SSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 4, April 2018

Cyberbullying Detection based on Semantic-Enhanced Marginalized Denoising Auto-Encoder

[1] D. Swathi, [2] S. Babu
[1] M.Tech Student, [2] Assistant Professor
[1][2] Dept. Of Computer Science & Eng. V.R. Siddhartha Engineering College Kanuru, Vijayawada-07, AP. .

***Abstract*- Social Networking is a group of Internet based applications that allow the creation and exchange of user-generated content. Via social media, people can enjoy enormous information, convenient communication experience and so on. Since, social media may have some side effects such as cyberbullying, which may have negative impacts on the life of people, especially children and teenagers. Cyberbullying can be defined as aggressive, intentional actions performed by an individual or a group of people via digital communication methods such as sending messages and posting comments against a victim. Different from traditional bullying that usually occurs at school during face-to- face communication, cyberbullying on social media can take place anywhere at any time. Most of the individuals involved in these activities belong to the younger generations, especially teenagers, who in the worst scenario are at more risk of suicidal attempts. Here, we propose the technique for detection and avoidance of cyberbullying words in social media when cyberbullying takes place. And also proposing the technique for detection and blocking the accessing of a predator in social media.**

***Keywords*: Cyberbullying, Detection, Social Media, Traditional bullying, user-generated content.**

I. INTRODUCTION

Social networking sites have become immensely popular in the last few years. More than millions of users have used these websites as communication tools and as real-time, dynamic data sources. Where users can create their own profiles and communicate with other users regardless of location and physical limitations. Bullies are utilizing social media as a new platform for spreading rumors.

Cyberbullying is emerged as a major problem along with the recent development of online communication and social media. Cyberbullying actions are intentionally performed by an individual via social networking sites such as sending messages or posting comments against a victim. Cyberbullying is recognized as a serious problem, because the victims might face health issues like stress, depression and there is a high risk of suicidal attempts. Cyberbullying is different from traditional bullying that takes place any time. Social media provides users not only a good platform for communication and information sharing, but also an easy access to the current information. However, these platforms are also places where users experience bullying as victims, bullies or predators. The most common places where cyberbullying occurs are:

• Social Media, such as Facebook, Instagram, Snapchat, and Twitter.
• SMS (Short Message Service) also known as Text Message sent through devices.
• Instant Message (via devices, email provider services, apps, and social media messaging features).
• Email.

Cyberbullying Tactics:

Some of the cyberbullying actions that takes place over social media is:

• Posting victims personal information without his permission and spreading rumors about victim.
• Telling someone to kill themselves.
• Posting a hurtful picture or video.
• Posting hateful names, comments, or content about any race, religion, ethnicity, or other characteristics online.
• Creating a hurtful webpage about someone.

II. LITERATURE REVIEW

According to the authors Mohammed Ali Al-garadi et al. [1] applied Random Forest to detect cyberbullying. A systematic approach to develop Online Patrol by automatically spotting suspicious entries and reporting them to PTA members and therefore help them do their job. Random forest is an ensemble learning technique that builds multitudes of decision tress. A decision tree is a graphic method in which each branch node represents a choice between alternatives. A graphic approach is employed in decision trees to compare competing alternatives.

According to the authors Ying Chen et al. [2] applied Lexical Syntactic Feature (LSF) architecture to detect offensive content and identify potential offensive users in social media. The overall accuracy is 77.8% in user offensive detection. According to the authors Maral Dadvar, Dolf Trieschnigg, Roeland Ordelman & Franciska de Jong, [3] applied SVM to detect cyberbullying, and

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 4, April 2018

determined that incorporating user-based content improved the detection accuracy of SVM. Using data sets from MySpace, Dadvar et al. developed a gender-based cyberbullying detection approach that used the gender feature in enhancing the discrimination capacity of a classifier. Dadvar et al. and Ordelman et al. included age and gender as features in their approach; however, these features were limited to the information provided by users in their online profiles. Moreover, most studies determined that only a few users provided complete information about themselves in their online profiles. Alternatively, the tweet contents of these users were analyzed to determine their age and gender. The paper [4] proposes an alternative approach to cyberbullying: a system composed of multiple agents that control users' norm adherence within virtual societies. Being physically present in the virtual society, the agents continuously monitor the behavior of the visitors, communicate with each other to maintain shared beliefs of the visitors' characteristics, and apply punishments and rewards to influence their behavior. Computer software was developed to detect the presence of cyberbullying in online chat conversations.

In existing system even after detecting the cyberbully attack it wouldn't avoid and also there is no detection and avoidance of a predator.

III. METHDOLOGY

The main goal of this project is to detect the cyberbullying messages and avoiding the cyberbullying messages to be posted. This project also detects the predator or bully, and blocking the accessing of a predator. Here we are proposing N-grams algorithm to detect the cyberbullying messages. N-grams algorithm is considered to be an improved approach to detect cyberbullying messages. It is a contiguous sequence of 'n' items typically collected from a text or speech corpus. An n-gram of size 1 is referred to as a "unigram"; size 2 is a "bigram" (or, less commonly, a "digram"); size 3 is a "trigram". Bi-gram and Tri-gram are the most popular N-grams used in text mining. For example, for the sentence "The cow jumps over the moon". If N=1(known as unigrams), then the n-grams would be:

- The
- Cow
- Jumps
- Over
- The
- Moon

If N=2 (known as bigrams), then the n-grams would be:

- The cow
- Cow jumps
- jumps over
- Over the
- The moon

IV. EXPERIMENTAL RESULTS

The project mainly focuses on detecting cyberbullying messages and avoiding those messages and also detecting bullies and blocking the access of a bully. Initially, the user has to register by register as in an online social network, by entering his/her details. After registration is done, the users can login with their login credentials. The existing users can send messages privately and publicly, options are built. Users can also share post with others. The user can able to search the other user profiles and public posts. In this module users can also accept and send friend requests. If in case, the user posts information containing bullying words or foul language the post gets blocked and will not be displayed. The details of the bullies, who tries to post information containing bullying words are detected and sent to admin. The admin can block the access of a bully.

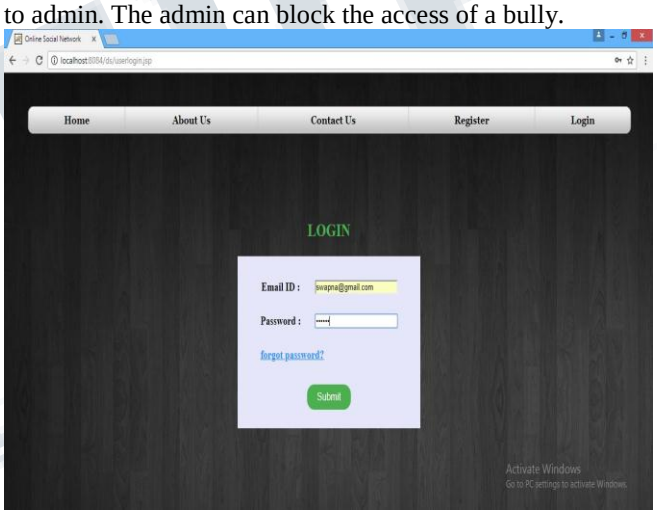

Fig 1 : User Login

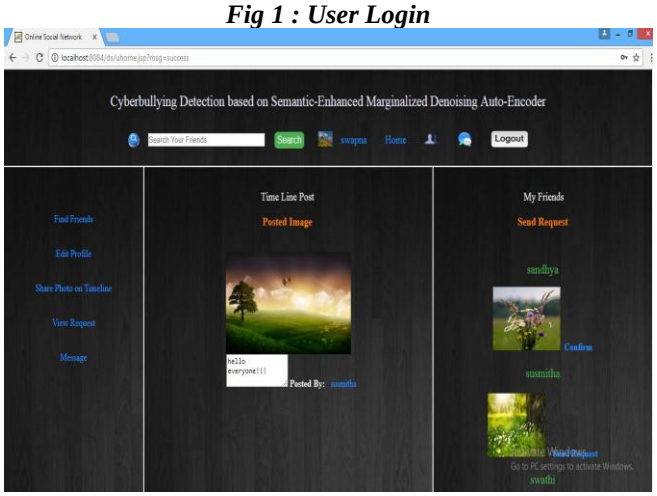

Fig 2: User Account

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 4, April 2018

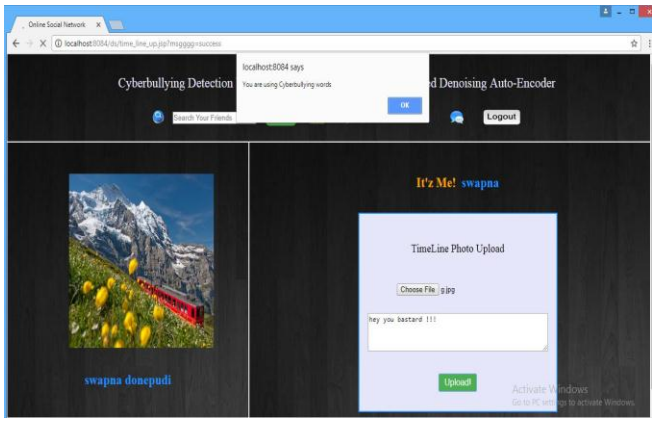

Fig 3: Detection of cyberbullying words

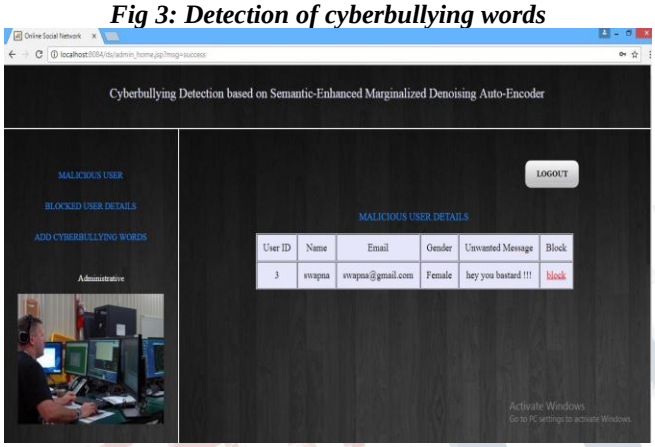

Fig 4: Detection of bully

V. CONCLUSION

This project mainly focus on the text-based cyberbullying detection problem, where robust and discriminative representations of messages are critical for an effective detection system. Study about effects of bullying is present but implementation for monitoring social network to detect cyberbullying activities is less. Hence, the proposed system focuses on detecting the presence of cyberbullying activity in social networks and avoiding the cyberbullying words to be posted in social networks. The proposed system also focuses on detecting the predator and blocking the accessing of a predator in social networks. This could help to construct a healthy and safe social media environment.

REFERENCES

[1] Mohammed Ali Al-garadi, Kasturi Dewi Varathan, Sri Devi Ravana, "Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network", 2016, Elsevier Journal, pp. 433-443.

[2] Ying Chen, Yilu Zhou, Sencun Zhu1, HengXu, "Detecting Offensive Language in Social Media to Protect Adolescent Online Safety", 2012.

[3] Maral Dadvar, Dolf Trieschnigg, Roeland Ordelman & Franciska de Tong, "Improving cyberbullying detection with user context", 2013, Springer, pp. 693-696.

[4] TiborBosse and Sven Stam, "A Normative Agent System to Prevent Cyberbullying", 2011, IEEE.

[5] Divyashree, Vinutha H, Deepashree N S, "An Effective Approach for Cyberbullying Detection and Avoidance", 2016, International Journal of Innovative Research in Computer and Communication Engineering, pp. 8005-8010.

[6] B. Sri Nandhini, J.I.Sheeba, "Online Social Network Bullying Detection Using Intelligence Techniques", 2015, Science Direct Journal, pp. 485-492.

[7] A. Saravanaraj, J.I. Sheeba, S. Pradeep Devaneyan, "Automatic Detection of Cyberbullying from Twitter", 2016, International Journal of Computer Science and Information Technology & Security, pp. 26-32.

[8] Ramzan Talib, Muhammad Kashif Hanif, Shaeela Ayesha, Fakeeha Fatima, "Text Mining: Techniques, Applications and Issues", 2016, International Journal of Advanced Computer Science and Applications, pp. 414-418.

[9] Roshan Jabee, M. Afshar Alam, "Issues and Challenges of Cyber Security for Social Networking Sites (Facebook)", 2016, International Journal of Computer Applications, pp. 36-40.

[10] Rita Dewanjee, Dr. R. Vyas, "Cyber Crime: Critical View", 2016, International Journal of Science and Research, pp. 85-87.

[11] Shabnoor Siddiqui, Tajinder Singh, "Social Media its Impact with Positive and Negative Aspects", 2016, International Journal of Computer Applications Technology and Research, pp. 71-75.

[12] Parvathy J, Suchithra R, "Impact of Usage of Social Networking Sites on Youth", 2015, International Journal of Computer Applications, pp. 33-34.

[13] Dr. S Kavitha, R Bhuvaneswari, "Impact of Social Media on Millennials- A Conceptual Study", 2016, Journal of Management Sciences and Technology, pp. 80-86.