WhatsApp, Emoji, Python libraries

INTRODUCTION

Recently, internet and its associated technologies have continued to revolutionize information exchanged. Social Media sites such as WhatsApp, Facebook, Twitter and Blogs etc. are becoming indispensable platforms for all forms of communications; individuals co-operate bodies, organizations, governments among others. The need to analyze such huge generated data to determine views and sentiments for desired purposes such as whether information shared are relevant or not, or to make prediction about occurrence of an event cannot be over emphasized. In most cases, messages exchange by group members appears irrelevant as investigated by Ahmad et al. (2021). Their findings revealed that of the total messages of over sixteen thousand, only 8.7% were found to be relevant messages, which is very insignificant compared to a significant percentage of messages found to be irrelevant constituting 43.3% of the total messages posted over a period of fourteen months. The result further exposed that messages such as jokes, spam, forwarded messages, greeting messages, blank emotions, devotional messages, social event wishes, personal comments etc. dominate most social network group (Ahmad et al. 2021). These and other factors compel groups to set out rules and regulations to guide exchange of information with limited or no compliance by recalcitrant members. The process involved in discovering and extracting sentiment or opinion from within text is called opinion mining or knowledge discovery. Opinion mining is a type of natural language processing to track opinions of people about a particular event or subject (Harshal et al., 2018).

Opinion mining is the process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc., is positive, negative, or neutral. (Oxford, 2019). The goal of opinion mining is to determine the view point of individual or people towards any topic or an event by analyzing their views on social networking sites. This is possible through knowledge discovery which is synonymous to data or text mining in databases. It is a way of discovering hidden pattern or beneficial knowledge from data. The mining of huge dataset also called Big data is a trending area of research involving methods at the intersection of machine learning, statistics and mathematics. Today, Companies, organizations and individuals leverage mining techniques to look for patterns in large batches of data to improve their business, marketing strategies, learn about customers, increase sales and decrease cost among other benefits. (Court, 2015)

Classification is a process of sorting and categorizing the data into various groups, forms or any other distinct class. This classification process provides separation and categorization of data according to data set requirements for various purposes. Thus target class can be treated for each and every case in the data. Algorithms driving this data management process are termed as 'classifiers'. In machine learning and statistics, classification is a supervised learning approach in which a

program learns from the data given to it and then uses that experience to classify new observations.

The goal of this paper is to present detailed analysis of the WhatsApp group chat to ascertain the level of involvement and participation by members in the group. To achieve this, we analyzed the FAM WhatsApp group dataset and extracted the thoughts and opinions of lecturers expressed on the group through their posts and classified them into different categories, namely; text messages, media messages, website links and Emojis. The results were used to determine the level of participation of members in the group chat. The analysis involved; the most active date in the group, the number of messages sent on the most active date, the overall most active user, total number of users, number of posts made by each individual on the group, and the most used word on the platform. In addition, an analysis of the top 10 users was also done and results visualized using python code in Jupyter notebook.

The remaining parts of this paper were organized as follows; in section 2, we reviewed literatures related to this paper. Section 3 describes methodology for WhatsApp sentiment analysis utilized in this research. In section 4, we demonstrate the experimental results conducted on FAM WhatsApp datasets. Finally, section 5 concludes and suggests some recommendations.

RELATED WORK

There are so many works done to date on sentiment analysis particularly on classifying twitter messages into positive, negative or neutral and for prediction purposes. Sentiment analysis has been handled as a Natural Language Processing task at many levels of granularity, word level, sentence level or document level. (Liu, 2010.). Sentiment classification techniques can broadly be grouped into Machine Learning (ML) approach, lexicon-based approach and hybrid approach (Diana & Adam, 2011).

The ML text classification approach is divided into supervised and unsupervised learning methods. The supervised methods exploit large number of labeled training documents exploiting classifiers namely; Decision Tree, Linear classifiers, Rule-based classifiers and Probabilistic classifiers. While the unsupervised method is used when it is difficult to find labeled training documents (Li & Tsai, 2011).

For Lexicon-based approach, it relies on a sentiment lexicon, a collection of known and precompiled sentiment terms. It is classified into dictionary-based approach and corpus-based approach. While the former approach relies on finding opinion seed words, and then searches the dictionary of their synonyms and antonyms, the later approach uses statistical or semantic methods to find sentiment polarity. The combination of the two approaches (Machine Learning and Lexicon Based approaches) referred to as hybrid approach get better classification results (Mudinas et al., 2012).

Other concept-level sentiment analysis approaches such as pSenti is integrated into opinion mining lexicon-based and machine learning-based approaches. The system achieved higher accuracy in sentiment polarity classification as well as sentiment strength detection compared with pure lexicon-based systems using two real-world data sets (CNET software reviews and IMDB movie reviews). It outperformed the proposed

hybrid approach over state-of-the-art systems like SentiStrength (Mudinas et al., 2012).

Kontopoulos et al. (2013) proposed the use of ontology-based techniques toward a more efficient sentiment analysis of twitter posts. They worked on the domain of smart phones. Their architecture gives more detailed analysis of post opinions regarding a specific topic. Other techniques not categorized as either ML approach or Lexicon-based approach includes; Formal Concept Analysis (FCA) which is a mathematical approach used for structuring, analyzing and visualizing data and Fuzzy Formal Concept Analysis (FFCA) developed in order to deal with uncertainty and unclear information (Walaa et al., 2014).

In related researches, multiple techniques are combined to improve performance. Thakare and Sachin (2016) combined both Lexicon based and Machine Learning approach and achieved an increased performance in precision and recall for twitter data. Generally, bag-of-words has been used for mining sentiments online. In this approach, individual words are considered instead of complete sentences. Therefore, Traditional machine learning algorithms such as Support vector Machines, Naive Bayes' and Maximum entropy etc. are commonly used to solve such classification problems with features such as unigrams, n-grams, Part-Of-Speech (POS) tags (Paridhi et al., 2018).

Other researchers proposed enhancements of some approaches such as the ensemble model aimed to improve performance metrics of the existing algorithms like Naïve Bayes, SVM and Linear Regression model (Mohanavalli et al., 2018). Bhattacharjee et al. (2019) worked on boosting TF-IDF for effectiveness. Recently, Darwich et al. (2019) presented a comprehensive review on the notable research works that focus on the corpus-based approach for sentiment lexicon generation. The authors arrived at eight reasons in favor of corpus-based approach over a dictionary-based approach and concluded on the note that corpus-based techniques are considered as a vital part of any modern sentiment analysis system.

Recently, Cambria et al. (2020) built new version of SenticNet, called SenticNet 6; a commonsense knowledge base for sentiment analysis which integrate top-down and bottom-up learning via an ensemble of symbolic and sub symbolic AI tools for logical reasoning within deep learning architectures. This system exploits both Artificial Intelligence and Semantic Web technologies for recognizing meaningful patterns in natural language text and, hence, represent these in a knowledge base using symbolic logic.

In this research, the authors adopted the Supervised Machine Learning method of text classification in a unique way. Most of the researches focused on binary classification of sentiments, i.e. weather sentiments are positives or negatives, or even neutral in some cases. However, the authors' approach, seek to find the level of involvement and participation by members in a chat group. Find out interesting insights like the most used form of messaging, the most active date and time of the day, most active users. These would be interesting insights to uncover. To accomplish that, the dataset created in the previous work of Ahmad et al. (2021) was adopted and utilized with exclusion of "Remark" column which contained the labeled attribute in the dataset.

METHODOLOGY

In this paper, the authors exploit text classification method. There are many kinds of supervised classifiers in literatures. Some of the most frequently used classifiers in Sentiment Analysis are; Decision Tree, Linear, Rule-based and Probabilistic classifiers. Since the focus in this paper is not

prediction, Attention is not paid to classifiers. The text classification steps in this research involves; collection of data from the FAM WhatsApp, transformation of the data and then exploration, analysis and eventual visualization of the data using python in Jupyter Notebook. The processes are depicted in figure 1 below. The different phases in text classification are explained in the following subsection.

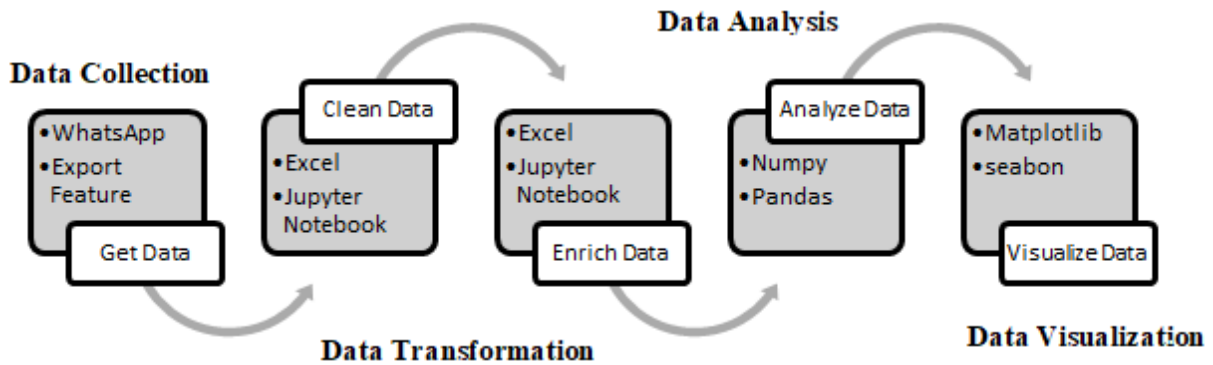


Figure 1: Text Classification Technique showing different Processes

Data Collection Process.

The first step was to obtain the data from target source. WhatsApp provides a feature for exporting chats through a .txt format; this was achieved via the chat group (FAM). In exporting the WhatsApp data file, the procedures involved the following; on the FAM group page, click on the settings, select export data and then select without media. This means that media file is excluded to narrow the scope of the work. Additionally, exporting along with the media files, will lead to use of larger volume of data and waste of time for data collection. The text file is then transferred into pandas data frame. The following depicts how the file looks like in .txt format;

```

WhatsAppChat01 - Notepad
File Edit Format View Help
9/20/18, 9:06 PM - Messages to this group are now secured with end-to-end encryption. Tap for more info.
8/17/18, 7:33 PM - +234 [REDACTED] 4555 created group "FUDMA ASUU MATTERS"
9/20/18, 9:06 PM - +234 [REDACTED] 4555 added +66 80 956 0207
9/20/18, 9:41 PM - +234 [REDACTED] 4278 left
9/20/18, 10:58 PM - +234 [REDACTED] 8706: <Media omitted>
9/21/18, 12:54 PM - +234 [REDACTED] 5764: <Media omitted>
9/21/18, 12:56 PM - +234 [REDACTED] 5764: <Media omitted>
9/21/18, 11:08 PM - +234 [REDACTED] 5764 added +234 806 817 7887
9/21/18, 9:41 PM - +234 [REDACTED] 9328: <Media omitted>
9/21/18, 10:32 PM - +234 [REDACTED] 6877: Master Scholarships at the University of Maiduguri in Nigeria, 20
the-university-of-maiduguri-nigeria-2018/2018/09/20/#.W6VIMnM0eeo.whatsapp
9/21/18, 10:40 PM - +234 [REDACTED] 6508: <Media omitted>
9/22/18, 7:15 AM - +234 [REDACTED] 1187: ???
9/22/18, 8:02 AM - +234 [REDACTED] 6877: ?
9/22/18, 8:18 AM - +234 [REDACTED] 6508: I'm so sorry colleagues this video was mistakenly sent to this pla
Thanks
9/22/18, 8:54 AM - Dr. A [REDACTED] kara: This message was deleted
9/22/18, 10:10 AM - +234 [REDACTED] 7442: Admin please add this number 08039474709 umar Musa kankara biolog
9/22/18, 10:16 AM - +234 [REDACTED] 4360: Masha Allah! Wishing you all the best.
9/22/18, 10:26 AM - +234 [REDACTED] 0804: In need of 5 young programmers - fresh graduates who are good at
A lot of travel and travel allowances.
  
```

Figure 2: FAM WhatsApp Group Chats exported to Notepad

Data Preparation and Transformation

At this stage, the dataset was uploaded to Jupyter Notebook for exploration, analysis and presentation. Details for the visualization were presented in 'Result and Discussion' section of the paper. The plain text file is parsed and tokenized in a meaningful manner to be stored in pandas data frame. The raw messages are broken down into four tokens; *date*, *time*, *phone number* (author) and

message representing the table head. A function is written that detects whether each line begins with a date and time, it then splits the line into date, time, author and message tokens. The following depicts how the table is structured in pandas data frame.

```
df.head(10)
```

	Date	Time		Author	Message
0	2018-08-17	7:33 PM	+234	4555	created group "FUDMA ASUU MATTERS"
1	2018-09-20	9:06 PM	+234	4555	added +66 80 956 0207
2	2018-09-20	9:41 PM	+234	4278	left
3	2018-09-20	10:58 PM	+234	8706	<Media omitted>
4	2018-09-21	12:54 PM	+234	5764	<Media omitted>
5	2018-09-21	12:56 PM	+234	5764	<Media omitted>
6	2018-09-21	11:08 PM	+234	5764	added +234 806 817 7887
7	2018-09-21	9:41 PM	+234	9328	<Media omitted>

Figure 3: FAM WhatsApp Group Chats in data frame

The data frame splits the line into the date, time, phone and message tokens. The data available to this research covered from 17th of August 2018 when the group was created to 21st October 2019 when the group migrated to Telegram from Whatapp. A total of sixteen thousand, three hundred and ninety nine (16,399) messages were generated and transform to form the dataset for this research.

It is important to note that the exported chat file was without a media. Therefore, the phrase "<Media omitted>" is used in place of media messages. Other messages are text, emojis and link messages. The implementation tools were explained in the following sub section.

Data Mining Platform and Tools

Python programming language was utilized in this research not only because of its simplicity and popularity but its ubiquitous collection of libraries handling various purposes such as imputation of data, transformation of data, exploration of data and data visualization. Python packages are directory of python scripts; each script is a module which can be a function, methods or new python type created for particular functionality. *Numpy* was used to handle the multidimensional arrays and functions required for the classification of chats into days, hours, minutes and seconds. *Pandas* were utilized for data extraction, manipulation, cleaning, and analysis. It was suitable for the kind of data used in this research. Another package utilized in this research is Python *Matplotlib*. *matplotlib.pyplot* is a plotting library used for 2D graphics. It was used for the various graphs plotted in this work. In addition, *Seaborn*, which

is a dominant data visualization library, was also imported in this work. Being a higher-level library, it was able to expand the plot and better beautify it. It doesn't work alone hence it works on *Matplotlib* foundation.

Ethical Considerations

It is evidently clear that the data in WhatsApp file contain phone numbers and messages sent by members of the group. There are therefore certain privacy considerations that must be taken into account. Most notably, individual phone numbers should not be collected and/or released to public. The authors sought the consent of administrators to use only the information for research purposes. As a result, the identities of individual members are hidden.

RESULTS AND DISCUSSION

Figure 4a and 4b show the distribution of the various messages in all posts of the group chat. 63.4% of all the chats amounting to 10,393 messages were *Emojis*. It is the highest with a very wide merging almost doubled the other types of messages put together. Followed by is the media with a total of 2,264 messages accounting for 13.8% of the entire messages. *Text* messages which ought to be in the fore front is proceeding with 12.8% constituting 2,096 messages, a little behind media post. The least of all the posts were *Links* posts which a total of 1,646 messages resulting to 10% of the entire posts. This leads to the conclusion that FAM group chats are mainly used to communicate via *Emojis* post over the period reviewed, while text, media and links posts were used to some extent.

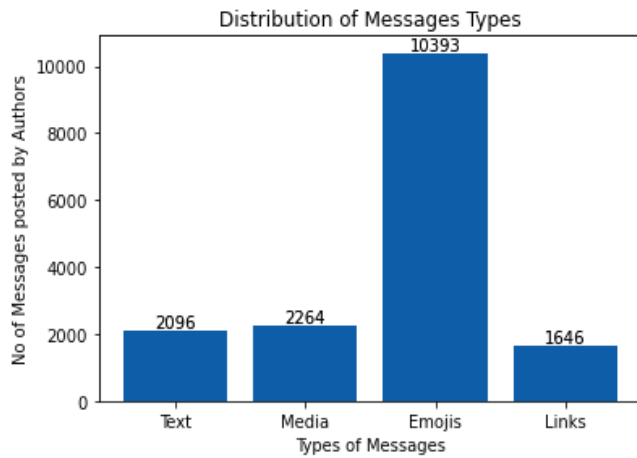


Figure 4a: Bar Chart showing distribution of various messages types

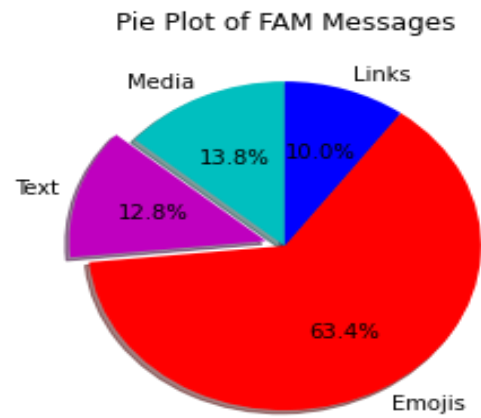


Figure 4b: Pie Chart showing percentage distribution of message types

Furthermore, the frequency of posts at a particular date was investigated. Figure 5a illustrates the dates (y-axis) in connection with the relative frequency of posts (x-axis). Most posts were sent on 15th and 20th of December, 2018 with over 350 messages and about 300 messages respectively. Further checks on the records revealed that these were the period when funds for Earned Academic Allowance (EEA) were released to universities. The graph as a whole presented the top ten most active dates over the period under review. This finding supports the statement that FAM WhatsApp constitutes a very fast communication medium among members.

Similarly, the frequency of posts at a particular time of day was investigated. Figure 5b illustrates the day times in hours and minutes (y-axis) in connection with the relative frequency of posts (x-axis). Most posts were sent around 10:34 AM and 2:36 PM. The chart shows the trend that members communicate more after 10:30 AM in the morning, 2:36 PM in the afternoon and mostly around 9:54 PM in the night.

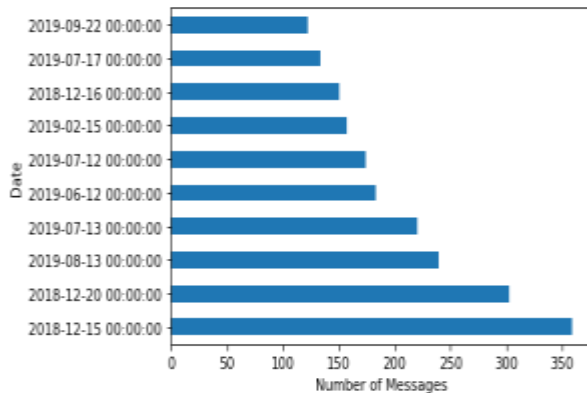


Figure 5a: A Bar chart showing top ten date with most frequent posts in FAM Chat group

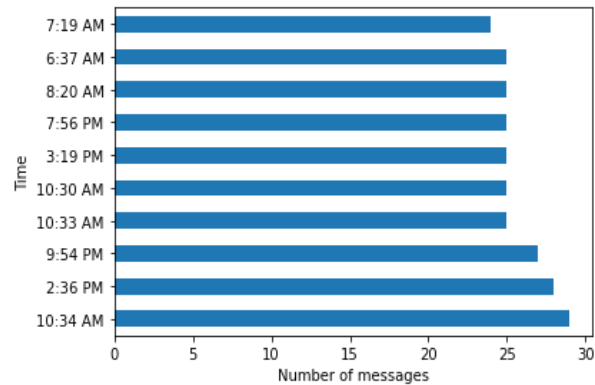
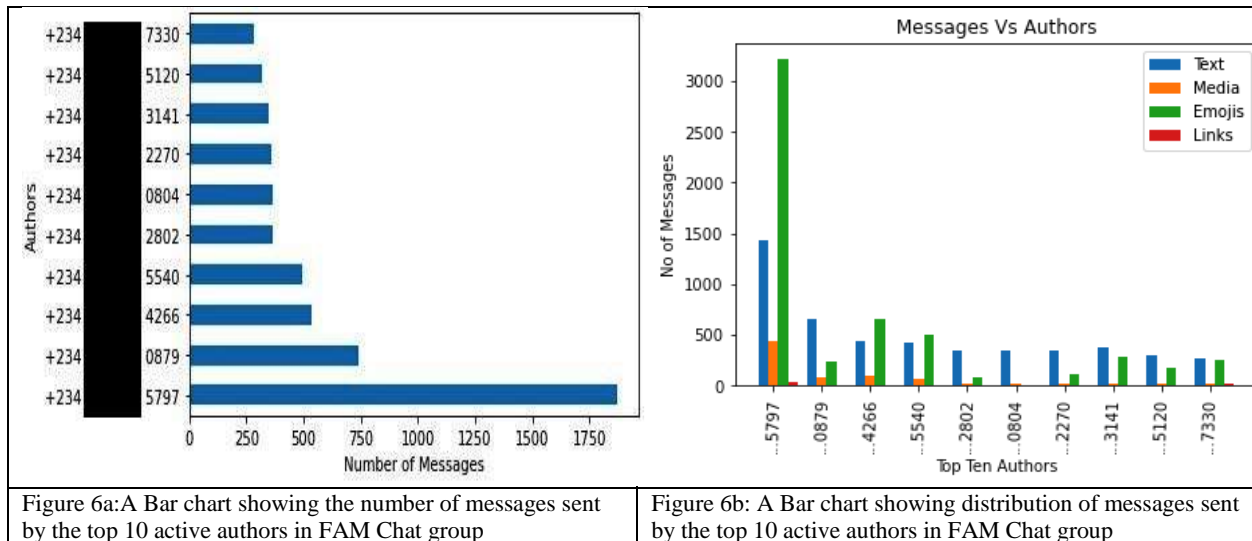


Figure 5b: A Bar chart showing top ten times in hours and minutes with most frequent post in FAM Chat group

Figure 6a depicts the total messages sent by top ten authors; it is astonishing to know that of the 16,399 cumulative messages, a total of 11,290 messages were sent by only ten top authors, meaning that by far, more than half of the messages were attributed only to top ten members. What is even more striking is the fact that of the 11,290 messages sent by the top ten members, almost half of these messages were sent by the highest contributing member. In comparison with the cumulative messages, the highest contributing member contributed more than one third of the cumulative messages, simply put 31.2% of the cumulative message over the period reviewed.

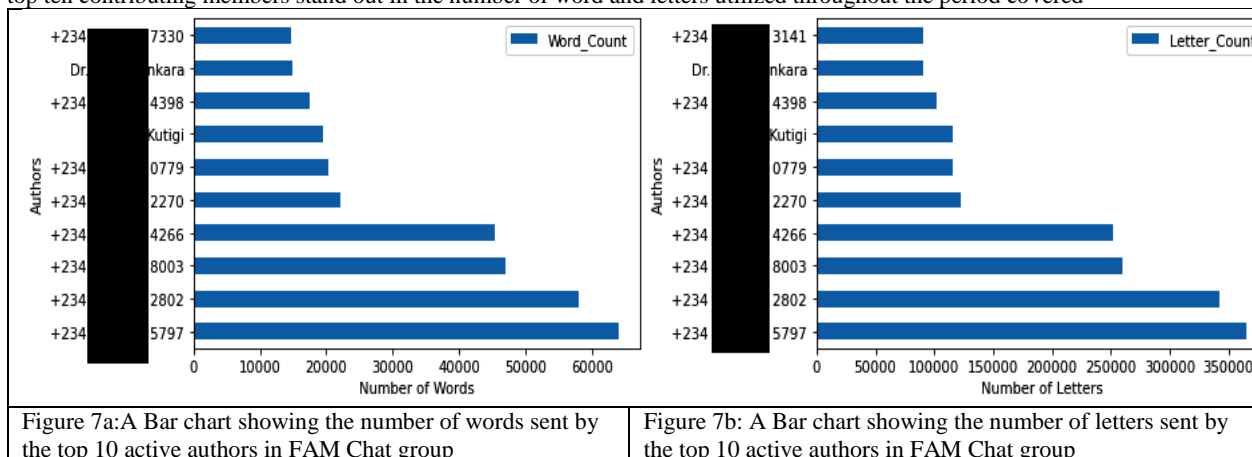


Furthermore, figure 6b shows exemplarily the highest category of messages sent by the ten top authors is the emoji with 48.8% of the total messages equivalent to (5,507) messages, this cut across other authors with the exception of few. Followed to emojis did text message constituting 43.7% (4,934) and then media messages with 6.9% (780). The least is link messages having 0.6% (69) as presented in figure 6b above. This reaffirmed the fact that members utilized emojis more than other forms of messages. Details of the statistics is presented in table 1 below;

Table 1: Details statistics of ten top Authors

Author	Messages	Media	Emojis	Links	Total
+234 XXXXXX 7330	260	21	251	13	545
+234 XXXXXX 5120	295	24	180	2	501
+234 XXXXXX 3141	372	18	278	4	672
+234 XXXXXX 2270	346	13	114	7	480
+234 XXXXXX 0804	345	17	3	0	365
+234 XXXXXX 2802	349	15	74	4	442
+234 XXXXXX 5540	429	63	504	5	1001
+234 XXXXXX 4266	444	93	658	4	1199
+234 XXXXXX 0879	658	83	232	3	976
+234 XXXXXX 5797	1436	433	3213	27	5109
Total	4934	780	5507	69	11290

In figure 7a and 7b, the distribution of number of words and letters respectively are presented. It is indicative that four out of the top ten contributing members stand out in the number of word and letters utilized throughout the period covered



SUMMARY OF FINDINGS

This paper investigated FAM group-based communication. The result shows that of the 16,399 cumulative messages, only 12.8% were found to be *text* messages, which is very insignificant compared to a significant percentage of messages found to be *emojis* consisting of 63.4% of the total messages posted, it concludes that members use emojis more than any form of messaging. Accordingly, 13.8% of the total messages were classified as *media* messages testifying that members send more videos than posting link message to the group. Similarly, the result further revealed that much traffic of messages was recorded on 15th and 20th of December 2018 as a result of Earned Academic Allowance (EAA). The result add that members communicate more after 10:30 AM in the morning, 2:36 PM in the afternoon and mostly around 9:54 PM in the night. With regards to posting on the FAM platform, an individual member alone contributed more than one third of the 16,399 messages, while top four members accounted for more than half of the entire messages. The ten top members contributed 69% of the cumulative messages confirming the fact that they are the most active members posting more often on the platform.

Future Work

In the future, there is need to extend the scope of this research to ascertain the level of non-involvement and in-activity by members in the chat group. Find out interesting insights like the most passive member. The scope of the dataset could be expanded to include multi-media messages such as voice, video, empty messages among others for a more detailed analysis.

CONCLUSION

Conclusively, WhatsApp application along with the power of python in Jupyter notebook has shown its strength as a tool for analyzing any form of dataset. This working instigated WhatsApp application, its potentials and how data from the platform could be exported, prepared, transformed, analyzed and later visualized. The analysis was done with Jupyter notebook, with the following libraries; Pandas, Numpy, Matplotlib and Seaborn. Text classification Technique was used. At the end of the work the results obtained shows comprehensive details of most active user(s), date, and time of the day and the most used format of posting messages on the FAM group platform.

RECOMMENDATION

The authors encourage members to keep to the purpose of the group, exchange information from a verified source, and involve actively in meaningful and relevant contributions in the platform. Passivity, sycophancy and praise singing by members is discouraged.

REFERENCES

Ahmad, A., Mukhtar, A., Akinyemi, O. O. (2021, JANUARY). Sentiment Analysis and Classification of ASUU WhatsApp Group Post Using Data Mining. *JOURNAL OF CONFLICT RESOLUTION AND SOCIAL ISSUES*, VOL. 1 (NO. 2), 18 - 27.

Bhattacharjee, U., Srijiith, P. K. & Maunendra, D. (2019). Term Specific TF-IDF Boosting for Detection of Rumours in Social Networks. In D. o. Engineering (Ed.), *In Proceedings of the*

Sixth Social Networking Workshop, SN@COMSNETS 2019, Bengaluru., 116 , pp. 10–19. IIT Hyderabad India.

Cambria, E., Yang, L., Xing, Z. F., et al. (2020). SenticNet 6: Ensemble Application of Symbolic and Subsymbolic AI for Sentiment Analysis. *CIKM*, (pp. 105-114). Ireland.

Court, D. (2015). *Marketing & Sales Big Data, Analytics, and the Future of Marketing & Sales*. McKinsey & Company.

Darwich, M., Mohd, N., Shahrul, A., Omar, N., & Osman, N. (2019). Corpus-Based Techniques for Sentiment Lexicon Generation: A Review. *Journal of Digital Information Management.*, 17, 296. 10.6025/jdim/2019/17/5/296-305.

Diana, M. & Adam, F. (2011). Automatic detection of political opinions in tweets. *Proceedings of the 8th international conference on the semantic web, European Semantic Web Conference (ESWC'11)* (pp. 88–99). Springer.

Harshal, K., Kalyani, G. & Tanmay, S. (2018, March 03). A review on: Sentiment polarity analysis on Twitter data from different Events. *International Research Journal of Engineering and Technology (IRJET)*, 05 (03 | Mar-2018), Page 1479.

Kontopoulos, E., Berberidis, C., Dergiades, T. & Bassiliades, N. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, 40(10), 4065-4074.

Li, S. & Tsai, F. (2011). Noise control in document classification based on fuzzy formal concept analysis. In: Presented at the IEEE. *International Conference on Fuzzy Systems (FUZZ)*. IEEE.

Liu, B. (2010.). *Sentiment Analysis and Subjectivity*. In *Handbook of Natural Language Processing, Second Edition*. Taylor and Francis Group, Boc.

Mohanavalli, S., Karthika, S., Srividya, K.R., & Uthayan, N. S. (2018). Categorisation of Tweets Using Ensemble Classification Methods. *nternational Journal of Engineering & Technology*, 7 (3.12), 722-725.

Mudinas, A., Zhang, D. & Levene, M. (2012). Combining lexicon and learning based approaches for concept-level sentiment analysis Presented at the. *WISDOM'12*. Beijing, China.

Neumann, G. (2006). A Hybrid Machine Learning Approach for Information Extraction from Free Text. *From Data and Information Analysis to Knowledge Engineering* (pp. 390 - 397). Springer, Berlin, Heidelberg.

Oxford. (2019). *Oxford online dictionary*. accessed, 12:53PM, 16th, October 2019: https://www.lexico.com/en/definition/sentiment_analysis.

Paridhi, P. N., Dinesh D. P. & Yogesh, S. P. (2018). Sentiment Classification of Twitter Data: A Review. *International Research Journal of Engineering and Technology (IRJET)*.05,

pp. 929 - 931. p-ISSN: 2395-0072: ISO 9001:2008 Certified Journal.

Patil, S. (2016). WhatsApp Group Data Analysis with R. *International journal of Computer Applications*, 0975 – 8887.

Walaa, M., Ahmed, H. & Hoda, K. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5, 1093 - 1113.



©2021 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.