

Cyber bullying Detection & Prevention for Social Media Using Data Mining

^[1] Nangare Ravi B., ^[2] Chaitali R. Kamthe, ^[3] Awade Milind Shravan, ^[4] Kadam Aishwarya Mahesh
^[1] Assistant Professor RDTSC SCSCOE, Dhangawadi, Pune, ^{[2],[3],[4]} RDTSC- SCSCOE, Dhangawadi

Abstract- The increasing use of social communication networks by their users leads to huge amount of user-generated communication data. Due to the popularity of social media cyberbullying become the major problem in online communication and cyberbullying behavior received more and more attention. Cyberbullying may cause many serious and negative impacts on person's life and even leads to teen suicide. In the existing system the set of unique features derived from Twitter such as network, activity, user and tweet contents. By using these features the cyberbullying words which are presented in the comment contents are detected using data mining algorithms. The rumor comments are detected using syntactic and semantic techniques. The cyberbully detection and rumor detection on social network are done separately in the existing technique. In the proposed work the detection of cyberbully words and rumor comments on social media are integrated into a single application, along with these the cyberbully contents in the post. Comments will be detected using Pattern Matching algorithm.

Keywords— Cyber bullying, social-network, Cyber harassment, Text mining.

I. INTRODUCTION

Cyberbullying was defined by Patching and Hinduja as "willful and repeated harm inflicted through the medium of electronic text". According to the definition of the National Crime Prevention Council, cyberbullying is the use of the Internet, cell phones or other technologies to send or post a text or images intended to hurt or embarrass another person. Flooding It consists of the bully frequently sending the same comment, nonsense comments, or press the enter key in order to not allow the victim to contribute to the conversation. Masquerade involves the bully pretends to be someone who they are not. This would make it appear with the purpose of bully a victim directly. Flaming or bashing is a kind of online fight. The bully sending or posting electronics message which are enticingly insulting, vulgar to one or several persons either privately or publicly to an online group. Harassment is the kind of conversation that the bully frequently sends insulting and rude messages to the victim. Cyber stalking and cyber threats occur when the poster sends intimidating or offensive messages. Denigration also called "dissing" happens when an electronic bully sends or publishes gossip or untrue statement about a victim in order to damage the victim's friendship or reputation. Outing occurs when a person sends or publishes private or embarrassing information in a public chat room or forum. This type of cyberbullying is similar to the denigration. However in outing the relationship between bully and victim are close.

II. LITERATURE SURVEY

a) Detection of cyberbullying in messages Yin et al., conducted experiments on three different data sets (My

Space, Slashdot and Congregate) provided by Content analysis for Web 2.0 (CAW 2.0) in order to detect harassment. For harassment detection they used content, sentiment, and contextual features of the documents to train a support vector machine (SVM) classifier for a corpus of online posts. Various methods were used to develop the attributes of the entrance to the classifier, such as: standard text mining techniques based on weights of term (in this case - words), rule-based systems for detection of feelings and context analysis. The obtained results demonstrated that the use of the combined model, which besides text mining included methods for adding context and detection of feelings, improved the detection of cyberbullying.

B) Detection of cyberbullying on Twitter A framework for the detections of cyberbullying on Twitter was created by Sanchez and Kumar. Text that was used in messages (tweets, twitter message) requires intensive pre-processing prior to classification, including identification of syntax errors, emotions, and use of slang. The idea was to classify emotions contained in a message using a Sentiment analysis and opinion mining, and then to visualize the changes in the message over time. The messages were classified using Naive Bayes algorithm as negative or positive, with respect to some frequently used words. Bag-of- words model was used in the classification. The aim of the authors was to identify the victims.

C) Detection of cyberbullying in comments from YouTube video clips Dadvar et al. detected cyberbullying in comments from YouTube video. They used combination of content based, cyberbullying specific and user based features. They have shown that using user context (user's comments history and user characteristics) improves

cyberbullying detection accuracy. For training they used SVM binary classifier.

D) Detection of cyberbullying in social networks Nahar et al. proposed semi-supervised approach for detection in social networks, by devised new framework automatic detection of cyberbullying for streaming data with insufficient labeling. They conducted experiments on three different data sets (My Space, Slashdot and Congregate) provided by Content analysis for the enriched features sets were generated based on user context, linguistic knowledge and baseline keywords. They proposed fuzzy SVM algorithm for cyberbullying detection.

III. PROPOSED ARCHITECTURE

In the proposed architecture the process of detecting cyberbully activities begins with input dataset from social network. Input is text conversation collected from social Q&A website. Input is given to data pre-processing which is applied to improve the quality of the research data and subsequent analytical steps, this includes removing stop words, extra characters and hyperlinks. After performing pre-processing on the input data, it is given to Feature Extraction. Feature Extraction is done to obtain features like Noun, Adjective and Pronoun from the text and statistics on occurrence of word (frequency) in the text. The extracted features are given to Learning Algorithm. The Learning algorithm unit is the central element of the architecture and is composed of a genetic algorithm for modeling adaptive and exploratory behavior. Knowledge is given as Fuzzy rule set. The main functionality is to adjust the representation of the information needed for classification and yet retains the essential knowledge from the past. This knowledge is kept in a population of chromosomes, which is processed by the genetic algorithm. All the chromosomes in the population are competing to predict the classification of cyberbully activities. The output from learning unit is given to Classifier technique classifies the cyberbully activities using the fitness value of chromosome. The ability of a chromosome to classify the activity is called the fitness of the chromosome. The chromosome with higher fitness value gives the classified output. The output is classified bullying words present in the conversation.

IV. MATHEMATICAL MODEL

This consists of the total mathematical model gives idea about the analysis of the proposed system in terms of mathematical notations.

$$S = \{U, D, C, E, SA\}$$

Where, S = System.

U = User login system.

D = Data Center.

$$U = \{U1, U2, U3 \dots Un \dots\}$$

$$D = \{D1, D2, D3 \dots Dn \dots\}$$

C=Comment.

SA = Main System.

E= Eliminate.

SE = I, Q, T, F

SE = Server for operation.

I = Input (Data Comment.)

Q = Display Remove comment.

T=Task process.

When,

F= if negative comment then do not post

Else Post the comment

SA = Comment Uploading, Remove Comment, Identify Data Center, Task Assignment Task Processing

Success Conditions: As per user input desired output is generated

Failure Conditions: Desired output is not obtained

V. ALGORITHM

The Brute-Force Algorithm consists in comparing two strings of characters. This algorithm compares from left to right each word the user writes with each letter of the name of the file found inside of the route the user specifies. The process that this algorithm performs is the following [3]:

- Takes the character with which the pattern starts.
- Starts to compare it with each of the text characters, until the first match is found.
- It stops in said position and from there it starts to verify if the pattern matches with the rest of the text

Naive-String-Matcher (T,P)

$$n = T.length \quad m = P.length$$

$$\text{for } s = 0 \text{ to } n - m \quad \text{if } P[1..m] == T[s + 1..s + m]$$

print "Pattern occurs with shift" s

VI. RESULTS



Fig1:-User Login by enter the email-id and password

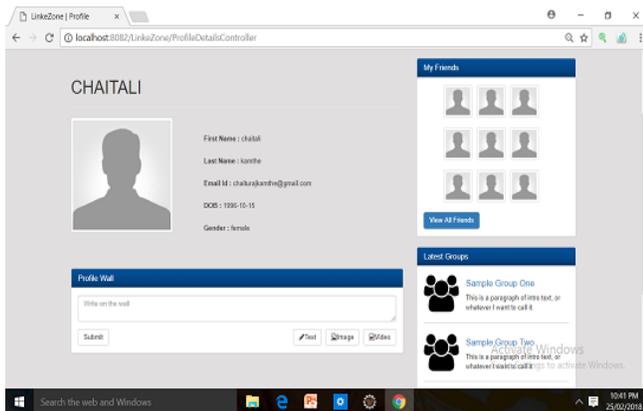


Fig2:-Login User Profile

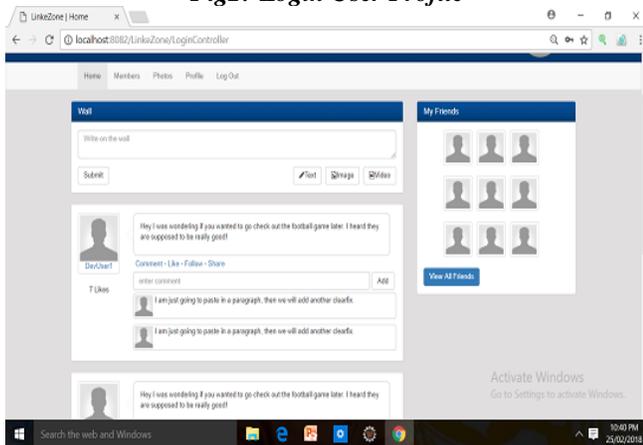


Fig3:-Comment entered on User Profile

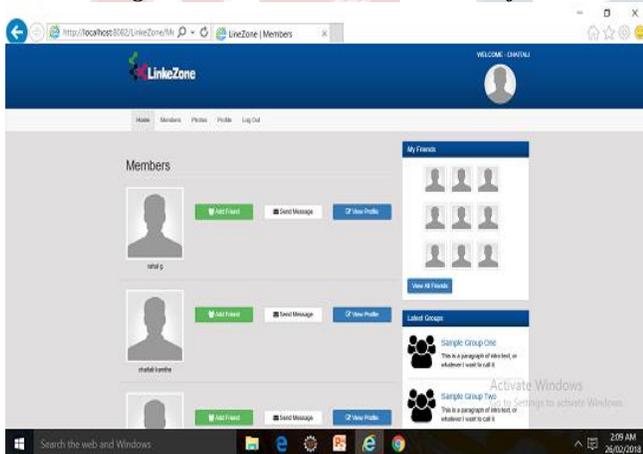


Fig4:-SMS send to User Friend

VII. CONCLUSION

In this paper we represented a survey on the current scenario of cyberbullying and various methods available for the detection and prevention of cyber harassment. Our concept depends upon the text analysis, the data which is uploaded or text written by any user is first analyzed and after that, we

estimate the roles of user, is it a bully? Or a victim? And then provide help as required by the user using data mining techniques. Also we will be using a User Identity for registration on our site i.e. one will have to provide an identity proof for registering on our site else they will not be able to make an account. With this feature we will be able to check the problem of fake accounts and also cyberbullying will be controlled to a limit as user accounts will be directly linked to their original identity. This mechanism will be very helpful for our society and the victims.

REFERENCES

- [1] cyberbullying.org/about-us/ (Accessed 26th August)
- [2] <http://www.endcyberbullying.org/india-ranks-third-on-global-cyberbullying-list/> (Accessed 28th August)
- [3] Chavan, V. S., & Shylaja, S. (2015). Machine learning approach for detection of cyber aggressive comments by peers on social media network. In Advances in computing, communications and informatics (ICACCI), 2015 International Conference on (pp.2354-2358 IEEE).
- [4] K.Dinakar, R.Reichart, H.Lieberman,"Modelling the Detection of Textual Cyberbullying." In: ICWSM 2011, Barcelona, Spain, July 17-21-2011.
- [5] Rui Zhao, Anna Zhou, Kezhi Mao, Automatic Detection of Cyberbullying on Social Networks based on Bullying Features, ICDCN '16 Article No. 43, January 2016, ACM.
- [6] <http://indianexpress.com/article/technology/technologyot-hers/alarmed-50-indian-youths-have-experienced-cyberbullying/> (Twitter", /JABM. Vol. I. No.1, April 2015