A Technique For Perceiving Abusive Bangla Comments

Article · July 2019

DOI: 10.5281/zenodo.3544583

2 authors:

Md Gulzar Hussain
Green University of Bangladesh
19 PUBLICATIONS **11** CITATIONS

Tamim Al Mahmud
Aston University
17 PUBLICATIONS **16** CITATIONS

GUBJSE: ISSN: 2409-0476

A TECHNIQUE FOR PERCEIVING ABUSIVE BANGLA COMMENTS

Md Gulzar Hussain and Tamim Al Mahmud

***Abstract*—Most of the research on abusive comments or text detection is conducted in English, some of which are intended to detect humiliating or insulting text. But a few works are found in the Bangla language. Detecting abusive text for Bangla language will be helpful to prevent cyber crimes such as online blackmailing, harassment and cyber bullying which are nowadays becoming the main concern in Bangladesh. Our goal is to detect abusive Bangla comments that are gathered from different social sites where people share their views, feelings, opinions, etc. in this paper. In order to classify a bangla comments is abusive or not, we proposed a root level algorithm and also proposed uni-gram string features to achieve a better result. We have collected several comments from renowned social media Facebook for our work.**

***Index Terms*—Abusive, Bangla, Comment, Natural Language Processing (NLP), Unigram.**

I. Introduction

NATURAL languages are the languages of human, but computer programming languages e.g. C and C++, are not. For example, Bangla, French, English, and Chinese are natural languages. In computer science, making a computer to understand natural languages may be the most challenging issue. Natural language processing (NLP) is a pathway for computers to understand, explore, and derive meaning from human language in a smart and useful path. By utilizing NLP, developers can organize and structure knowledge to perform tasks such as automatic summarizing, translation, relationship extraction, named entity recognition, sentiment analysis, topic segmentation, and speech recognition. A large portion of the examination being done on natural language processing revolves around search, particularly enterprise search. This includes enabling users to query data sets in the form of a question that they might pose to another person. The machine translates the important elements of the human dialect sentence, such as those that might compare to particular highlights in a data set, and returns an answer. NLP can be used to interpret and analyze free text. There is an enormous amount of information, such as medical records of patients, stored in free text files. Before deep learning NLP models, this information was not accessible for computer - assisted analysis and could not be systematically analyzed. However, NLP enables analysts to find relevant information in the files by means of massive troves of free text. Another primary use case for NLP is sentiment analysis. Data scientists can evaluate comments on social media using sentiment analysis to see how their business's brand is performing, for example, or review notes from customer service teams to identify areas where people want the company to perform better. Google and other search engines use NLP deep learning models to base their machine translation technology. This allows algorithms to read text on a web page, interpret its meaning and translate it into different language. In this field, research work is increasing day by day. Natural Language Processing contributes to almost every sector such as Customer Service, Healthcare and Automotive etc. According to the report by Tractica(2017), market profit on NLP software, hardware, and various kinds of services would be around $22.3 billion by 2025. It forecasts that this Artificial Intelligence software service solution will rise from $136 (million) in 2016 to a market value of $5.4 (billion) by 2025 [1].

In 2010 Bangla Language was spoken by 205 million native speakers and is the seventh most frequently spoken native language in the world [2]. This total number covers 3.05 percent of the world's total population. And in 2017 there will be around 250 - 300 million total speakers worldwide [3]. There are 81.7 million Internet users in Bangladesh, 30.0 million of whom are active social media users and 28.0 million of whom use mobile phones in January 2018 to access social media [4]. Facebook is the world's most popular social networking sites with 2.23 billion active monthly users, YouTube is second with 1.9 billion active monthly users and Instagram is third with 1 billion active monthly users in August 2018 [5].

A lot of research has been done in the field of abusive text detection in English Language using social networks but limited amount of research has been done in Bangla Language. Though we found some works on sentiment analysis on Bangla language but recently very few research has been done to detect abusive Bangla text using social network sites. So, in this field, there are lots of research scope for us. In this article, we are proposing a new technique to detect

This paper was received on 22 April 2019 and accepted on 15 July 2019. This work was supported financially by GUB.

Md Gulzar Hussain is with the Computer Science & Engineering, Green University of Bangladesh, Dhaka, Bangladesh. E-mail: gulzar@cse.green.edu.bd

Tamim Al Mahmud is with the Computer Science & Engineering, Green University of Bangladesh, Dhaka, Bangladesh. E-mail: tamim@cse.green.edu.bd

abusive Bangla text. We proposed two algorithms using machine learning idea. One of them is to train our system using training dataset and another one is to classify the test dataset. We are just classifying a Bangla text as if it is abusive or not. Our proposed system giving a satisfied accuracy of 71.7%.

II. Literature Review

We briefly discussed the methodologies of different research carried out on the English and Bangla language. There are a large number of approaches that has been developed to date for classifying sentiments or polarities in English texts. These methods can be classified in two categories- (1) machine learning or statistical-based approach and (2) unsupervised lexicon based approach. But detecting abusive text in social sites is one of the challenging work due to the changing nature and the variation in the language used. Researchers tried to develop many approaches to detect abusive or offensive text to get a better result. When it comes to work with Bangla language it becomes more difficult to detect abusive text.

A. Work In English Language

In paper [6], they developed a machine learning based method for detecting hate speech on online user comments and categorized the sentences into Hate Speech, Derogatory and Profanity categorizes. They used Vowpal Wabbits regression model to measure different aspects of the user comment and used N-grams, Linguistic, Syntactic features. Using multi-class classifier, [7] categorized tweets into hate speech, offensive and neither of these two and differ hate speech from offensive language. The authors of [8] proposed a Lexical Syntactic Feature (LSF) architecture for detecting offensive content and identify potential offensive users in social media. In [9], the authors proposed a statistical topic modeling to detect profanity-related offensive content in Twitter.

[10] is one of first papers to apply supervised machine learning methods to sentiment classification. The authors perform the classification on movie reviews and show that MaxEnt and SVM outperform Nave Bayes (NB) classifier. One of the first papers on the automatic classification of sentiments in Twitter messages, using machine learning techniques, is by [11].Through distant supervision, the authors use a training corpus of Twitter messages with positive and negative emoticons and train this corpus on three different machine learning techniques- SVM, Nave Bayes, and MaxEnt, with features such as N-grams (unigrams and bigrams) and Part of Speech (POS) tags. They obtain a good accuracy of above 80%. [12] follow the same procedures as [11] to develop the training corpus of Twitter messages, but they introduce a third class of objective tweets in their corpus and form a dataset of 3 classes- positive sentiments, negative sentiments, and a set of objective texts (no sentiments). They use multinomial NB, SVM, and Conditional Random Field (CRF) as classifiers with N-grams and POS-tags as features. The authors of [13] use 50 hashtags and 15 emoticons to train a supervised sentiment classifier using the K-Nearest Neighbors (KNN) algorithm as sentiment labels. In [14], the authors integrate a two-steps sentiment detection framework by first distinguishing non-subjective tweets from subjective tweets and then further classify the subjective tweets into negative and positive polarities. The authors find that using meta-features (POS tags) and tweet-syntax features (emoticons, punctuations, links, retweets, hashtags, and uppercases) to train the SVM classifiers enhances the sentiment classification accuracy by 2.2% compared to SVMs trained from unigrams only. Although supervised machine learning methods have been widely employed and proven effective in sentiment classification, they normally depend on a large amount of labeled data, which is both time consuming and labor intensive work.

Unsupervised lexicon-based methods rely on manually or semi-automatically constructed lexical resources, such as lexicons, to identify the polarity of texts in general. Lexicon is the collection of strong sentiment-bearing words or phrases, which are labeled with their prior polarity, or the context-independent polarity most commonly associated with the lexicon entries. There are several lexicons in English which are available online. One of the initial works to apply unsupervised techniques to sentiment classification problem is by [15]. In the paper, the average semantic orientation of sentences containing adjectives or adverbs classifies a document as positive or negative. The semantic orientation of a phrase is calculated as the Pointwise Mutual Information (PMI) with a positive seed word excellent minus the PMI with a negative seed word poor. This approach achieves 84% accuracy in automotive reviews and 66% accuracy in film reviews. In [16], the authors develop a sentiment lexicon manually consisting of negative and positive sentiment bearing words annotated with their POS tags. This sentiment lexicon, along with a set of regulations, is used to first classify the tweets as subjective or objective and then further classify the subjective tweets as positive, negative or neutral. They use a corpus of political tweets collected over the UK pre-election period in 2010. For the task of correctly identifying that a document contains a political sentiment and then correctly identifying its polarity, they get 62% Precision and predict 37% Recall. However, methods based on lexical resources often have the problem of obtaining low recall values because they depend on the presence of the words comprising the lexicon in the message to determine the orientation of opinion. And due to the varied and changing nature of the language used on Twitter, this approach is not suitable for our thesis work. Moreover, as such lexical resources are not available for many other languages spoken in social media, such as Bangla, hence this approach often becomes unsuitable for scarce-resource languages.

B. Work In Bangla Language

The authors of [17] calculated the total positivity, negativity of sentence or document with regard to the total meaning of the sentence. Tf. Idf (term frequency

- inverse document frequency) was used to find a better solution in this process of obtaining information from a document. They wanted to determine some patterns in this experiment so that positive and negative sentences could be categorized. In paper [18], they tried to extract from Bangla Micro-blog's posts the negative or positive opinion or feeling of a full text. They used Support Vector Machine (SVM) and Maximum Entropy (MaxEnt) for classification and developed the training corpus using a semi - supervised bootstrapping approach. By combining the results of word2vec word co - occurrence score with the sentiment polarity score of the words [19], the authors tried to classify Bengali comments with sentiment and found 75.5% accuracy.

Authors of [20] proposed an root leveled algorithm to find out the abusiveness of a Bangla comment. They worked with 300 Bangla comments but didn't found any accuracy using those data. We found a paper which tried to detect abusive Bangla text is [21], used Random Forest(RF), Multinomial Nave Bayes(MNB), Support Vector Machine (SVM) with Radial Basis Function (RBF), Polynomial, Linear, and Sigmoid kernel and compared with uni-gram, bi-gram and tri-gram based Count-Vectorizer and Tf.idf-Vectorizer features to detect Bengali abusive text. They found that with the Tf.idf - Vectorizer trigram functions, the SVM Linear kernel performs best.

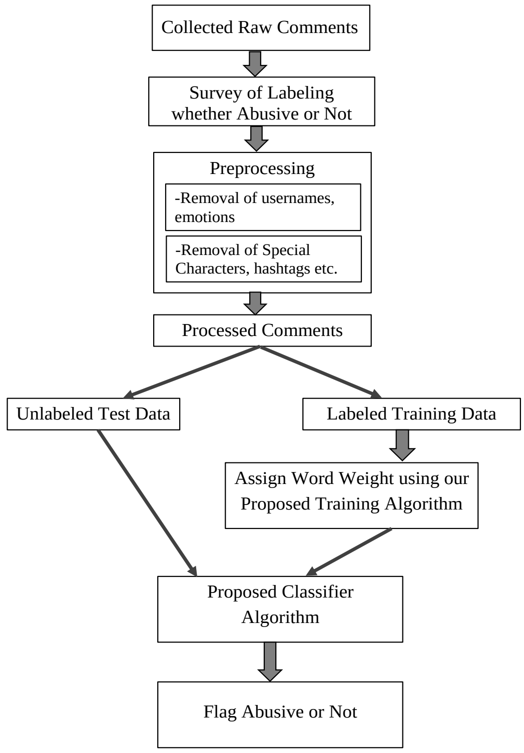


Fig. 1: System Architecture

III. METHODOLOGY

Abusive text classification techniques may be categorized into two types, i.e., binary classification & multi-class classification. In binary, we can just determine if a comment is abusive or not. But in multi-class classification a comment can be categorized as hate speech, anger to someone or some group of people, criticism, insult etc. Binary classification is much easier than multi-class classification for English Language. It becomes more harder in case of Bangla Language. So we decided to work with the binary classification in Bangla Language text. For conducting the experiment we developed a manual algorithm to detect abusive Bangla text. We implemented our proposed algorithms and automate the procedure to get the result. The system architecture is the basic structure and general vision of a system. Our system architecture is shown in Figure 1, which outlines the entire process.

A. Dataset Collection

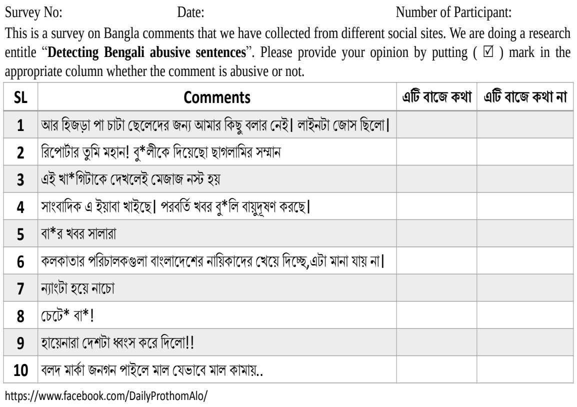
For conducting experiment, we have collected comments on posts from Facebook pages, Prothom-Alo news, and YouTube channels e.g. Prothom Alo [22], Mashrafe Bin Mortaza [23], Shakib Al Hasan [24], SalmoN TheBrownFish [25], Naila Nayem [26] and Prothom Alo News Portal [27]. Only public comments are collected without the commenters information due to protect privacy. In total, we collected 300 comments as we will do the whole experiment manually. We have done the experiment in three different set of 100, 200 and 300 comments. We used 80% of the comments to train the term weighting using proposed Training Algorithm and remaining 20% of the comments will be tested using proposed Classifier Algorithm. However, the comments we collected contained English text as well; instead of filtering out the English text from the comments, we take in them as part of our training and testing sets. Since English words express strong abusiveness, they are likely to contribute to the classification of our data set.

B. Survey

Survey is used to gather the opinions, beliefs and feelings of collected comments in our research. This survey for collecting comments is the part which no other authors had done. They used rule based classifier and methods to label the training data. But we were doing this survey to ensure the fact that the comments given the right polarity of abusiveness. We ran the survey in our varsity campus and students of various age participated in our survey. To label every comment we had to run survey on the comments. We created a survey form as shown in Figure 2 and ran survey in Green University of Bangladesh. For every comment, we were taken opinion from at least 50 persons to take a decision whether the comment is abusive or not. Finally we shown found results for 300 comments in Figure 3 and labeled every comment based on the result.

C. Preprocessing

Preprocessing is an significant task and a crucial step in Text mining, Natural Language Processing (NLP). The information requirement of the user is represented by a query or profile and includes one or more search terms and additional information such as



Survey No: Date: Number of Participant:

This is a survey on Bangla comments that we have collected from different social sites. We are doing a research entitle "**Detecting Bengali abusive sentences**". Please provide your opinion by putting (☑) mark in the appropriate column whether the comment is abusive or not.

SL	Comments	এটি বাজে কথা	এটি বাজে কথা না
1	আর হিজড়া পা চাটা ছেলেদের জন্য আমার কিছু বলার নেই। লাইনটা জোস ছিলো।		
2	রিপোর্টার তুমি মহান! বু*লীকে দিয়েছো ছাগলামির সম্মান		
3	এই খা*গিটাকে দেখলেই মেজাজ নস্ট হয়		
4	সাংবাদিক এ ইয়াবা খাইছে। পরবর্তি খবর বু*লি বায়ুদূষণ করছে।		
5	বা*র খবর সালারা		
6	কলকাতার পরিচালকগুলা বাংলাদেশের নায়িকাদের খেয়ে দিচ্ছে,এটা মানা যায় না।		
7	ন্যাংটা হয়ে নাচো		
8	চেটে* বা*!		
9	হায়েনারা দেশটা ধ্বংস করে দিলো!!		
10	বলদ মার্কা জনগন পাইলে মাল যেভাবে মাল কামায়..		

https://www.facebook.com/DailyProthomAlo/

Fig. 2: A Sample Survey Form

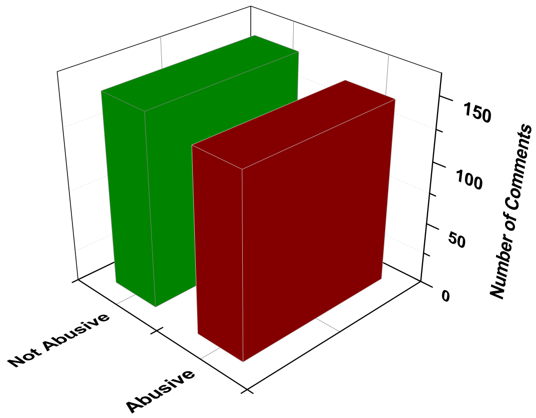


Fig. 3: Survey Result for all 300 Comments

the weight of the words. The decision to retrieve is therefore taken by comparing the query terms with the index terms (important words or phrases) that appear in the document itself. The decision may be binary (reclaim / reject) or may involve an assessment of the degree of relevance that the document needs to query. Unfortunately, the words in documents and queries often have many structural variations. Therefore, the data preprocessing techniques are used on the target data set to reduce the size of the data set, which increases the efficiency of the IR system, before the information is retrieved from the documents. The aim of this study is to analyze the problems of pre-processing methods such as tokenization, word deletion and stemming for the text documents.

1) Normalization: Text needs to be standardized before further processing. Normalization generally refers to a series of related tasks which means that all text is placed on a level playing field: all text is converted to the same case (upper or lower), punctuation is removed, numbers are converted to their word equivalents, etc. Normalization places all words on the same footing and allows uniform processing. The raw comments contain special characters e.g. @, #, - etc., punctuation, and emotions. At the time of prepossessing, we removed these special characters, white spaces, numbers, punctuation and Unicode emotions manually. Conjunctions are also removed as they are unnecessary to abusive text detection.

2) Noise Removal: Bear in mind once again that we are not dealing with a linear process whose steps must be applied exclusively in a given order. Therefore, noise removal can take place before or after the previously outlined sections or somewhere between them. In this step we removed user names, hash tags, unwanted signs etc.

D. Proposed Algorithm

The proposed algorithm have two parts. One is for training and another is for classify the comments. The training algorithm is applied on the labeled training data for term weighting and the classifier algorithm is applied on the unlabeled data to classify them if it is abusive or not.

1) Training Algorithm: Our Training Algorithm is based on bag-of-words model approach [28]. In this model, A text is represented as an un-ordered collection of words, disregarding grammar and even word order. It is counting the frequency of a word and creating a term weighting table. To train Algorithm 1 is proposed-

---

Algorithm 1: Training Algorithm

Step 1: Start
Step 2: For each comments of the data-set steps 3 and 4 is taken
Step 3: Initial,

$$nWc = NumberOfWordsInTheComment$$

Step 4:
for *Each words in a comment* **do**
  if *Comment is labeled as abusive* **then**
    if *Word is not in the list of weighted-words* **then**
      $Weight_{Abusive} = \frac{1}{nWc}$
    else
      $Weight_{Abusive} = (old)Weight_{Abusive} + \frac{1}{nWc}$
    end
  else
    if *Word is not in the list of weighted-words* **then**
      $Weight_{NotAbusive} = \frac{1}{nWc}$
    else
      $Weight_{NotAbusive} = (old)Weight_{NotAbusive} + \frac{1}{nWc}$
    end
  end
end

---

2) Classifying Algorithm: Our Classifying Algorithm is calculating the summation of the abusiveness or non-abusiveness of a text using term weighting table. To classify the test comments Algorithm 2 is proposed-

E. Feature Extraction

We can conduct the experiment with three types of string features to find out, with what kind of features our proposed algorithm performs better. They are uni-gram, bi-gram, and tri-gram [29]. In uni-gram characteristics, the relationship between words in a

Algorithm 2: Classifying Algorithm

Step 1: Start
Step 2: For each comments of the data-set steps 3 and 4 is taken
Step 3:
for *Each words in the comment* **do**
  if *Word is not in the Term-Weight List* **then**
    $TotalWeight_{Abusive} = TotalWeight_{Abusive} + 0$
    $TotalWeight_{NotAbusive} = TotalWeight_{NotAbusive} + 0$
  else
    $TotalWeight_{Abusive} = Weight_{Abusive} + TotalWeight_{abusive}$
    $TotalWeight_{NotAbusive} = Weight_{NotAbusive} + TotalWeight_{NotAbusive}$
  end
end
Step 4:
if

$$TotalWeight_{Abusive} > TotalWeight_{NotAbusive}$$

then
  Set the label of that comment as abusive that is 1
else
  Set the label of that comment as not abusive that is 0
end

sentence is not considered. But using this feature, it can be found which words are more abusive. Bi-gram features consider the relationship between two consecutive words in a sentence. In the tri-gram, the relationship is considered in a sentence between three consecutive words.

F. Illustration using test comments

A step-by-step illustration of our proposed algorithm is given below using some sample comment. Consider unlabeled comments of figure 4,

Serial	Comments
01	কু*র বাচ্চা কোথাকার।
02	বাবুকে আরও খাবার দাও।
03	কু*র বাচ্চা দেরেকে খাবার দাও।

Fig. 4: Unlabeled sample comments

Here think the first and second comments as training data. After running the survey suppose we found the following labeled comments in figure 5-

Now the part of preprocessing the comments and use of training algorithm comes into play. So the following table found given in figure 6-

For the third comment, if we run the test algorithm then the following table of figure 7 calculated-

Serial	Comments	Label
01	কু*র বাচ্চা কোথাকার।	1
02	বাবুকে আরও খাবার দাও।	0
03	কু*র বাচ্চা দেরেকে খাবার দাও।	

Fig. 5: Labeled sample comments after survey

Comment	Weight $_{abusive}$	Weight $_{notabusive}$
কু*র	0.33	0.0
বাচ্চা	0.33	0.0
কোথাকার	0.33	0.0
বাবুকে	0.0	0.2
আরও	0.0	0.2
খাবার	0.0	0.2
দাও	0.0	0.2

Fig. 6: Term Weighting Table after preprocessing and running the training algorithm

Comment	কু*র	বাচ্চা	দেরেকে	খাবার	দাও	Sum
TotalWeight $_{abusive}$	0.33	0.33	0	0	0	0.66
TotalWeight $_{notabusive}$	0.0	0	0	0.2	0.2	0.4

Fig. 7: After applying the classifying algorithm

Figure 7 shows that for the third comment the sum of

$$TotalWeight_{abusive} > TotalWeight_{notabusive}$$

So that the third comment is abusive.

IV. Result Analysis And Evaluation

We have collected 300 random comments for our research work. Among them, We divided our data set into three sets with 100 comments, 200 comments, and 300 comments. We provide the experimental results of Correct Abusive, Wrong Abusive, Correct not-Abusive, Wrong not-Abusive for the binary classification task using our proposed algorithm with uni-gram feature for 20% test data in TABLE I and Figure 8. From TABLE I and Figure 8 we can see that the number of accurateness is increasing with the increase of number of comments. //

TABLE I: Correct and wrong result for abusive and not abusive class for 20% test data of three sets of comments

Number of Comments	Correct Abusive	Wrong Abusive	Correct not Abusive	Wrong not Abusive
100 comments	10	0	4	6
200 comments	16	4	10	10
300 comments	25	5	18	12

A. Evaluation Metrics

In order to examine the performance of the proposed algorithm, we first use the standard precision, recall and F-measure to measure the abusive or the not abusive for the classifier algorithm using uni-gram features. We then use the accuracy metric to find the overall performance of the proposed classifier algorithm. Abusiveness analysis task can be interpreted as a classification task where each classification label represents a Abusiveness. Therefore, the four metrics for each label (abusive and not abusive) are defined and calculated in the same way as in the general classification task.

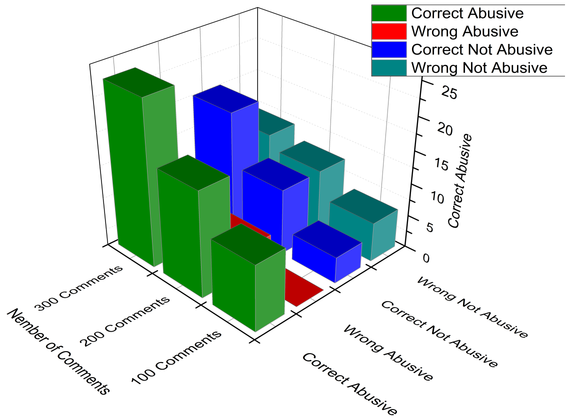


Fig. 8: Number of correct abusive, correct not abusive, wrong abusive and wrong not abusive class for three sets of comments

In a classification task, precision, recall, F-measure and accuracy are explained by four terms - true positive, true negative, false positive and false negative.

- **True Positive** (TP)**:** is defined as the number of comments from the test set that the classifier correctly labels as belonging to a specific class or label.
- **True Negative** (TN)**:** is defined as the number of comments from the test set that the classifier correctly labels as not belonging to a specific class or label.
- **False Positive** (FP)**:** is defined as the number of comments from the test set that the classifier incorrectly labels as belonging to a specific class or label.
- **False Negative** (FN)**:** is defined as the number of comments from the test set which the classifier does not label but should have belonged to a specific class or label.

Using these four terms, we now define the evaluation metrics as follows:

- **Precision** is the number of comments in the test set that is correctly labeled by the classifier from the total comments in the test set that are classified by the classifier for a specific class. That is,

$$Precision(P) = \frac{TP}{TP + FP}$$

- **Recall** is the number of comments in the test set that is correctly labeled by the classifier from the total comments in the test set that are actually labeled for a specific class. That is,

$$Recall(R) = \frac{TP}{TP + FN}$$

- **F-measure** is the weighted harmonic mean of precision and recall for a specific class. That is,

$$F - measure = \frac{2 * P * R}{P + R}$$

- **Accuracy** is the percentage of comments in the test set that the classifier correctly labels. That is,

$$Acuracy(A) = \frac{TP + TN}{TP + TN + FP + FN} * 100\%$$

To calculate precision, recall, F-Measure and accuracy, we manually calculated the values of TP, TN, FP, FN.

B. Result

We divide our data set into three sets with 100 comments, 200 comments, and 300 comments. We provide the experimental results of precision, recall, F-measure, and accuracy for the binary classification task using our proposed algorithm with unigram feature in Table II, Table III and Table IV. In table II and table III, we can govern that, we found the best F-measure score of 0.75, for both the abusive as well as the not abusive label for 300 comments. Table IV shows the average values of table II and table III

TABLE II: Experimental outcome of Precision, Recall and F-measure for abusive comments

Serial	Number of Comments	Precision	Recall	F-measure
1	100	0.625	1.0	0.769
2	200	0.62	0.8	0.7
3	300	0.68	0.83	0.75

TABLE III: Experimental outcome of Precision, Recall and F-measure for non abusive comments

Serial	Number of Comments	Precision	Recall	F-measure
1	100	1.0	0.286	0.445
2	200	0.71	0.5	0.59
3	300	0.78	0.6	0.75

TABLE IV: Average experimental outcome of Precision, Recall and F-measure from Table II and Table III

Serial	Number of Comments	Precision	Recall	F-measure
1	100	0.81	0.64	0.61
2	200	0.67	0.65	0.65
3	300	0.73	0.72	0.75

The accuracy of our proposed algorithm for the three sets of comments and their average are given in Table V. We can see that, we achieved the best accuracy of 71.7% for 300 comments.

TABLE V: Accuracy for the experiment result

Serial	Number of Comments	Accuracy	Average Accuracy
1	100	70%	68.9
2	200	65%	
3	300	71.7%	

From our experimental results in Table V, we can say that using more comments is in fact very effective and offers promising performance for both the proposed classifier algorithm.

V. CONCLUSION

In this paper, we discuss how we collect the training data and test data in a manual way and perform abusive text analysis on Bangla comment data. We tried to do that using our proposed classifying algorithm. Though this type of root level algorithm is not appropriate for nowadays, we achieve a satisfying maximum accuracy of 71.7% for 300 comments and average accuracy of 68.9% for overall test data set. We hope our future plan will give a better result than already developed ideas. There are still many opportunities to improve our experimental methodology. Natural language processing using various machine learning algorithms and techniques, but work in Bangla language is not increasing as expected due to limited resources and mentoring.

A. Future Works

Since our technique have many limitation, the proposed method can be extended with the following future works.

- We will try to implement the whole idea to make it faster and automated.
- Our algorithm can be integrated with various Machine learning algorithms such as Nave Bayes, Random Forest, and Support Vector Machine etc. to observe if the result become more accurate than the previous methods.
- New features can be integrated to get more accurate results.
- The proposed algorithm can be modified to differentiate funny sentences, hate speeches, angry sentences, and abusive sentences.
- An application can be developed to detect abusive texts when people browse various social sites using browsers or mobile app.
- The number of comments in the data set has to be increased to get a more accurate result.

REFERENCES

[1] R. Madhavan. (2018) Natural language processing current applications and future possibilities. [Online]. Available: https://www.techemergence.com/nlp-current-applications-and-future-possibilities/

[2] Wikipedia. (2010) List of languages by number of native speakers. [Online]. Available: https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers

[3] wikipedia. (2017) Bengali language. [Online]. Available: https://en.wikipedia.org/wiki/Bengali_language

[4] W. A. Social and Hootsuite. (2018) 2018 digital yearbook. [Online]. Available: https://digitalreport.wearesocial.com/

[5] P. KALLAS. (2018) Top 15 most popular social networking sites and apps [august 2018]. [Online]. Available: https://www.dreamgrow.com/top-15-most-popular-social-networking-sites/

[6] C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad, and Y. Chang, "Abusive language detection in online user content," in *Proceedings of the 25th International Conference on World Wide Web*, ser. WWW '16. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2016, pp. 145–153. [Online]. Available: https://doi.org/10.1145/2872427.2883062

[7] T. Davidson, D. Warmsley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," *arXiv preprint arXiv:1703.04009*, 2017.

[8] Y. Chen, Y. Zhou, S. Zhu, and H. Xu, "Detecting offensive language in social media to protect adolescent online safety," 09 2012, pp. 71–80.

[9] G. Xiang, B. Fan, L. Wang, J. Hong, and C. Rose, "Detecting offensive tweets via topical feature discovery over a large scale twitter corpus," in *Proceedings of the 21st ACM international conference on Information and knowledge management*. ACM, 2012, pp. 1980–1984.

[10] S. V. Wawre and S. N. Deshmukh, "Sentiment classification using machine learning techniques," *International Journal of Science and Research (IJSR)*, vol. 5, no. 4, pp. 819–821, 2016.

[11] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N Project Report, Stanford*, vol. 1, no. 12, 2009.

[12] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining." in *LREc*, vol. 10, no. 2010, 2010, pp. 1320–1326.

[13] D. Davidov, O. Tsur, and A. Rappoport, "Enhanced sentiment learning using twitter hashtags and smileys," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, ser. COLING '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 241–249. [Online]. Available: http://dl.acm.org/citation.cfm?id=1944566.1944594

[14] L. Barbosa and J. Feng, "Robust sentiment detection on twitter from biased and noisy data," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, ser. COLING '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 36–44. [Online]. Available: http://dl.acm.org/citation.cfm?id=1944566.1944571

[15] P. D. Turney, "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews," in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ser. ACL '02. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002, pp. 417–424. [Online]. Available: https://doi.org/10.3115/1073083.1073153

[16] D. Maynard and A. Funk, "Automatic detection of political opinions in tweets," in *Extended Semantic Web Conference*. Springer, 2011, pp. 88–99.

[17] M. M. Nabi, "Detecting sentiment from bangla text using machine learning technique and feature analysis," 2016.

[18] S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in bangla microblog posts," in *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*. IEEE, 2014, pp. 1–6.

[19] M. Al-Amin, M. S. Islam, and S. Das Uzzal, "Sentiment analysis of bengali comments with word2vec and sentiment information of words," 04 2017.

[20] M. G. Hussain, T. A. Mahmud, and W. Akthar, "An approach to detect abusive bangla text," in *2018 International Conference on Innovation in Engineering and Technology (ICIET)*, Dec 2018, pp. 1–5.

[21] S. C. Eshan and M. S. Hasan, "An application of machine learning to detect abusive bengali text," in *2017 20th International Conference of Computer and Information Technology (ICCIT)*, Dec 2017, pp. 1–6.

[22] Prothom alo - facebook home. [Online]. Available: https://www.facebook.com/DailyProthomAlo/

[23] Mashrafe bin mortaza - facebook home. [Online]. Available: https://www.facebook.com/Official.Mashrafe/

[24] Shakib al hasan - facebook home. [Online]. Available: https://www.facebook.com/Shakib.Al.Hasan/

[25] Salmon thebrownfish. [Online]. Available: https://www.youtube.com/user/salmanmuqtadir

[26] Naila nayem - facebook home. [Online]. Available: https://www.facebook.com/artist.nailanayem/

[27] Prothom alo - online news portal. [Online]. Available: https://www.prothomalo.com/

[28] wikipedia. (2018) Bag-of-words model. [Online]. Available: https://en.wikipedia.org/wiki/Bag-of-words_model
[29] Wikipedia. (2018) n-gram. [Online]. Available: https://en.wikipedia.org/wiki/N-gram

Md. Gulzar Hussain was born in Nimnagar, Dinajpur City, in 1990. He received the Bachelor of Science degree (2018) in Computer Science and Engineering from the Green University of Bangladesh, Dhaka, Bangladesh.

Since 2019, he has been a Lecturer with the Computer Science and Engineering Department, Green University of Bangladesh, Dhaka, Bangladesh. His research interests include Artificial Intelligence, Machine Learning, Natural Language Processing, Text Mining, Topic Modeling etc.

Mr. Hussain was a recipient of seven times Vice-Chancellor awards and once Dean award for excellent performance in trimester results during his bachelor study.

Tamim Al Mahmud was born in East Ratanpur village, Kazirhat, Barisal, Bangladesh, in 1990. He received his Bachelor Degree (2012) in Computer Science and Engineering from Patuakhali Science and Technology University, Patuakhali, Bangladesh and Master Degree (2016) in Information Technology from University of Dhaka, Dhaka, Bangladesh.

From 2013 to 2016, he was a Lecturer in Computer Science at Dhaka International University. Since 2017, he has been working as an Assistant Professor in the Department of Computer Science and Engineering, Green University of Bangladesh. His research interests include Machine Learning, Artificial Intelligence, Software Engineering and Human Computer Interaction . He is a Member of IEEE, Bangladesh Computer Society and Asian Business Consortium. He is also a member and team leader of HCI research group of Computer science Department at Green University of Bangladesh.