


Analysis of Cyber Bullying on Facebook Using Text Mining

Nasiru I. Aliyu^{1,*} , Musbau D. Abdulrahaman¹ , Fatimah O. Ajibade¹, Tosho Abdurauf² 

¹Department of Information and Communication Science, Faculty of Communication and Information Sciences, University of Ilorin, Ilorin, Nigeria

²Department of Computer Science, Faculty of Natural and Applied Sciences, Al-Hikmah University, Ilorin, Nigeria

Received: 21.11.2020 • Accepted: 24.12.2020 • Published: 30.12.2020 • Final Version: 30.12.2020

Abstract: Cyberbullying is a form of online crime that involves the use of the internet and other information technology resources to deliberately insult, embarrass, harass, bully, and threaten people online. Nowadays, cyberbullying is becoming more rampant in our society due to the increase in the internet connectivity and social media usage among the populace. More importantly, the inability of the existing cyberbullying detection techniques to effectively differentiate between bully and non-bully words on social media such as Facebook makes detection and identification of cyberbullying more difficult, thereby, exposing more people to the life threatening crime. Text mining is one of the trending data mining techniques capable of analysing every single word in a text. The aim of this work is to analyse and predict cyberbullying words in Facebook posts and comments using Naïve Bayes based text mining technique. The technique analyses every instance of Facebook posts and predicts the occurrence as well as the category of the bully related words. To measure the performance of the technique, the score accuracy, confusion matrix as well as classification reports including precision, recall and F1 – score were employed on the classifier. The result shows that the Naïve Bayes based text mining technique has an accuracy of 0.95%. As a result, the experimental analysis shows that the technique is capable of predicting the occurrence of bully and non-bully words in Facebook posts and comments effectively and accurately.

Keywords: Cyberbullying, Text Mining, Naïve Bayes, Cybercrime, Analysis.

1. Introduction

Social networking sites are evolving and continuously gaining widespread adoption among thousands and millions of people around the world daily. The widespread adoption is due to its ability to connect people by eliminating distance among the parties and enables real-time information dissemination to a large number of people [1]. Social media is a form of computer-mediated communication that is used to connect a large number of people in different geographical locations. Anyone with enable technology and internet access can be part of social media and share opinions [2]. According to [3], Facebook is one of the popular social media platforms to reach over 1 billion active users monthly as at the third quarter of 2012 and 2.7 billion active users in the second quarter of 2020. Studies show that users spend about 35minutes on Facebook every day.

The growing use of social media worldwide exposes more people to attack from cybercriminals who normally capitalize on their vulnerabilities and private information. Cyberbullying is a form of

* Corresponding Author: aliyu.in@unilorin.edu.ng

cybercrime that involves deliberate insulting, threatening, embarrassing, or harassing people on the internet through smartphone applications [4]. In another words, it involves the use of computers and the internet to inflict harm on people by sending offensive, abusive and demoralizing messages which reduce the morale of the victim. The rate at which digital crime occurs leaves little hope for human intervention to prevail without human error [5]. Cyberbullying usually takes place on several Information and Communication Technology (ICT) platforms such as social websites, chat forums like Facebook. Some of the popular cyberbullying activities or crimes include: harassment, impersonation, and cyber stalking. Most of the times, this could cause the victim to feel insecure and Commit suicide. The issue of cyberbullying is becoming a major threat to every individual who uses social media, especially Facebook for various activities such as entertainment, marketing or businesses. Cyberbullying is different from traditional bullying because its perpetrator can go undetected due to the fact that it is usually committed online where it does not require the physical presence of the criminal or victim. The issue of cyberbullying is becoming rampant leading to low self-esteem and demoralization of its targets [6, 7]. Hence there is a need for a more flexible and robust approach for solving this ever-growing cybercrime

Text mining is an emerging technique that involves extraction of hidden, previously unknown, and potentially useful information from a large amount of textual data. It is the process of structuring the input text, deriving patterns within the structured data, and finally evaluates and interprets the output [1]. The use of text mining technique on Facebook can uncover bully words, thereby leading to a tremendous reduction in the number of cyberbullying activities on Facebook. Therefore, this study proposed a text mining-based technique for curbing the menace of cyberbullying using data from Facebook users for analysis for the purpose of reducing this ever-growing crime.

2. Related Work

Recently, several studies related to cyberbullying analysis and identification based on text mining on social media have been carried out. The study of [8] analysed cyber-crime on social media using data mining techniques. The mined data was collected from Facebook pages and pre-processed to reduce noise, incomplete and inconsistent statistics. After which a clustering technique was adopted to fetch information about hackers. It uses K-means clustering algorithm to get the structured data from unstructured databases. It concludes that data mining technique is effective for intelligent analysis of cybercrimes. In another study carried out by [9], the authors explored how social media monitoring and the application of text mining technique can be used in curtailing crimes in Nigeria. It explored text mining, its basic concepts, different application areas and an approach to curb cyber-crime by extracting relevant information using the inductive approach. It gives details on cybercrime, the need for social media monitoring as a way of curbing cyber and related crime.

More so, [10] created a response grading system that mapped bullying on Twitter to high, medium and low. It pops up a message when it predicts a bully message is about to be sent. A Support Vector Machine (SVM) was used to classify the data into a bully or non-bully post and produced an accuracy of 6.88%. A sentimental analysis system was then used to perform classification on words that were not previously classified. [11] Analysed the rate of cyber bullying in Indonesia on twitter and what type of cyberbullying is often used by abusers to do the bullying using data mining techniques. The results show that the content containing bullying is 86.97% and the types of cyberbullying most Indonesian used are psychological related words that is 61.63%.

According to the study conducted by [12], the authors used SVM to implement different experimental setups with various feature groups and determine the importance of each feature group

to the classification. Lexical features were used to gain insight into the difficulty and learnability of the detection as well as the classification of cyberbullying. The study concluded that it was difficult to classify cyberbullying due to the data sparseness and the degree to which the categories are lexicalized. Similarly, [13] proposed a classification model with optimal accuracy for identifying cyberbully conversation in social media platform known as Formspring.me implementing Naive Bayes method and Support Vector Machine (SVM), and later applied n-gram 1 to 5 for the number of class 2, 4, and 11 for each method. The result yields an average accuracy of 92.81% through the Naive Bayes and the SVM with a poly kernel yields an average accuracy of 97.11%. The study concluded that SVM with poly kernel yields higher accuracy than SVM with other kernels. This results is encouraging, however, the study needs to be analyzed with more dataset to further examine the accuracy of the methods.

The relationship between cyber victimization and suicide ideation was investigated by [14] based on the data collected from 1,062 adolescents using the psychosocial maladjustment variables including loneliness, depressive symptomatology, perceived stress, and psychological distress. The analysis of the results shows that cyber victimization is directly and indirectly related to suicide attempt. The study of [15] proposed an automatic mechanism for cyberbullying detection by tapping into Twitter users' psychological features namely personality, emotion and sentiment as determined by the use of Big Five and Dark Triad models. A total number of 5453 tweets were gathered from Twitter, categorized into four classes: bully, aggressor, spammer and normal and then analyzed using three different classifiers including, Naïve Bayes, Random Forest and J48. The result shows the combination of personalities and sentiments categories yielded improved detection rate than when sentiment was used.

3. Research Methodology

This section describes the methodology framework and the dataset used for the detection of bullying words on Facebook platform.

3.1. Text Mining Process

Text mining is a knowledge discovery technique that involves some important processes to uncover hiding knowledge. These steps are shown in figure 1 and describes as follows:

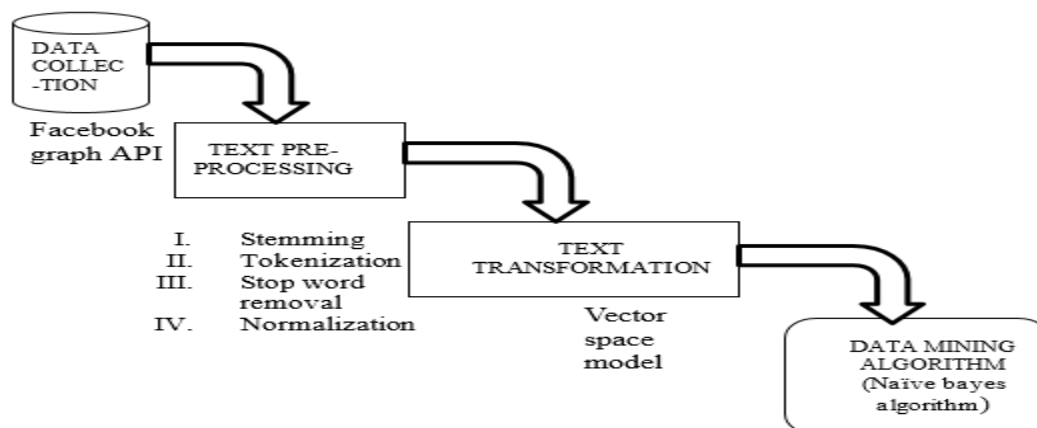


Figure 1. Diagram showing the steps of text mining process

Data Collection: To collect raw data from Facebook, an Application Programming Interface (API) is needed. The raw and unstructured data is gotten from Facebook graph API which requires having an active Facebook account and an access token that allows you to gain access to the Facebook APIs. Facebook Graph API provides a way of interacting with the social graph which is a large graph data structure representing social interactions and consisting of nodes and connections between nodes. A software called FacePager was used to extract the data from the Facebook Graph API. The text data based on public posts and comments was collected using the Facebook Identification numbers of pages. Over 2,000,000 rows of data from sixty Facebook groups were collected and used for a more efficient prediction. The targeted groups for the analysis include: Political Groups; Relationship groups; Football groups; Teens group; Religious groups; and Entertainment groups.

All these categories were chosen due to the form of discussion held and the posts and comments present. This form of data is more effective in getting the activities of cybercriminals because it shows the distress of the user in their post or the messages gotten from these criminals. The extracted data were then stored in a database and exported to a CSV form.

Text Pre-processing: Text pre-processing is one of the most vital steps in text analysis. It is the process of cleaning and preparing the text for classification. The data collected as a CSV format is pre-processed using python 3 on Jupyter notebook. The hypothesis of having the data properly pre-processed is to reduce the noise in the text that should help improve the performance of the classifier and speed up the classification process, thus aiding in real time sentiment analysis. Data gathered from the Facebook are usually incomplete, noisy and inconsistent. Usually, incomplete data lacks attribute values and comprises of errors and outliers [16]. The data pre-processing was performed using software called rapid miner. Text pre-processing mainly involves:

- a. **Tokenization:** This is the process of segmenting the texts into words by removing blank spaces and punctuation marks that are not useful. Tokenization probe the sentences and makes a list of tokens which are used as input for further algorithm [16]. The aim of tokenization is to discover words in a sentence [17]
- a. **Stemming:** Stemming is the process of merging the variant forms of a word into a common representation [17]. It removes all the affixes (suffixes, prefixes, infixes and circumfixes). The hypothesis behind stemming is that words with the same root words can be conflated by using the stems.
- b. **Stop words Removal:** Stop-words are terms in a document that appears often as to make them irrelevant to differentiate document with respect to their content [18]. In this process, word which does not play an important role in providing intelligent pattern or information is removed. However, the development of such stop word list is difficult and inconsistent between textual sources. Every text document deals with these words which are not necessary for text mining applications [17]. In English, there are about 400-500 stop words. Examples of such include 'the', 'to', 'of' etc.
- a. **Normalization:** This involves converting a list of words to a more uniform sequence which involves eliminating punctuation, converting the entire text into lower case. It can be seen as the process of transforming text into a single normal form that it might not have had before. It is used in the cleaning of user-generated text for analysis and prediction [19]. This is not always necessary but due to the information we are trying to get, it is quite important.
- a. **Feature Selection:** The document term matrix contains the set of documents as rows and set of documents as columns. These terms can also be called features because they uniquely identify the document. In this step, features that are considered irrelevant for mining

purposes are removed. It gives more advantages to small data set size, fewer computations, and minimum search space required. The main assumption when using a feature selection technique is the data contains irrelevant and redundant features, both either providing no extra or no relevant information in the document.

Text Transformation: This is the stage where data is modelled and technique for data mining activity is selected. Text transformation allows document to be converted into a bag of words or vector space document model notation to be used in the next phase [20]. Typically, vector space model gives a numeric value corresponding to each term appearing in the document. Vector space model (VSM) was proposed by Salton, it uses vectorization method to convert data into tabular form to enable efficient analysis of large collection of documents. Also, VSM allows vector space documents to be represented by a variable with numerical value, which indicates the weight of the word in the document. In vector space model a set of documents is converted to vectors. The document i can be represented with vector d_i , $d_i = (w_{1,i}, w_{2,i}, \dots, w_{N,i})$, Where $w_{j,i}$ represents the weight word for j that occurs in document i . N is the number of words used for vectorization. In the Term Frequency-Inverse document frequency (tf-idf) scheme, $w_{j,i} = tf_{j,i} \times idf_j$ where $tf_{j,i}$ is the frequency of word j in document i . as represented by [21] and presented in equation 1 thus:

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{i,k}} \quad (1)$$

Where idf_j , is the inverse frequency of word j across all documents and represented thus:

$$idf_j = \log_2 \frac{|D|}{|\{\text{document} \in D | j \in \text{document}\}|}$$

The equation 1 states that the logarithm of the total number of documents divided by the number of documents that contain word j is equal to the inverse frequency of word j across all documents [22]. Therefore, this study used Python programming language to calculate the tf-idf.

Text Mining Algorithm: To analyze and identify Cyber bullying related words from the training data, text mining algorithm is required for the analysis. This study adopts the use of supervised learning algorithm to learn from the sample dataset based on the already identified set of cyberbullying related words and attributes.

Supervised Algorithm: Supervised learning algorithm divides data inputs into training and testing dataset. Supervised learning algorithm uses training dataset (that contains both input and output variables) for learning hiding patterns from the data. In other hands, testing dataset contains the input variables alone. The testing data is used to evaluate the performance of the learning algorithm in predicting every instance of the input data as either normal or cyber bullying word. The workflow for a typical supervised learning is represented in the Figure 2 as follows:

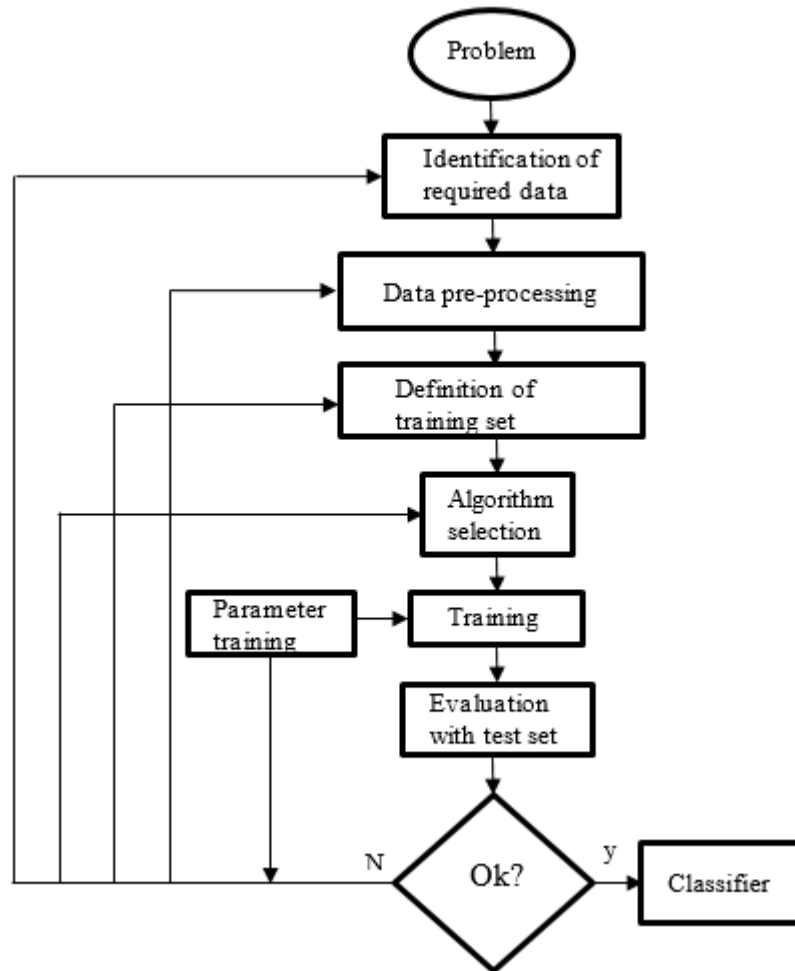


Figure 2. Workflow of Supervised learning algorithm source: [23]

Naïve Bayes: Naïve Bayes is one of the simplest and widely used classification algorithms. It models the distribution of documents in each class using a probabilistic model and making an assumption that the distribution of different terms is independent from each other [24]. These assumptions make the computation of Bayesian classification approach more efficient at the same time limits its applicability. An advantage of the Naïve Bayes classifier is that it requires a small amount of training data to estimate the parameters necessary for classification. Naïve Bayes arrives at the correct classification as long as the correct category is more likely to be true than others. Its main disadvantage is that it has relatively low classification performance compared to other discriminative algorithms such as support vector machine. However, its simplicity and satisfactory performance makes it more attractive in various tasks [25]. The Naïve Bayes theorem states that given two random variable C and D:

$$P\left(\frac{C_i}{D}\right) = \frac{P(C_i)P\left(\frac{D}{C_i}\right)}{P(D)} \quad (2)$$

$$P\left(\frac{D}{C_i}\right) = \prod_{j=1}^n P(d_j|c_i) \quad (3)$$

$$\text{Where } P(C_i) = P(C = c_i) = \frac{N_i}{N}, \text{ and } P(d_j|c_i) = \frac{1+N_{ji}}{M + \sum_{k=1}^M N_{ki}}$$

D represents the class variable and C represents the instance features, C_i represents the value of feature i and D_j represents the value of the variable j . [22]. The Naïve Bayes theorem requires training data or features under which the words would be categorized. The training data that has been chosen are based on the type of popular cyberbullying that occurs on Facebook. In this study, the bullying words are classified into Sexual words, Insult words, Swear words and Animal. The labels include sexual, swear word, unpleasant, offensive and demoralizing. The algorithm for Naïve Bayes algorithm includes:

- INPUT: training set T, hold-out set H, initial number of components k_0 , and convergence thresholds δEM and δAdd .
- Initialize M with one component. $k \leftarrow k_0$
- Repeat
- Add k new mixture components to M, initialized using k random examples from T.
- Remove the k initialization examples from T.
- Repeat
- E-step: Fractionally assign examples in T to mixture components, using M.
- M-step: Compute maximum likelihood parameters for M, using the filled-in data.
- If $\log P(H|M)$ is best so far, save M in M_{best}
- Every 5 cycles, prune low-weight components of M.
- Until $\log P(H|M)$ fails to improve by ratio δEM .
- $M \leftarrow M_{best}$
- Prune low weight components of M.
- $k \leftarrow 2k$
- Until $\log P(H|M)$ fails to improve by ratio δAdd .
- Execute E-step and M-step twice more on M_{best} , using examples from both H and T.
- Return M_{best}

4. Experimental Analysis and Results

The experimental data was collected from the pages of each Facebook group. Each page on Facebook has a unique Identification number collected from a website called 'Find my Facebook Id'. The url of the page was imported to get the Id. On the FacePager, add nodes were used to input the Id for data collection. The parameters to be collected, maximum pages and access token was specified before the data was collected. The access token is unique for each Facebook account and is gotten from Facebook. After the data has been fetched, it was exported to an excel file and saved on the computer as shown in figure 3:

A total number of data with 2,000,000 rows of text data was collected from sixty groups and pre-processed using python3 on jupyter notebook. All these analyses were carried out on windows 8(64-bit) operating system with core i3 processor and 4 GB RAM. Due to the RAM space the data was reduced to 400 rows of text data as a system of 16GB RAM is needed to process the initial text data. For the learning process, the data was divided into two sets, 70% of the dataset for training and 30% of the dataset for testing. The parameters used for classifying the data includes sexual, unpleasant, demoralizing, and offensive. The classification was done on jupyter notebook and the various libraries used include nltk, sklearn, textblob, pandas, numpy etc. As depicted in Figure 3, the occurrence of bullying words =150, insults = 94, Sexual bullying = 32, Swear words = 22, Threat = 1, and animal bully = 1.

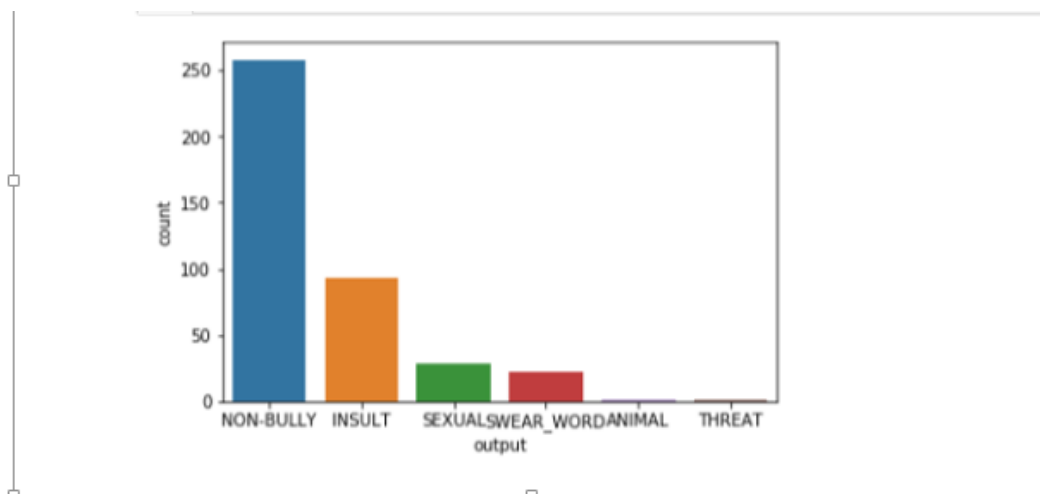


Figure 3. Bar chart showing the number of bullying related words.

During the analysis, only the columns (sexual, offensive, demoralizing, unpleasant and output) were used to get the result. These columns were pre-processed and converted to the form understands by computer (0s and 1s). The Sexual, Offensive, Demoralizing, Unpleasant, and Output were manually imputed as the computer requires the parameters to determine if there is an occurrence of cyberbullying in a sentence. Each message passed through the test of if; i.e. are the words sexual, offensive, demoralizing, unpleasant etc.? The values for true were converted to 1 and the values for false instances were converted to 0. Each unique output was assigned a given number (non_bully = 0, insult = 1, sexual = 2, swear_word = 3, threat = 4, animal = 5). All other columns were dropped using the drop function as they were no longer needed. The astype was used to convert each output and parameter to an array format for it to be well understood by the machine.

After the id, parent_id, object_type, name, query_type columns were deleted, the remaining labels which would be used during the prediction and analysis include sexual, demoralizing, unpleasant, offensive, swear word and the output. The data was then converted to lower case, punctuations were removed before the conversion process. Pandas library was imported and used for the cleaning process.

The pre-processed data was fed into Naïve Bayes algorithm to identify or classify the cyber-bullying related words from the text data. After the training and testing phase, the performance of the Naïve Bayes based cyber-bullying detection model was measured using some of the state-of-the-art performance measures in the text mining field.

4.1. Performance measurement

To measure the performance of the proposed classifier on bullying based dataset on Facebook group pages, several performance metrics were used. The performance metrics used are:

Confusion Matrix: This metric is used in measuring the performance of the algorithm. The predicted classes are represented in the columns of the matrix and the actual classes are in the row. Figure 4 shows the components of the confusion matrix for the performance of Naïve Bayes on the cyber-bullying dataset. The matrix indicates that for every row there are 4 cases, True positive, True Negative, False Positive, and False Negatives. True positive is the case for which the classifier

predicted bully and the words were bully. True Negative is the case for which the classifier predicted not bully and the words were not bullied. False Positive is the case for which the classifier predicts 'bully' but the words are Nonbully. False Negative is the cases for which the classifier predicted 'not bully' but the words were bully.

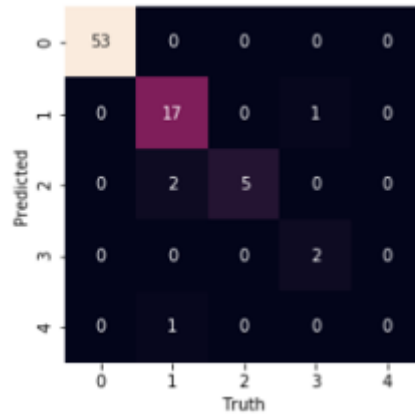


Figure 4. Graph of the confusion matrix

As depicted in figure 5: Row 1: Out of the 53 actual instances of 'Non-bully' the classifier predicted 53 correctly. Row 2: Out of the 18 actual instances of 'Insult' the classifier predicted 17 correctly and 1 was predicted wrongly. Row 3: Out of the 7 actual instances of 'Sexual' the classifier predicted 2 correctly and 5 incorrectly. Row 4: Out of the 2 actual instances of 'Swear word' the classifier predicted none correctly. Row 5: Out of the 1 actual instances of 'animal' the classifier predicted 1 correctly.

This indicates that out of the 88 instances, the classifier predicted 81 of them correctly. The threat was not included in the dataset classified by the classifier.

Score Accuracy: This is one of the popular ways of measuring the performance of any text mining algorithm. The score accuracy answers the question of how often the classifier predicts correctly. In another word, it measures how many times the algorithm correctly classified the words. Sklearn was used to obtain the accuracy using accuracy score measure. The output of the accuracy is 95%. This means that for every 100 words classified by the algorithm, 95 instances were correctly classified as non-bully, sexual, insult, swear-word, threat, and animal.

Classification Report: This shows the precision, recall, and f1-score. The precision answers the questions "When the classifier predicts, how often is the prediction correct?" That is, how many times is the prediction made by the classifier correct? The recall answers the question "How many times is the prediction made by the classifier correct when the result is true?" The f1-score is the harmonic mean of precision and recall.

The performance of the Naïve Bayes algorithm on the text data in terms of predicting the occurrence of normal and cyber-bullying related data is presented in table 1.

Table 1. Classification Report Showing the Precision, Recall, and F1- Score of the Naïve Bayes Algorithm on Cyber-Bullying Dataset

Text Category	Precision	Recall	F1 - score
0	1.00	1.00	1.00
1	0.94	0.85	0.89
2	0.71	1.00	0.83
3	1.00	0.67	0.80
4	0.00	0.00	0.00
Micro avg	0.95	0.95	0.95
Macro avg	0.73	0.70	0.71
Weighted avg	0.97	0.95	0.96

In row 1 of table 1, the Naïve Bayes recorded a precision, recall and F1 – score of 1.00, 1.00, and 1.00 respectively. This shows the naïve bayes algorithm can differentiate between normal and cyberbullying related words accurately. Row 2 shows that Naïve Bayes measured 0.94, 0.85, and 0.89 as precision, recall and F1 – score respectively on insult related data, and recorded 0.71, 1.00, and 0.89 as precision, recall, and F1 – score for sexually related data. On the swear_word, the algorithm recorded 1.00, 0.67, and 0.80 as precision, recall and F1 – score respectively, while recorded no measurement for the threat-related words. These results indicate that Naïve Bayes is effective and efficient in predicting cyberbullying related words using text mining techniques.

5. Conclusion and Future Work

Cyberbullying is a serious online crime that has led to numerous suicides among youths in the world. Facebook, as one of the social media platforms where cyberbullying activities take place, has a privacy policy that allows a user to choose the number of people that view their posts and comments but many users are ignorant of the privacy policy. Due to the failure of the existing Cyberbullying detection techniques to effectively analyse users posts and comments on social media, many people have fall victim of this life threatening online crime. Text mining is a form of data mining techniques that is gaining a tremendous adoption among security experts and researchers. Due to its effectiveness in identifying every instance of bully related words in a text, this study adopts Naïve Bayes based text mining technique for predicting bully words in Facebook posts and comments. The performance of the Naïve Bayes classifier has been measured using score accuracy, confusion matrix and classification reports including precision, recall and F1- score. The analysis of the Naïve Bayes based text mining technique shows the accuracy of the classifier to be 0.95%. The results also show that Naïve Bayes recorded a precision, recall and F1 – score of 1.00, 1.00, and 1.00 respectively for normal words. More so, it recorded 0.94, 0.85, and 0.89 as precision, recall and F1 – score respectively on insult related data, 0.71, 1.00, and 0.89 as precision, recall, and F1 – score for sexually related data, and recorded 1.00, 0.67, and 0.80 as precision, recall and F1 – score respectively on swear-related words.

This experimental analysis result shows that Naïve Bayes is effective in classifying every instance of Facebook posts into a bully or non-bully word and was able to identify the category of a bully word that is being sent. For future work, we plan to further develop a technique to analyse and detect bullying related pictures and videos. Besides, a novel algorithm could also be used to improve the detection accuracy of the cyber-bullying model. Finally, a big data analytic based technique could also be used to handle the huge amount of data that characterized modern social media.

References

- [1] Salloum, S. A., Mostafa Al-Emran, Azza Abdel Monem, and Khaled Shaalan. (2017):"A survey of text mining in social media: Facebook and twitter perspectives." *Adv. Sci. Technol. Eng. Syst. J* 2, no. 1 127-133.
- [2] Sajithra, K., & Patil, R. (2013). Social media–history and components. *Journal of Business and Management*, 7(1), 69-74.
- [3] Clement, J. (2020). *Number of monthly active Facebook users worldwide as of 3rd quarter 2020*. Retrieved from <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>
- [4] Kandlapalli, N., Shinde, S., Shriramoji, P., Uke, P., and Chaudhary, S. (2017). Defending Mechanism for cyber bullying, *International journal of sscientific research in computer science, Engineering and Information Technology*, 2(2),716-719
- [5] Kasali, F., Kuyoro, A. (2015). Text mining approach in curtailing cyber-crime in Nigeria, *International Journal of computer science trends and technology*, 3(6), 1-8..
- [6] Rosa, H., Pereira, N., Ribeiro, R., Ferreira, P. C., Carvalho, J. P., Oliveira, S., Coheur, L., Paulino, P., VeigaSimão, A.M., & Trancoso, I. (2019). Automatic cyberbullying detection: A systematic review. *Computers in Human Behavior*, 93, 333-345.
- [7] Chun, J., Lee, J., Kim, J., & Lee, S. (2020). An international systematic review of cyberbullying measurements. *Computers in Human Behavior*, 106485.
- [8] Ganesan, M., & Mayilvahanan, P., (2017) Cybercrime Analysis in social media using data mining technique, *International journal of pure and applied mathematics*, 116, 413-424
- [9] Funmilayo, K., Afolashade, K. & Oludele, A. (2015) Text mining approach in curtailing cyber- crimes in Nigeria, *International journal of computer science trends and technology*, 3(6), 1-8
- [10] Sugandhi, R., Pande, A., Agrawal, A. & Bhagat, H. (2016). Automatic monitoring and prevention of Cyberbullying. *International Journal of Computer Application*, 144(8), 17-19.
- [11] Hariani, Riadi, I. (2017). Detection of cyberbullying on social media using data mining techniques. *International Journal of Computer Science and Information Security (IJCSIS)*, 15(3).
- [12] VanHee, C., Lefever, E., Verhoeven, B., Mennes, J., Desmet, B., Pauw., G, Daelemans, W., &Hoste, V.,(2015). Detection and Fine-Grained Classification of Cyberbullying Events. *Proceedings of Recent Advances in natural language Processing*, 672 – 680.
- [13] Noviantho, Isa, S. M., & Ashianti, L. (2017). Cyberbullying classification using text mining. In *2017 1st International Conference on Informatics and Computational Sciences (ICICoS)* (pp. 241-246). IEEE.
- [14] Iranzo, B., Buelga, S., Cava, M. J., & Ortega-Barón, J. (2019). Cyberbullying, psychosocial adjustment, and suicidal ideation in adolescence. *Psychosocial Intervention*, 28(2), 75-81.
- [15] Balakrishnan, V., Khan, S., & Arabnia, H. R. (2020). Improving cyberbullying detection using Twitter users' psychological features and machine learning. *Computers & Security*, 90, 101710.
- [16] Nayak, A. S., Kanive, A. P., Chandavekar, N., & Balasubramani, R. (2016). Survey on pre-processing techniques for text Mining. *International Journal Of Engineering And Computer Science*, ISSN, 2319-7242.
- [17] Kanan, T., Aldaaja, A., & Hawashin, B. (2020). Cyber-Bullying and Cyber-Harassment Detection Using Supervised Machine Learning Techniques in Arabic Social Media Contents. *Journal of Internet Technology*, 21(5), 1409-1421.
- [18] Ferilli, S., Esposito, F., & Grieco, D. (2014). Automatic learning of linguistic resources for stopword removal and stemming from text. *Procedia Computer Science*, 38, 116-123.
- [19] Jianqiang, Z. and Xiaolin, G., (2017). Comparison Research on Text Preprocessing Methods on Twitter Sentiment Analysis. *IEEE Access*, 5, pp.2870-2879
- [20] Maheswari, M. U., &Sathiaseelan, J.G.R.(2017). Text mining: Survey on Techniques and Applications. *International Journal of Science and Research*, 6(6), 1660-1664.pro
- [21] Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.

- [22] Zafarani, R., Abbasi, M. A., & Liu, H. (2014). *Social media mining: an introduction*. Cambridge University Press.
- [23] Dey, A. (2016). Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*, 7(3), 1174-1179.
- [24] Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint arXiv:1707.02919*.
- [25] Khan, K., Baharudin, B., & Khan, A. (2014). A review of unsupervised approaches of opinion target extraction from unstructured reviews. *Research Journal of Applied Sciences, Engineering and Technology*, 7(12), 2400-2410.